PROCEEDINGS OF THE 14th STUDENT SESSION OF THE EUROPEAN SUMMER SCHOOL FOR LOGIC, LANGUAGE, AND INFORMATION

JULY 20-31, BORDEAUX, FRANCE

THOMAS ICARD (EDITOR)

Preface

The Student Session has been a part of the ESSLLI tradition every year since its inaugural session in 1996 in Prague, making this the fourteenth. The quality of papers in these proceedings is a testament to the fact that the Student Session continues to be an excellent venue for students to present original work in quite diverse areas of logic, language, and computation. This year's authors also comprise a very international group, with students from universities in the Netherlands, Germany, France, Spain, United Kingdom, United States, Japan, and India.

I would first like to thanks each of the co-chairs, as well as our distinguished area experts, for all of their help in the reviewing process and organizational matters in general. Thanks are also due to the long list of generous anonymous reviewers. Other individuals who have helped on numerous counts from FoLLI and ESSLLI Organizing Committee, and from previous Student Sessions, include Ville Nurmi, Kata Balogh, Christian Retoré, Richard Moot, Uwe Mönnich, and especially Paul Dekker and Sophia Katrenko. As in previous years, Springer-Verlag has graciously offered prizes for 'Best Paper' awards, and for this we are very grateful. Finally, thanks to all those who submitted papers, as they are what really make the Student Session such an exciting event.

Thomas Icard
Chair of the 2009 Student Session

Stanford, June 2009

Program Committee

Chair

Thomas Icard (Stanford University)

Area Co-Chairs

Logic and Language
Salvador Mascarenhas (New York University)

Language and Computation
Bruno Mery (University of Bordeaux)

Logic and Computation
Marija Slavkovik (University of Luxembourg)

Area Experts

Logic and Language
Nathan Klinedinst (University College London)
Reinhard Muskens (Tilburg University)

Language and Computation
Makoto Kanazawa (National Institute of Informatics, Japan)
Jens Michaelis (University of Bielefeld)

Logic and Computation
Arnon Avron (Tel-Aviv University)
Alexandru Baltag (Oxford University)

Table of Contents

Description Logics for Relative Terminologies

or

Why The Biggest City Is Not a Big Thing

Szymon Klarman
Department of Computer Science, Vrije Universiteit Amsterdam

Abstract. Context-sensitivity has been for long a subject of study in linguistics, logic and computer science. Recently the problem of representing and reasoning with contextual knowledge has been brought up in the research on the Semantic Web. In this paper we introduce a conservative extension to Description Logic, the formalism underlying Web Ontology Languages, supporting representation of ontologies containing *relative terms*, such as 'big' or 'tall', whose meaning depends on the selection of *comparison class* (context). The solution rests on introduction of modal-like operators in the language and an additional modal dimension in the semantics, which is built upon the standard object dimension of the Description Logic languages and whose states correspond to selected subsets of the object domain. We present the syntax and semantics of the extended language and elaborate on its representational and computational features.

1. Introduction

It is a commonplace observation that the same expressions might have different meanings when used in different contexts. A trivial example might be that of the concept The_Biggest. Figure 1 presents three snapshots of the same knowledge base that focus on different parts of the domain. The extension of the concept visibly varies across the three takes. Intuitively, there seem to be no contradiction in the fact that individual London is an instance of The_Biggest, when considered in the context of European cities,

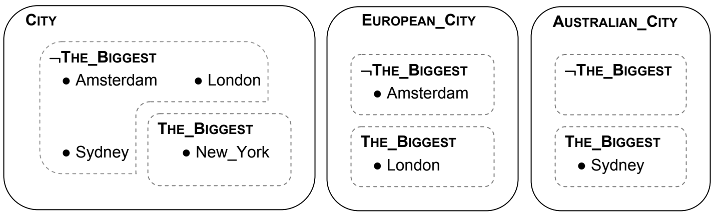


Figure 1: Example of a context-sensitive concept The_Biggest.

an instance of ¬The_Biggest, when contrasted with all cities, and finally, not belonging to any of these when the focus is only on Australian cities. Natural language users resolve such superficial incoherencies simply by recognizing that certain terms, call them *relative*, such as The_Biggest, acquire definite meanings only when put in the context of specified *comparison classes* (Shapiro 2006, van Rooij to appear, Gaio 2008).

The problem of context-sensitivity has been for a long time a subject of studies in linguistics, logic and even computer science. Recently, it has been also encountered in the research on the Semantic Web (Bouquet, et al. 2003, Caliusco, et al. 2005, Benslimane, et al. 2006) where the need for representing and reasoning with imperfect information becomes ever more pressing (Lukasiewicz & Straccia 2008, Laskey, et al. 2008). Relativity of meaning appears as one of common types of such imperfection. Alas, *Description*

Logics (DLs), which form the foundation of the Web Ontology Language (Horrocks, et al. 2003), the basic knowledge representation formalism on the Semantic Web, were originally developed for modeling crisp, static and unambiguous knowledge, and as such, are incapable of handling the task seamlessly. Consequently, it has become clear that it is necessary to look for more expressive, ideally backward compatible languages to meet the new application requirements on the Semantic Web. Current proposals focus mostly on the problems of uncertainty and vagueness (Lukasiewicz & Straccia 2008, Straccia 2005), with several preliminary attempts of dealing with different aspects of contextualization of DL knowledge bases (Grossi 2007, Goczyla, et al. 2007, Benslimane et al. 2006). In this paper we propose a simple, conservative extension to the classical DLs, which is intended for representation of relative, context-sensitive terminologies, where by *contexts* we understand specifically the *comparison classes* with respect to which the terms acquire precise meanings.

To take a closer look at the problem consider again the scenario from Figure 1. On a quick analysis it should become apparent there is no straightforward way of modeling the scenario within the standard DL paradigm. Asserting both $\mathsf{London} : \mathsf{The\_Biggest}$ and $\mathsf{London} : \neg\mathsf{The\_Biggest}$ in the same knowledge base results in an immediate contradiction, which is obviously an unintended outcome. To avoid this consequence one can resort to the luring prospect of indexing, and instead assert $\mathsf{London} : \mathsf{The\_Biggest}_{\mathsf{European\_City}}$ and $\mathsf{London} : \neg\mathsf{The\_Biggest}_{\mathsf{City}}$, with an implicit message that the two indexed concepts are meant to be two different 'variants' of $\mathsf{The\_Biggest}$, corresponding to two possible contexts of its use. The contradiction is indeed avoided, but unfortunately the baby has been thrown out with the bath water, for the two 'variants' become in fact two unrelated concept names, with no common syntactic or semantic core. More precisely, using this strategy one cannot impose global constraints on the contextualized concepts, for instance, to declare that regardless of the context, $\mathsf{The\_Biggest}$ is always a subclass of Big. Even if this goal was achieved by rewriting constraints over all individual contexts, another source of problems is reasoning about the contexts themselves, for example, deciding whether an individual occurs in a given comparison class or not.

The extension proposed in this paper is, to our knowledge, unique in addressing this particular type of context-sensitivity in DL, and arguably, it cannot be simulated within any of the approaches present in the literature. Technically, the solution rests on the presence of special modal-like operators in the language and a second modal dimension in the semantics of the language, which is defined over the standard object dimension of DL and whose states correspond to selected subsets of the object domain. In the following section we formally define the language, next we elaborate on some of its basic representational and computational features, and finally, in the last two sections, we shortly position our work in a broader perspective and conclude the presentation.

2. Representation Language

We start be recalling preliminary notions regarding DLs and follow up with presentation of the syntax and the semantics of the proposed extension, for brevity denoted as $\mathrm{DL}^{\mathcal{C}}$.

2.1. Basic Description Logics

Description Logics are a family of knowledge representation formalisms, designed particularly for expressing terminological and factual knowledge about a domain of application (Baader, et al. 2003). For instance, the following DL formula defines the meaning of the concept European city by equating it to the set of all and only those individuals that are cities and are located in Europe; the next one asserts that New York is not in fact an instance of that concept:

$$\mathsf{European\_City} \equiv \mathsf{City} \sqcap \exists \mathsf{located\_in}.\{\mathsf{Europe}\}$$
$$\mathsf{New\_York} : \neg \mathsf{European\_City}$$

Formally DLs can be seen as fragments of modal and first-order logic, with an intuitively appealing syntax (Schild 1991). A DL language $\mathcal{L}$ is specified by the vocabulary $\Sigma = (N_I, N_C, N_R)$, consisting of a set of *individual names* N_I, *concept names* N_C, and *role names* N_R, and by a selection of logical operators Π (e.g.: $\sqsubseteq$, $\sqcap$, $\sqcup$, $\neg$, $\exists$, $\forall$, $\ldots$), which allow for constructing complex expressions: concept descriptions, complex roles and axioms. A knowledge base $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, expressed in $\mathcal{L}$, consists of the TBox $\mathcal{T}$, containing terminological constraints, (typically) of the form of inclusion $C \sqsubseteq D$ or equivalence axioms $C \equiv D$, for arbitrary concept descriptions C and D, and the ABox $\mathcal{A}$, including concept assertions $a : C$ and role assertions $(a, b) : r$, for individual names a, b, a concept C and a role r.

The semantics is defined in terms of an *interpretation* $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a non-empty *domain* of individuals, and $\cdot^{\mathcal{I}}$ is an *interpretation function*, which specifies the meaning of the vocabulary by mapping every $a \in N_I$ to an element of $\Delta^{\mathcal{I}}$, every $C \in N_C$ to a subset of $\Delta^{\mathcal{I}}$ and every $r \in N_R$ to a subset of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$. The function is inductively extended over complex terms in a usual way, according to the fixed semantics of the logical operators. An interpretation $\mathcal{I}$ *satisfies* an axiom in either of the following cases:

- $\mathcal{I} \models C \sqsubseteq D$ iff $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$
- $\mathcal{I} \models a : C$ iff $a^{\mathcal{I}} \in C^{\mathcal{I}}$
- $\mathcal{I} \models (a, b) : r$ iff $\langle a^{\mathcal{I}}, b^{\mathcal{I}} \rangle \in r^{\mathcal{I}}$

Finally, $\mathcal{I}$ is said to be a *model* of a DL knowledge base, i.e. it makes the knowledge base true, if and only if it satisfies all its axioms.

2.2. Syntax

We consider an arbitrary DL language $\mathcal{L}$ and extend it with special operators for expressing contextualized concept descriptions, roles and axioms.

Let C and D be concept descriptions in $\mathcal{L}$ and r a role. Then the following are proper concept and role descriptions, respectively:

$$\begin{aligned} C, D \rightarrow &\quad \Diamond C \mid \Box C \mid \langle D \rangle C \mid [D]C \\ r \rightarrow &\quad \Diamond r \mid \Box r \mid \langle D \rangle r \mid [D]r \end{aligned}$$

The modal operators give access to the meaning of the bound terms in *specific subcontexts* of the current context of representation. The diamonds point at *certain* subcontexts, which might be either *anonymous* ($\Diamond$-modality) or *named* by the qualifying concepts placed inside of the operators ($\langle\cdot\rangle$-modality). In the former case the designated comparison class is unspecified, whereas in the latter it consists of all and only those individuals that, in the current context, are instances of the qualifying concept. Analogically, the dual box operators refer to *all* anonymous ($\Box$-modality), or named subcontexts ($[\cdot]$-modality) of the current context.

For instance, $\langle\mathsf{City}\rangle\mathsf{The\_Biggest}$ describes the individuals that are the biggest as considered in the context of cities, $\Box\neg\mathsf{The\_Biggest}$ refers to the individuals that are never $\mathsf{The\_Biggest}$ regardless of the considered subcontext, while $\exists\langle\mathsf{City}\rangle\mathsf{nicer}.\mathsf{European\_City}$ uses the meaning of the role nicer as interpreted when talking about cities, and denotes those individuals which in this sense are nicer than some European cities.

In a similar manner we allow for contextualization of DL axioms. If α is a TBox/ABox axiom and D is a concept description, then the following are also proper TBox/ABox axioms in DL$^{\mathcal{C}}$, respectively:

$$D, \alpha \rightarrow \quad \Diamond\alpha \mid \Box\alpha \mid \langle D\rangle\alpha \mid [D]\alpha$$

For example, $\langle\mathsf{Australian\_City}\rangle(\mathsf{Sidney} : \mathsf{The\_Biggest})$ asserts that there exists a context, namely that of Australian cities, in which Sidney is considered the biggest. The TBox axiom $\Box(\mathsf{The\_Biggest} \sqsubseteq \mathsf{Big})$ enforces that regardless of the comparison class the concept $\mathsf{The\_Biggest}$ is always subsumed by Big.

2.3. Semantics

The central semantic notion underlying DL$^{\mathcal{C}}$ is *context structure*, which can be seen as a special type of an interpretation of a *multi-dimensional DL* (Wolter & Zakharyaschev 1999b).

Definition 1 (Context structure) *A* context structure *over a DL$^{\mathcal{C}}$ language $\mathcal{L}$, with a set of operators Π and the vocabulary $\Sigma = (N_I, N_C, N_R)$, is a tuple $\mathcal{C} = \langle W, \lhd, \Delta, \{\mathcal{I}_w\}_{w\in W}\rangle$, where:*

- *$W \subseteq \wp(\Delta)$ is a set of possible contexts, such that $\Delta \in W$ and $\emptyset \notin W$;*
- *$\lhd \subseteq W \times W$ is an accessibility relation, such that for any $w, v \in W$, $w \lhd v$ if and only if $v \subseteq w$. In such cases we say that v is a* subcontext *of w;*
- *Δ is a non-empty domain of interpretation;*
- *$\mathcal{I}_w = (\Delta^{\mathcal{I}_w}, \cdot^{\mathcal{I}_w})$ is a (partial) interpretation of $\mathcal{L}$ in the context w:*
  - *$\Delta^{\mathcal{I}_w} = w$ is the domain of interpretation in w,*
  - *$\cdot^{\mathcal{I}_w}$ is a standard interpretation function for language $\mathcal{L}_w$, defined by Π and a subset $\Sigma_w \subseteq \Sigma$ of the vocabulary of $\mathcal{L}$.*

Note that the contexts are uniquely identifiable by their corresponding domains of interpretation and are ordered by $\lhd$ according to the decreasing size of the domains, i.e. for every context structure and every $w, v \in W$ the following conditions hold:

- $w \lhd v$ iff $\Delta^{\mathcal{I}_v} \subseteq \Delta^{\mathcal{I}_w}$,
- if $\Delta^{\mathcal{I}_w} = \Delta^{\mathcal{I}_v}$ then $w = v$.

Finally, we observe there exists a special element $\hat{w} \in W$, denoted as the *top context*, such that $\Delta^{\mathcal{I}_{\hat{w}}} = \Delta$. Given the conditions above, it follows that $\lhd$ imposes a partial order (reflexive, asymmetric and transitive) on the set of contexts, with $\hat{w}$ as its least element. Thus context structures are built upon rooted partially ordered Kripke frames.

For an arbitrary context structure $\mathcal{C}$, a context w, concept descriptions C, D and a role r, the meaning of contextualized terms is inductively defined as follows:

$$
\begin{aligned}
(\Diamond C)^{\mathcal{I}_w} &= \{x \in \Delta^{\mathcal{I}_w} \mid \exists w \lhd v:\ x \in C^{\mathcal{I}_v}\}\\
(\Box C)^{\mathcal{I}_w} &= \{x \in \Delta^{\mathcal{I}_w} \mid \forall w \lhd v:\ x \in \Delta^{\mathcal{I}_v} \rightarrow x \in C^{\mathcal{I}_v}\}\\
(\Diamond r)^{\mathcal{I}_w} &= \{\langle x, y\rangle \in \Delta^{\mathcal{I}_w} \times \Delta^{\mathcal{I}_w} \mid \exists w \lhd v:\ \langle x, y\rangle \in r^{\mathcal{I}_v}\}\\
(\Box r)^{\mathcal{I}_w} &= \{\langle x, y\rangle \in \Delta^{\mathcal{I}_w} \times \Delta^{\mathcal{I}_w} \mid \forall w \lhd v:\ x, y \in \Delta^{\mathcal{I}_v} \rightarrow \langle x, y\rangle \in r^{\mathcal{I}_v}\}
\end{aligned}
$$

Terms contextualized via named contexts are interpreted analogically, except for an extra restriction imposed on the accessibility relation: only the subcontexts that match the extension of the qualifying concept are to be considered.

$$
\begin{aligned}
(\langle D\rangle C)^{\mathcal{I}_w} &= \{x \in \Delta^{\mathcal{I}_w} \mid \exists w \lhd v,\ \Delta^{\mathcal{I}_v} = D^{\mathcal{I}_w}:\ x \in C^{\mathcal{I}_v}\}\\
([D]C)^{\mathcal{I}_w} &= \{x \in \Delta^{\mathcal{I}_w} \mid \forall w \lhd v,\ \Delta^{\mathcal{I}_v} = D^{\mathcal{I}_w}: x \in \Delta^{\mathcal{I}_v} \rightarrow x \in C^{\mathcal{I}_v}\}\\
(\langle D\rangle r)^{\mathcal{I}_w} &= \{\langle x, y\rangle \in \Delta^{\mathcal{I}_w} \times \Delta^{\mathcal{I}_w} \mid \exists w \lhd v,\ \Delta^{\mathcal{I}_v} = D^{\mathcal{I}_w}:\ \langle x, y\rangle \in r^{\mathcal{I}_v}\}\\
([D]r)^{\mathcal{I}_w} &= \{\langle x, y\rangle \in \Delta^{\mathcal{I}_w} \times \Delta^{\mathcal{I}_w} \mid \forall w \lhd v,\ \Delta^{\mathcal{I}_v} = D^{\mathcal{I}_w}: x, y \in \Delta^{\mathcal{I}_v} \rightarrow \langle x, y\rangle \in r^{\mathcal{I}_v}\}
\end{aligned}
$$

Noticeably, the modalities involving named contexts nearly collapse, as there can always be only one such subcontext that matches the qualifying concept. Thus the inclusion $\Box(\langle D\rangle C \sqsubseteq [D]C)$ is valid in DL$^{\mathcal{C}}$, although its converse $\Box([D]C \sqsubseteq \langle D\rangle C)$ is not, as there might be individuals that are instances of $[D]C$ simply because they do not exist in the subcontext designated by D. In fact, it is easy to prove that for any C, D and r the following equivalences hold: $[D]C = \neg D \sqcup \langle D\rangle C$ and $\langle D\rangle C = [D]C \sqcap D$.

3. Reasoning with relative terminologies

In this section we define the problem of satisfiability and discuss some issues concerning computational properties of DL$^{\mathcal{C}}$. Next, we present two examples embedded in a decidable subset of the language.

3.1. Satisfiability and computational properties

As for most formalisms within the DL paradigm, the basic reasoning service being of interest for DL$^{\mathcal{C}}$ is *satisfiability checking*. The notion of *satisfaction* of an axiom is relativized here to the context structure and a particular context. For a context structure $\mathcal{C}$, a context w, and a TBox/ABox axiom α, we say that α is *satisfied* in $\mathcal{C}$ in the context w, or shortly $\mathcal{C}, w \models \alpha$ iff $\mathcal{I}_w \models \alpha$. Consequently, satisfaction of contextualized axioms conservatively extends the definition of satisfaction used in the basic DLs:

- $\mathcal{I}_w \models \Diamond\alpha$ iff $\exists w \lhd v:\ \mathcal{I}_v \models \alpha$

- $\mathcal{I}_w \models \Box\alpha$ iff $\forall w \lhd v : \mathcal{I}_v \models \alpha$
- $\mathcal{I}_w \models \langle D\rangle\alpha$ iff $\exists w \lhd v,\ \Delta^{\mathcal{I}_v} = D^{\mathcal{I}_w} : \mathcal{I}_v \models \alpha$
- $\mathcal{I}_w \models [D]\alpha$ iff $\forall w \lhd v,\ \Delta^{\mathcal{I}_v} = D^{\mathcal{I}_w} : \mathcal{I}_v \models \alpha$

We say that a knowledge base is satisfied in a context w whenever all its axioms are satisfied in w. Finally, a context structure $\mathcal{C}$ with the top context $\hat{w}$ is a *model* of a knowledge base when all axioms in the knowledge base are satisfied in $\hat{w}$. Considering the satisfiability conditions and the formal properties of the underlying Kripke frames, we strongly suspect that decidability of the satisfaction problem, and consequently of other standard reasoning problems in DL$^{\mathcal{C}}$ (e.g. subsumption, instance checking, etc. (Baader et al. 2003)), should be preserved. In the next section we will discuss a syntactic restriction of the language whose decidability we show by a simple argument.

As an interesting consequence of the formulation of the framework, we are able to define the notions of *global* (context-independent) and *local* (context-dependent) terms.

Definition 2 (Globality/locality) *A DL term τ is* global *in a context structure $\mathcal{C}$* iff *for every $w, v \in W$ such that $w \lhd v$ it holds that:*

- *if τ is an individual name a then $a^{\mathcal{I}_v} = a^{\mathcal{I}_w}$* iff *$a^{\mathcal{I}_w} \in \Delta^{\mathcal{I}_v}$, else $a^{\mathcal{I}_v}$ is unspecified,*
- *if τ is a concept description C then $C^{\mathcal{I}_v} = C^{\mathcal{I}_w} \cap \Delta^{\mathcal{I}_v}$,*
- *if τ is a role r then $r^{\mathcal{I}_v} = r^{\mathcal{I}_w} \cap \Delta^{\mathcal{I}_v} \times \Delta^{\mathcal{I}_v}$,*

Otherwise, τ is local *in $\mathcal{C}$.*

Notably, the dichotomy of *global* vs. *local* terms, in the above sense, follows the distinction between *rigid* and *non-rigid* designators, as they are often denoted in the philosophy of language. Rigid designators are terms which designate the same things in all possible worlds in which those things exist, and do not designate anything in the remaining worlds. Non-rigid designators are exactly those terms which fail to satisfy the same condition. A suitable and explicit selection of assumptions regarding globality/locality of the employed vocabulary is of a great importance from the perspective of reasoning with relative terminologies. On the one hand, the rules of the reasoning calculus should be properly aligned with the modeling intentions with respect to which parts of the represented terminology are actually context-dependent and which are to be interpreted rigidly.[1] On the other one, the choice of assumptions is known to directly affect computational properties of the resulting models, i.e. decidability and complexity of reasoning.[2]

3.2. Reasoning in $\mathcal{ALC}^{\mathcal{C}}$ – examples

Let us finally present two small examples of (in)valid inferences in the DL$^{\mathcal{C}}$. To this end we will first specify a small, yet still sufficiently expressive subset of DL$^{\mathcal{C}}$, which can be easily shown to be decidable. As the basis we will use the DL $\mathcal{ALC}$, whose concept constructors and their semantics are presented in Table 1. We extend $\mathcal{ALC}$ to $\mathcal{ALC}^{\mathcal{C}}$ by posing the following requirements:

[1] E.g. in typical applications ambiguity of individual names is not considered an issue, hence it is natural to impose their rigid interpretation on the level of inference rules.

[2] Some authors report, for instance, that presence of global roles dramatically increases the complexity of the decision problem (Wolter & Zakharyaschev 1999b, Wolter & Zakharyaschev 1999a).

$$\begin{aligned}
(\neg C)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}} \\
(C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
(C \sqcup D)^{\mathcal{I}} &= C^{\mathcal{I}} \cup D^{\mathcal{I}} \\
(\exists r.C)^{\mathcal{I}} &= \{x \mid \exists_y(\langle x, y\rangle \in r^{\mathcal{I}} \wedge y \in C^{\mathcal{I}})\} \\
(\forall r.C)^{\mathcal{I}} &= \{x \mid \forall_y(\langle x, y\rangle \in r^{\mathcal{I}} \rightarrow y \in C^{\mathcal{I}})\}
\end{aligned}$$

Table 1: Concept constructors in the DL $\mathcal{ALC}$

1. Every axiom in $\mathcal{ALC}$ is a proper axiom in $\mathcal{ALC}^{\mathcal{C}}$.
2. If C, D are concept descriptions in $\mathcal{ALC}$, then $\langle D\rangle C$ and $[D]C$ are proper concept descriptions in $\mathcal{ALC}^{\mathcal{C}}$.
3. If α is a TBox axiom in $\mathcal{ALC}$, then $\Box\alpha$ is a proper axiom in $\mathcal{ALC}^{\mathcal{C}}$.
4. No other expressions are allowed in $\mathcal{ALC}^{\mathcal{C}}$.
5. (Only) individual names are interpreted rigidly in $\mathcal{ALC}^{\mathcal{C}}$.

The resulting language is decidable. First, observe that we strictly separate TBox axioms containing modalized concept descriptions from the ones bound by the $\Box$ operator. Moreover, note that axioms of the former type might contain only a finite number of $\langle\cdot\rangle$ and $[\cdot]$ operators (while no $\Diamond$ nor $\Box$), where each occurrence of $[D]C$ can be replaced by $\neg D \sqcup \langle C\rangle$ (see Section 2.3). Similarly, the ABox can be reduced to the form in which there is only a finite number of occurrences of $\langle\cdot\rangle$ and no other non-$\mathcal{ALC}$ constructs. Consequently, since every $\langle\cdot\rangle$ uniquely determines the accessible subcontext, it follows that every satisfiable DL$^{\mathcal{C}}$ knowledge base has to be satisfied in a model based on a finite context structure. Thus the problem of checking satisfiability of a DL$^{\mathcal{C}}$ knowledge base can be reduced to the problem of checking satisfiability of a finite number of $\mathcal{ALC}$ knowledge bases, which is of course decidable.[3]

Example 1 (The biggest city is not a big thing) *Consider the scenario presented in Figure 1 from the introductory section. Let us assert that New York is indeed the biggest city and further assume that in every possible context the concept* $\mathsf{The\_Biggest}$ *is subsumed by* Big*:*

1. $\mathcal{A} = \{\ \mathsf{New\_York} : \langle\mathsf{City}\rangle\mathsf{The\_Biggest}\ \}$
2. $\mathcal{T} = \{\ \Box(\mathsf{The\_Biggest} \sqsubseteq \mathsf{Big})\ \}$

Given no additional knowledge the following statement does not follow:

3. $\mathsf{New\_York} : \mathsf{Big}$

Since New York is the biggest in the context of cities (1) it must be also big in the same context (2). Nevertheless, the interpretation of Big in the context of cities is independent

[3] In (Klarman & Schlobach 2009) we present a tableau-based decision procedure for the same language without TBox axioms bound by the $\Box$ operator. Given the finiteness of the involved context structures, however, presence of these constructs obviously does not affect the complexity of reasoning and can be straightforwardly covered in the calculus.

from its interpretation in other contexts, in particular in the top context, in which our query (3) should be satisfied. Hence New York does not have to be an instance of Big in the top context. As an illustration consider the following canonical model invalidating the inference (x is any object different from New York):

$$
\begin{aligned}
W &= \{\hat{w}, w\} \\
\lhd &= \{\langle \hat{w}, w\rangle\} \\
\Delta &= \{\mathsf{New\_York}, \mathsf{x}\} \\
\hat{w} = \Delta^{\mathcal{I}_{\hat{w}}} &= \{\mathsf{New\_York}, \mathsf{x}\} \\
w = \Delta^{\mathcal{I}_{w}} &= \{\mathsf{New\_York}\} \\
\mathsf{City}^{\mathcal{I}_{\hat{w}}} &= \{\mathsf{New\_York}\} \\
\mathsf{The\_Biggest}^{\mathcal{I}_{\hat{w}}} = \mathsf{Big}^{\mathcal{I}_{\hat{w}}} &= \emptyset \\
\mathsf{The\_Biggest}^{\mathcal{I}_{w}} = \mathsf{Big}^{\mathcal{I}_{w}} &= \{\mathsf{New\_York}\}
\end{aligned}
$$

Example 2 (Are tall men tall people?) *Consider a simple terminology defining a person as a man or a woman, where the last two are disjoint concepts. Further, we assume the concepts* Tall *and* Short *are globally disjoint, and assert that individual John is tall as compared to men.*

1. $\mathcal{T} = \{\quad \mathsf{Person} \equiv \mathsf{Man} \sqcup \mathsf{Woman}$
2. $\qquad\quad \mathsf{Man} \sqcap \mathsf{Woman} \sqsubseteq \bot$
3. $\qquad\quad \Box(\mathsf{Tall} \sqcap \mathsf{Short} \sqsubseteq \bot)\ \}$
4. $\mathcal{A} = \{\quad \mathsf{John} : \langle \mathsf{Man} \rangle \mathsf{Tall}\ \}$

The following assertion is entailed by the knowledge base:

5. $\mathsf{John} : \langle \mathsf{Person} \sqcap \neg \mathsf{Woman} \rangle \neg \mathsf{Short}$

Notice that since the concept $\mathsf{Person} \sqcap \neg\mathsf{Woman}$ is equivalent to Man in the top context (1,2) then obviously both of them designate exactly the same context. Since John is tall in that context (4), and in every context tall objects cannot be short at the same time (3), it follows naturally, that John is in that context also an instance of $\neg\mathsf{Short}$. Observe, however, that it cannot be inferred that John is a non-short person, as nothing is known about the tallness of John in the context of all people.

4. Related work

The language $\mathrm{DL}^{\mathcal{C}}$ can be classified as an instance of *modal* or *multi-dimensional* (Wolter & Zakharyaschev 1999b, Wolter & Zakharyaschev 1999a) DLs, a family of expressive description languages being a fragment of *multi-dimensional modal logics* (Marx & Venema 1997, Kurucz, et al. 2003). To our knowledge, $\mathrm{DL}^{\mathcal{C}}$ constitutes a unique proposal explicitly employing this framework for the problem of contextualization of DL knowledge bases, and moreover, it is the only attempt of addressing the specific problem of relativity, i.e. contextualization of DL constructs by comparison classes. The most commonly considered perspectives on contextualization in DLs focus instead on:

1. *Integration of knowledge bases describing local views on a domain* (Bouquet et al. 2003, Benslimane et al. 2006, Borgida & Serafini 2003). In this perspective, one considers a finite set of DL knowledge bases related by bridge rules, of a certain form, which allow for relating concepts belonging to different sources.

2. *Contextualization as levels of abstraction* (Goczyla et al. 2007, Grossi 2007). Within this approach, a knowledge base is modeled as a hierarchical structure organized according to the levels of abstraction over the represented knowledge. The entailments of the higher levels hold in the more specific ones, but not vice versa.

Although the underlying models can be in both cases embedded in two-dimensional Kripke semantics, analogical to ours, the expressive power of the modal machinery is not utilized on the level of language, and thus the formalisms remain strictly less expressive than DL$^{\mathcal{C}}$. It can be expected, however, that restricted fragments of DL$^{\mathcal{C}}$ and enriched variants of the other approaches to contextualization might coincide on some problems in terms of expressiveness and semantical compatibility.

DL$^{\mathcal{C}}$ shares also some significant similarities with *dynamic epistemic logics*, in particular, with the *public announcement logic* (PAL) (van Ditmarsch, et al. 2007), which studies the dynamics of information flow in epistemic models. Interestingly, our modalities involving named contexts can be seen as public announcements, in the sense used in PAL, whose application results in a dynamic reduction of the description (epistemic) model to only those individuals (epistemic states) that satisfy given description (formula). Unlike in PAL, however, we allow for much deeper revisions of the models, involving also the interpretation function, e.g. it is possible that after contextualizing the representation by $\langle C\rangle$ there are no individuals that are C, simply because C gets essentially reinterpreted in the accessed world.

Finally, we should mention the loosely related problem of vagueness, inherent to the use of relative terms, such as considered in this paper. Traditionally, the problem has been analyzed on the grounds of *fuzzy logics*, which recently have been also successfully coupled with description languages, giving raise to *fuzzy DLs* (Straccia 2005). The ideas underlying fuzzy semantics, however, are orthogonal to the ones motivating our work, and thus can in principle complement each other. While fuzzy logic replaces a binary truth function with a continues one when defining an interpretation of a relative term, the semantics of DL$^{\mathcal{C}}$ allows for applying a number of truth functions instead of a single one, depending on the context of interpretation. Clearly, none of the two semantics can solve the problems handled by the other, while together they give a very broad account of the problem of relativity of meaning.

5. Conclusions

Motivated by the current challenges of the research on the Semantic Web, we have presented an extension to the standard Description Logics for representing and reasoning with relative terms, i.e. terms whose precise meaning depends on the choice of a particular comparison class. We have argued that the language is powerful enough to capture a number of intuitions associated with the natural use of such terms, and we moreover believe that a thorough investigation of its expressivity should reveal even more interesting applications. Naturally, the gain in expressivity is expected to come at a price of worse computational properties, the subject which we aim to study as part of the future research. It is likely that in order to achieve an optimal balance it will be necessary to restrict the syntax, possibly down to the most useful modalities $\Box$ and $\langle\cdot\rangle$, along the same lines as we have explored in the examples presented in the paper (Section 3.2.).

In principle, a clear advantage of the formalism is its backward compatibility with the standard DL languages. Note that every satisfiable DL knowledge base is at the same time a satisfiable DL$^{\mathcal{C}}$ knowledge base. Also, grounding the approach on multi-dimensional DLs gives good prospects for integrating it with other potential extensions embedded within the same paradigm, which slowly get to attract more attention in the context of the Semantic Web, as potentially useful for numerous representation and reasoning tasks.

Acknowledgements

The author is funded by a Hewlett-Packard Labs Innovation Research Program.

References

F. Baader, et al. (2003). *The description logic handbook: theory, implementation, and applications*. Cambridge University Press.

D. Benslimane, et al. (2006). 'Contextual Ontologies: Motivations, Challenges, and Solutions'. In *Proceedings of the Advances in Information Systems Conference, Izmir*.

A. Borgida & L. Serafini (2003). 'Distributed Description Logics: Assimilating Information from Peer Sources'. *Journal of Data Semantics* **1**:2003.

P. Bouquet, et al. (2003). 'C-OWL: Contextualizing Ontologies.'. In D. Fensel, K. P. Sycara, & J. Mylopoulos (eds.), *International Semantic Web Conference*, vol. 2870 of *Lecture Notes in Computer Science*, pp. 164–179. Springer.

M. L. Caliusco, et al. (2005). 'Contextual Ontologies for the Semantic Web — An Enabling Technology'. In *LA-WEB '05: Proceedings of the Third Latin American Web Congress*, p. 98, Washington, DC, USA. IEEE Computer Society.

S. Gaio (2008). 'Granular Models for Vague Predicates'. In C. Eschenbach & M. Grüninger (eds.), *Formal Ontology in Information Systems. Proceedings of the Fifth International Conference (FOIS 2008)*, pp. 253–265. IOS Press, Amsterdam.

K. Goczyla, et al. (2007). 'Contextualization of a DL Knowledge Base'. In D. Calvanese, E. Franconi, V. Haarslev, D. Lembo, B. Motik, A.-Y. Turhan, & S. Tessaris (eds.), *The Proceedings of the International Workshop on Description Logics (DL-2007)*.

D. Grossi (2007). *Desigining Invisible Handcuffs. Formal Investigations in Institutions and Organizations for Multi-Agent Systems*. Siks dissertation series 2007-16, Utrecht University.

I. Horrocks, et al. (2003). 'From SHIQ and RDF to OWL: The Making of a Web Ontology Language'. *Journal of Web Semantics* **1**:2003.

S. Klarman & S. Schlobach (2009). ‘Relativizing Description Logic Concepts to Comparison Classes’. In I. Horrocks, B. Motik, & B. Cuenca Grau (eds.), *The Proceedings of the International Workshop on Description Logics (DL-2009).*

A. Kurucz, et al. (2003). *Many-Dimensional Modal Logics: Theory and Applications*. No. 148 in Studies in Logic and the Foundations of Mathematics. Elsevier.

K. J. Laskey, et al. (2008). ‘Uncertainty Reasoning for the World Wide Web, W3C Incubator Group Report 31 March 2008’. http://www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331/.

T. Lukasiewicz & U. Straccia (2008). ‘Managing uncertainty and vagueness in description logics for the Semantic Web’. *Web Semantics* **6**(4):291–308.

M. Marx & Y. Venema (1997). *Multi Dimensional Modal Logic*. No. 4 in Applied Logic Series. Kluwer Academic Publishers.

K. Schild (1991). ‘A correspondence theory for terminological logics: preliminary report’. In *Proceedings of IJCAI-91, 12th International Joint Conference on Artificial Intelligence*, pp. 466–471, Sidney, AU.

S. Shapiro (2006). *Vagueness in Context.* Oxford University Press.

U. Straccia (2005). ‘A fuzzy description logic for the semantic web’.

H. van Ditmarsch, et al. (2007). *Dynamic Epistemic Logic*. Synthese Library. Springer.

R. van Rooij (to appear). *The Vagueness Handbook*, chap. Vagueness and Linguistics. Springer.

F. Wolter & M. Zakharyaschev (1999a). ‘Modal description logics: modalizing roles’. *Fundamenta Informaticae* **39**(4):411–438.

F. Wolter & M. Zakharyaschev (1999b). ‘Multi-Dimensional Description Logics’. In *IJCAI ’99: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pp. 104–109.

A NOTE ON THE THEORY OF COMPLETE MEREOTOPOLOGIES [*] [†]

Yavor Nenov
Department of Computer Science, University of Manchester

Abstract. We investigate theories of Boolean algebras of regular sets of topological spaces. By $RC(X)$, we denote the complete Boolean algebra of regular closed sets over a topological space X. By a *mereotopology* $\mathcal{M}$ over a topological space X, we denote every dense Boolean sub-algebra of $RC(X)$; $\mathcal{M}$ is called a *complete mereotopology* if it is a complete Boolean algebra.

In this paper we consider mereotopologies as $\mathcal{L}$-structures, where $\mathcal{L}$ is the language of Boolean algebras extended with the binary relational symbol C interpreted as the *contact* relation. We show that the $\mathcal{L}-$theories of complete mereotopologies and all mereotopologies are different. We also show that no complete mereotopology $\mathcal{M}$, over a connected, compact, Hausdorff topological space X, is elementarily equivalent to a mereotopology $\mathcal{M}'$, over X, that is a closed base for X and is finitely decomposable — i.e. every region in $\mathcal{M}'$ has only finitely many connected components.

1. Introduction

Formal systems for reasoning about space can be classified as point-based or region-based, depending on whether the variables of their formal languages are interpreted as points or sets of points. A notable example of a point-based theory of space is the decidable and complete theory of the Euclidean plane axiomatized by Tarski in (Tarski 1959). An early example of a region-based theory of space can be seen in another work of Tarski. In (Tarski 1956), he axiomatized the second-order theory of the regular closed sets of the 3-dimensional Euclidean space, with respect to the language consisting of the two predicates for the binary relation *part-of* and the property of *being a sphere*.

Authors usually motivate their interest in region-based spatial logics by arguing that they are more natural in comparison with point-based spatial logics, for people think in terms of objects, rather than in terms of the sets of points that these objects occupy. There are also practical advantages: greater expressive power, as noted in (Aiello, et al. 2007); ability to reason with incomplete knowledge, which was argued in (Renz & Nebel 2007); spatial reasoning free of numerical calculations.

In most region-based spatial logics, variables range over sets of a topological space, but it is a matter of choice whether arbitrary sets should count as regions. An example of a formal system for reasoning about arbitrary sets of topological spaces is given by McKinsey and Tarski in (McKinsey & Tarski 1944). The regular sets of a topological space are widely accepted as an appropriate choice for regions when it comes to spatial reasoning about real world objects. Regular closed (open) sets of a topological space X are those equal to the closure of their interior (the interior of their closure), and the set of all regular closed (open) sets is denoted by $RC(X)$ ($RO(X)$). $RC(X)$ and $RO(X)$ form complete

[*] The author is grateful to Dr Ian Pratt-Hartmann for his supervision, guidance and helpful comments on the content of the current document on all stages of its development.

[†] The author thanks the three anonymous reviewers for their useful comments, and especially the reviewer who suggested the proof of Lemma 4 ii).

Boolean algebras, as can be seen in (Koppelberg, et al. 1989). The aforementioned work of Tarski (Tarski 1956), is an example of a formal system in which variables range over all regular closed sets of the Euclidean space. Pratt and Schoop in (Pratt & Schoop 1998), restricted the variables of their formal system to range over the $ROP(\mathbb{R}^2)$ — the Boolean algebra of regular open polygons of $\mathbb{R}^2$, which is a dense Boolean sub-algebra of $RO(\mathbb{R}^2)$.

A *mereotopology* over a topological space X, is any dense Boolean sub-algebra of $RC(X)$; $\mathcal{M}$ is a *complete mereotopology* if it is a complete Boolean algebra — i.e. $\mathcal{M}$ is $RC(X)$. We consider this slightly weaker definition when compared to the one proposed by Pratt-Hartmann in (Pratt-Hartmann 2007), in order to allow mereotopologies over topological spaces that are not semi-regular.

In the recent years, the $\mathcal{L}-$theories of different classes of mereotopologies have been axiomatized, where $\mathcal{L}$ is the language of Boolean algebras extended with the binary relational symbol C interpreted as the *contact* relation. Roeper in (Roeper 1997) axiomatized the $\mathcal{L}-$theory of the mereotopologies over compact, Hausdorff topological spaces. Düntsch and Winter in (Düntsch & Winter 2005) established an axiomatization of the $\mathcal{L}-$theory of the mereotopologies over weakly regular, T_1 topological spaces. The $\mathcal{L}-$theory of the class of all mereotopologies was axiomatized by Dimov and Vakarelov in (Dimov & Vakarelov 2006).

To the best of our knowledge, there are no published results about the $\mathcal{L}-$theory of complete mereotopologies. In this paper we show that this theory is different from the theory of all mereotopologies. In particular, we introduce a sentence that is true in every complete mereotopology, but is not true in the incomplete mereotopology of the regular closed polygons of the real plane. As a corollary of our main result, we show that no complete mereotopology $\mathcal{M}$, over a Hausdorff topological space X, is elementarily equivalent to a mereotopology $\mathcal{M}'$, over X, that is a closed base for X and is finitely decomposable — i.e. every region in $\mathcal{M}'$ has only finitely many connected components.

We provide our main results in Section 4. The necessary definitions and basic facts about topological spaces and mereotopologies we give in Section 2. In Section 3, we summarize the main axiomatization results of classes of mereotopologies, established in (Dimov & Vakarelov 2006, Düntsch & Winter 2005, Roeper 1997), and some related results provided in (Pratt-Hartmann 2007). We discuss possible future work in Section 5.

2. Preliminary Notions

In this section we recall the definition and some examples of mereotopologies. We also prove a result about topological spaces that we use in Section 4. We assume that the reader is familiar with the basic definitions and results about Boolean algebra (see e.g. (Koppelberg et al. 1989)), and topological spaces (see e.g. (Kelley 1975)).

We start by defining the Boolean algebra of the regular closed sets over a topological space X.

Definition 1. Let X be a topological space with $\cdot^-$ and $\cdot^\circ$ the closure and interior operations in X. A subset A of X is called *regular closed* if it equals the closure of its interior, i.e. $A = A^{\circ-}$. The set of all regular closed sets in X is denoted by $RC(X)$. The Boolean operations, relations and constants can be defined in $RC(X)$ in the following way: for $a, b \in RC(X)$, $a + b = a \cup b$, $a \cdot b = (a \cap b)^{\circ-}$, $-a = (X \setminus a)^-$, $a \leq b$ iff $a \cdot b = a$, $0 = \emptyset$ and $1 = X$. The topological *contact* relation $C(x, y)$, is defined by: $C(a, b)$ iff $a \cap b \neq \emptyset$.

Recall that $\mathcal{B}$ is a *complete* Boolean algebra, if each set of elements of $\mathcal{B}$ has an infimum and a supremum. It is a well-know fact, that the structure $(RC(X), +, \cdot, -, 0, 1, \leq)$ is a complete Boolean algebra (see e.g. (Koppelberg et al. 1989)). For the definition of mereotopology, recall that, a Boolean sub-algebra $\mathcal{B}'$ of $\mathcal{B}$ is dense in $\mathcal{B}$, if for every non-zero element $a \in \mathcal{B}$ there is some non zero element $a' \in \mathcal{B}'$ such that $a' \leq a$.

Definition 2. A *mereotopology* over a topological space X is any dense Boolean sub-algebra, $\mathcal{M}$, of the complete Boolean algebra $RC(X)$.

In a dual way, one can define a *mereotopology of regular open sets*. Note that Definition 2 is weaker than the one given by Pratt-Hartmann in (Pratt-Hartmann 2007). We do not require the mereotopology to form a base for the topological space, in order to have mereotopologies over arbitrary topological spaces.

A well-studied example of an incomplete mereotopology of regular open sets, is that of the regular open polygons in the real plane (see (Pratt & Schoop 1998, Pratt & Schoop 2000, Pratt-Hartmann 2007)). The dual mereotopology, $RCP(\mathbb{R}^2)$, of the regular closed polygons of the real plane, plays an important role in proving our main result in Section 4. The formal definition of $RCP(\mathbb{R}^2)$ follows.

Definition 3. Each line in $\mathbb{R}^2$ divides the real plane into two regular open sets called open half-planes. The closure of an open half-plane is regular closed, and is called half-plane. The product in $RC(\mathbb{R}^2)$ of finitely many half planes is called a *basic polygon*. The sum of finitely many basic polygons is called a *polygon*. The set off all polygons is denoted by $RCP(\mathbb{R}^2)$.

We need the following lemma for Section 4. Recall that in a topological space X the non-empty sets $A, B \subseteq X$ are said to separate the set $C \subseteq X$ iff $C = A \cup B$, $A^- \cap B = \emptyset$ and $A \cap B^- = \emptyset$; a set $C \subseteq X$ is connected iff no pair of non-empty sets separates it; a *connected component* of a set $A \subseteq X$ is a maximal connected subset of A.

Lemma 4. Let X be a topological space and A, A_1 and A_2 be subsets of X such that A_1 and A_2 separate A. Then the following are true:

$i)$ A is closed iff A_1 and A_2 are closed;

$ii)$ A is regular closed iff A_1 and A_2 are regular closed.

Proof. First notice, that the right to left implications are obvious, since the union of two (regular) closed sets is a (regular) closed set.

$i)$ $(\rightarrow)$ From $A_1^- \subseteq A^- = A = A_1 \cup A_2$, we get $A_1^- \subseteq A_1^- \cap (A_1 \cup A_2) = (A_1^- \cap A_1) \cup (A_1^- \cap A_2) = A_1^- \cap A_1 = A_1$, so A_1 is closed. Similarly for A_2.

$ii)$ $(\rightarrow)$ From $i)$ it follows that A_1 and A_2 are closed. We want to show that $A^\circ = A_1^\circ \cup A_2^\circ$ because this implies $A_1 = A \cap X \setminus A_2 = A^{\circ-} \cap X \setminus A_2 = (A_1^{\circ-} \cup A_2^{\circ-}) \cap X \setminus A_2 = A_1^{\circ-} \cap X \setminus A_2 = A_1^{\circ-}$. The inclusion $A^\circ \supseteq A_1^\circ \cup A_2^\circ$ is trivial. Suppose $p \in A^\circ$ and w.l.o.g. let $p \in A_1$. Then $p \in X \setminus A_2$ since $A_1 \cap A_2 = \emptyset$. We get $p \in A^\circ \cap X \setminus A_2$. This set is open because A_2 is closed and subset of A_1, and, hence, $p \in A_1^\circ$.

$\square$

3. Representation Theorems for Mereotopologies

We consider mereotopologies as $\mathcal{L}-$structures, where $\mathcal{L}$ is the language $\{C, +, \cdot, -, 0, 1, \leq\}$ (see Definition 1). The $\mathcal{L}-$theories of different classes of mereotopologies were axiomatized in (Dimov & Vakarelov 2006, Düntsch & Winter 2005, Roeper 1997), although different terminology was used. In this section we give a translation of the original results in terms of mereotopologies in a way almost identical to the one in (Pratt-Hartmann 2007). Nice discussions on the algebraic approach taken in (Dimov & Vakarelov 2006, Düntsch & Winter 2005, Roeper 1997), can be seen in (Bennett & Düntsch 2007, Vakarelov 2007).

We assume the reader is familiar with some basic notions in Model Theory (see e.g. (Marker 2002)). Before we continue, we recall definitions of semi-regular and weakly regular topological spaces.

Definition 5. A topological space X is called *semi-regular*, if the set of all regular closed sets in X form a closed base for X. A semi-regular topological space is called *weakly regular* (Düntsch & Winter 2005), if for each nonempty open set $A \subseteq X$, there exists a nonempty open set B such that $B^- \subseteq A$.

Definition 6. We denote by Φ_{CA} the set of axioms for Boolean algebra, together with the following sentences:

$\psi_1 :=\ \forall x \forall y (C(x, y) \to x \neq 0);$

$\psi_2 :=\ \forall x \forall y (C(x, y) \to C(y, x));$

$\psi_3 :=\ \forall x \forall y \forall z (C(x, y + z) \leftrightarrow C(x, y) \vee C(x, z));$

$\psi_4 :=\ \forall x \forall y (x \cdot y \neq 0 \to C(x, y)).$

As we will see in Theorem 9, Φ_{CA} is an axiomatization for the class of all mereotopologies. Extending Φ_{CA} with different combinations of the axioms ψ_{ext}, ψ_{int} and ψ_{conn} (see bellow), leads to axiomatizations for mereotopologies over different classes of topological spaces.

In the following definition we abbreviate $\neg C(x, -y)$, by $x \ll y$.

Definition 7. Consider the following sentences:

ψ_{ext}	$:= \forall x (x \neq 0 \to \exists y (y \neq 0 \wedge y \ll x))$	- extensionality axiom;
ψ_{int}	$:= \forall x \forall y (x \ll y \to \exists z (x \ll z \wedge z \ll y))$	- interpolation axiom;
ψ_{conn}	$:= \forall x (x \neq 1 \wedge x \neq 0 \to C(x, -x))$	- connectedness axiom.

Theorem 8. *(Pratt-Hartmann 2007) Let $\mathcal{M}$ be a mereotopology over a topological space X, considered as an $\mathcal{L}-$structure.*

$i)$ $\mathcal{M} \models \Phi_{CA}$.

$ii)$ *If X is weakly regular, then $\mathcal{M} \models \psi_{ext}$.*

$iii)$ *If X is compact and Hausdorff and the elements of $\mathcal{M}$ form a closed base for X, then $\mathcal{M} \models \psi_{int}$.*

Theorem 9. *Let $\mathfrak{A}$ be an $\mathcal{L}$-structure.*

$i)$ *If $\mathfrak{A} \models \Phi_{CA}$, then $\mathfrak{A}$ is isomorphic to a mereotopology over a compact semi-regular T_0 topological space X. (Dimov & Vakarelov 2006)*

$ii)$ *If $\mathfrak{A} \models \Phi_{CA} \cup \{\psi_{ext}\}$, then $\mathfrak{A}$ is isomorphic to a mereotopology over a weakly regular and T_1. (Düntsch & Winter 2005)*

$iii)$ *If $\mathfrak{A} \models \Phi_{CA} \cup \{\psi_{ext,}\psi_{int}\}$, then $\mathfrak{A}$ is isomorphic to a mereotopology over a compact and Hausdorff. (Roeper 1997)*

Additionally, $\mathfrak{A} \models \psi_{conn}$ implies X is connected.

To the best of our knowledge, there are no results in the literature about the $\mathcal{L}-$theory of complete mereotopologies. It turns out that this $\mathcal{L}-$theory is different from the $\mathcal{L}-$theory of all mereotopologies. We devote the next section to establish this result.

4. The $\mathcal{L}-$theory of Complete Mereotopologies

In this section we show that the theory of complete mereotopologies differs from the theory of all mereotopologies. We accomplish this by introducing a first-order sentence that is true in each complete mereotopology but that is not true in $RCP(\mathbb{R}^2)$, which is an incomplete mereotopology. This result relies on the fact, that every non-trivial complete mereotopology satisfying $\{\psi_{ext}, \psi_{int,}\psi_{conn}\}$, has a pair of regions that are in contact, such that neither connected component of the first region is in contact with the second. The latter, however, is false for all finitely decomposable mereotopologies, including $RCP(\mathbb{R}^2)$, which on the other hand, is a non-trivial incomplete mereotopology satisfying the set of axioms $\{\psi_{ext}, \psi_{int,}\psi_{conn}\}$.

Connected regions play an important role in the proof of the main result, so we start by introducing a formula, which defines the set of connected regions in $RCP(\mathbb{R}^2)$ and each complete mereotopology $\mathcal{M}$. We make use of the fact that regular closed sets can be separated only by regular closed sets (Lemma 4).

Lemma 10. Let $\mathcal{M}$ be a complete mereotopology. Then for all $a \in \mathcal{M}$, a is connected iff $\mathcal{M} \models \psi_c[a]$, where

$$\psi_c(x) := (\forall y)(\forall z)(y \neq 0 \wedge z \neq 0 \wedge y + z = x \rightarrow C(y, z)).$$

Proof. $(\rightarrow)$ From $\mathcal{M} \not\models \psi_c[a]$ it follows that there are regular closed sets b, c that separate a, thus a is not connected.

$(\leftarrow)$ From a is not connected and Lemma 4, it follows that there are nonempty regular closed sets b, c such that $a = b + c$ and $\neg C(b, c)$. So b and c are witnesses for $\mathcal{M} \not\models \psi_c[a]$. □

In order to establish the same result for $RCP(\mathbb{R}^2)$, we have to show that a regular closed polygon can be separated only by regular closed polygons.

Lemma 11. Consider the mereotopologies $RC(\mathbb{R}^2)$ and $RCP(\mathbb{R}^2)$. For each $a \in RCP(\mathbb{R}^2)$ and $b, c \in RC(\mathbb{R}^2)$, if $a = b + c$ and $\neg C(b, c)$, then $b, c \in RCP(\mathbb{R}^2)$.

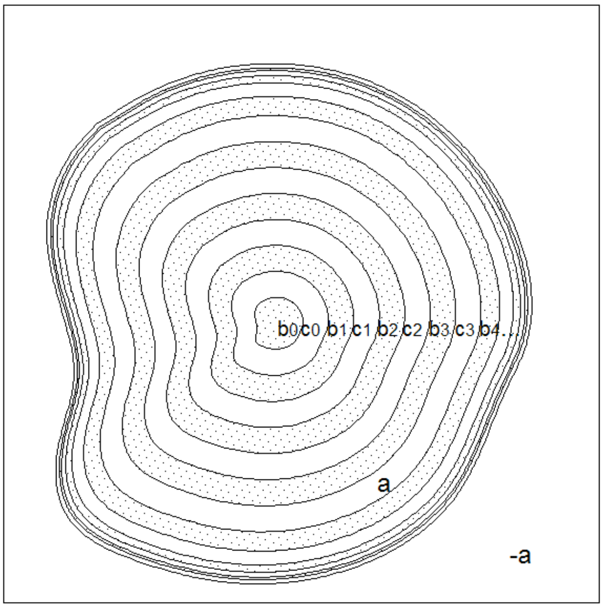


Figure 1: (Lemma 14) At least one of $b := \sum_{i\in\omega} b_i$ and $c := \sum_{i\in\omega} c_i$ is in contact with $(-a)$, but none of $\{b_i\}_{i\in\omega}$ and $\{c_i\}_{i\in\omega}$ is.

Proof. Since a is a regular closed polygon it is the sum of finitely many basic polygons, e.g. $a = \sum_{i=1}^{n} a_i$. Let $b_i := b.a_i$ and $c_i := c.a_i$. Since a_i is connected, $\neg C(b_i, c_i)$ and $a_i = b_i + c_i$, we get that $a_i = b_i$ or $a_i = c_i$. So b_i and c_i are basic polygons (either equal to 0 or to a_i). Since $b = \sum_{i=1}^{n} a_i$ and $c = \sum_{i=1}^{n} c_i$, we get that b and c are polygons, as finite sums of basic polygons. □

Lemma 12. For $a \in RCP(\mathbb{R}^2)$, a is connected iff $RCP(\mathbb{R}^2) \models \psi_c[a]$.

Proof. As in Lemma 10, considering Lemma 11. □

So far, we defined the set of connected regions in $RCP(\mathbb{R}^2)$ and each complete mereotopology by the formula ψ_c. Having shown that, we continue by constructing for every non-trivial complete mereotopology satisfying $\{\psi_{ext}, \psi_{int}, \psi_{conn}\}$, a pair of regions which are in contact, such that no connected component of the first is in contact with the second.

Lemma 13. Let $\mathcal{M}$ be a complete mereotopology such that $\mathcal{M} \models \psi_{ext} \wedge \psi_{int} \wedge \psi_{conn} \wedge \neg\psi_{triv}$, where $\psi_{triv} := (\forall x)(x = 0 \vee x = 1)$. Then there are elements a and $\{a_i\}_{i\in\omega}$ in $\mathcal{M}$ such that:

$$\begin{array}{llll} i) & C(a, -a); & ii) & a = \sum_{i\in\omega} a_i; \\ iii) & a_i \ll a_{i+1}, \text{ for } i \in \omega; & iv) & a_i \ll a, \text{ for } i \in \omega. \end{array}$$

Proof. From $\mathcal{M} \models \neg\psi_{triv}$ it follows that there is some $b \in \mathcal{M}$ such that $b \neq 0$ and $b \neq 1$. Now from $\mathcal{M} \models \psi_{ext}$, we get that there is some element $a_0 \in \mathcal{M}$ such that $a_0 \ll b$ and $a_0 \neq 0$. Considering that $\mathcal{M} \models \psi_{int}$, it follows that there is some a_1 such that $a_0 \ll a_1 \ll b$ and again by $\mathcal{M} \models \psi_{int}$, we get that there is some a_2 such that $a_1 \ll a_2 \ll b$. Arguing in a similar way one can construct a sequence $\{a_i\}_{i\in\omega}$ such that

$a_0 \ll a_i \ll a_{i+1} \ll b$ for $i \in \omega$. Now we take $a := \sum_{i\in\omega} a_i$, which is in $\mathcal{M}$, for $\mathcal{M}$ is complete. It is easy to see that $i) - iv)$ hold. We give details only in the case $i)$.

$i)$ From $a_0 \neq 0$ and $a_0 \leq a$, we get $a \neq 0$. On the other hand, $a \leq b$ since b is an upper bound of $\{a_i\}_{i\in\omega}$ and since $b \neq 1$, we get also that $a \neq 1$. Now considering ψ_{conn}, we get that $C(a, -a)$. □

In the following lemma we introduce an $\mathcal{L}-$ sentence, denoted by ψ_{cmp}, and show that it is true in each complete mereotopology.

Lemma 14. For each complete mereotopology $\mathcal{M}$, $\mathcal{M} \models \psi_{cmp}$, where

$$\begin{aligned} \psi_{cmp} :=&\quad \psi_{ext} \wedge \psi_{int} \wedge \psi_{conn} \wedge \neg\psi_{triv} \rightarrow (\exists x)(\exists y)(\psi_{\odot}(x,y)) \text{ and} \\ \psi_{\odot}(x,y) :=&\quad \psi_{\odot}(x,y) := C(x,y) \wedge (\forall x')(x' \leq x \wedge \psi_c(x') \rightarrow \neg C(x',y)). \end{aligned}$$

Proof. If $\mathcal{M} \models \psi_{ext} \wedge \psi_{int} \wedge \psi_{conn} \wedge \neg\psi_{triv}$, it follows from Lemma 13, that there are elements a and $\{a_i\}_{i\in\omega}$ in $\mathcal{M}$, such that $C(a,-a)$, $a = \sum_{i\in\omega} a_i$ and for $i \in \omega$, $a_i \ll a_{i+1}$ and $a_i \ll a$. Take $a_{-1} = 0$ and consider the following definitions:

$$\begin{aligned} &b_i = a_{2i} - a_{2i-1}, \quad b = \textstyle\sum_{i\in\omega} b_i, \quad b_{i-} = \sum_{j<i} b_j, \quad b_{i+} = \sum_{j>i} b_j, \\ &c_i = a_{2i+1} - a_{2i}, \quad c = \textstyle\sum_{i\in\omega} c_i, \quad c_{i-} = \sum_{j<i} c_j, \quad c_{i+} = \sum_{j>i} c_j. \end{aligned}$$

Since $\mathcal{M}$ is complete, it follows that $b, c, b_{i-}, c_{i-}, b_{i+}, c_{i+} \in \mathcal{M}$. (See Figure 1.)

Claim 1 For $i \in \omega$, $\neg C(b_i, b - b_i)$ and $\neg C(c_i, c - c_i)$

Proof. From $b_{i-} \leq a_{2i-2} \ll a_{2i-1}$ and $b_i \leq -a_{2i-1}$, we get that $\neg C(b_i, b_{i-})$. From $b_{i+} \leq -a_{2i+1} \ll -a_{2i}$ and $b_i \leq a_{2i}$, we get $\neg C(b_i, b_{i+})$. From $b - b_i = b_{i-} + b_{i+}$, we get that $\neg C(b_i, b - b_i)$. Similarly $\neg C(c_i, c - c_i)$.

Claim 2 From $b' \leq b$ and $\mathcal{M} \models \psi_c[b']$, it follows $b' \ll a$. From $c' \leq c$ and $\mathcal{M} \models \psi_c[c']$, it follows $c' \ll a$.

Proof. We will show that there is some $i \in \omega$ such that $b' \leq b_i$. Since $b' \leq b$ and $b = \sum_{i\in\omega} b_i$ there is some $i \in \omega$ such that $b' \cdot b_i \neq 0$. We have that $b' = b' \cdot b = b' \cdot (b + b_i - b_i) = b' \cdot b_i + b' \cdot (b - b_i)$. From Claim 1 it follows that $\neg C(b_i, (b - b_i)$ and thus $\neg C(b' \cdot b_i, b' \cdot (b - b_i))$. Now from $\mathcal{M} \models \psi_c[b']$ and $b' \cdot b_i \neq 0$ it follows that $b' \cdot (b - b_i) = 0$ and thus $b' = b_i \cdot b'$, which is $b' \leq b_i$. Finally, we get that $b' \leq b_i \ll a$. Similarly $c' \leq c$ and $\mathcal{M} \models \psi_c[c']$ imply $c' \ll a$.

Finally, from $a = b + c$ and $C(a, -a)$, it follows that either $C(b, -a)$ or $C(c, -a)$. W.l.o.g., let $C(b, -a)$ be the case. By Claim 2 we get that $\mathcal{M} \models \psi_{\odot}[b, -a]$ and so $\mathcal{M} \models \psi_{cmp}$. □

Lemma 15. $RCP(\mathbb{R}^2) \not\models \psi_{cmp}$.

Proof. It is well known that $RCP(\mathbb{R}^2) \models \psi_{ext} \wedge \psi_{int} \wedge \psi_{conn} \wedge \neg\psi_{triv}$, so it suffices to show that $RCP(\mathbb{R}^2) \not\models (\exists x)(\exists y)(\psi_{\odot}(x,y))$. Let a and b be regular closed polygons, such that $a \cap b \neq \emptyset$. Since a is a polygon, it can be represented as a finite sum of basic polygons, say $a = \sum_{i=1}^{n} a_i$. Since the sum of finitely many regular closed sets is just their union, we get that $a_i C b$ for some $i \leq n$. Since the basic polygons are connected and Lemma 12, we get that $RCP(\mathbb{R}^2) \not\models \psi_{\odot}[a, b]$. So, we get that $RCP(\mathbb{R}^2) \models (\forall x)(\forall y)(\neg\psi_{\odot}(x,y))$ and thus $RCP(\mathbb{R}^2) \models \neg\psi_{cmp}$. □

Theorem 16. *The $\mathcal{L}-$theory of complete mereotopologies is different from the $\mathcal{L}-$theory of:*

$i)$ *the class of all mereotopologies;*

$ii)$ *the class of mereotopologies over weakly regular topological spaces;*

$iii)$ *the class of mereotopologies over compact Hausdorff topological spaces.*

Proof. $RCP(\mathbb{R}^2)$ is a member of each of the above classes. □

Theorem 17. *Let X be a connected, compact, Hausdorff topological space and let the complete mereotopology, $\mathcal{M}$, over X be non-trivial. Then the $\mathcal{L}-$theory of $\mathcal{M}$ is different from the $\mathcal{L}-$theory of every finitely decomposable mereotopology, $\mathcal{M}'$, over X, that is a close base for X.*

Proof. From Theorem 8 and $\mathcal{M}$ being non-trivial, it follows that $\mathcal{M}' \models \psi_{ext} \wedge \psi_{int} \wedge \psi_{conn} \wedge \neg\psi_{triv}$. Since $\mathcal{M}'$ is finitely decomposable, one can show, as in Lemma 15, that $\mathcal{M}' \not\models \psi_{cmp}$. But since $\mathcal{M}$ is a complete mereotopology, we have that $\mathcal{M} \models \psi_{cmp}$. So $\mathcal{M}$ and $\mathcal{M}'$ have different $\mathcal{L}-$theories. □

Corollary 18. The $\mathcal{L}-$theories of $RC(\mathbb{R}^2)$ and $RCP(\mathbb{R}^2)$ are different.

5. Conclusions and Future Work

We showed that the theory of complete mereotopologies is different from the theory of all mereotopologies. As a future step, one can establish an axiomatization for the theory of complete mereotopologies or the theories of specific complete mereotopologies such as the mereotopologies of the regular closed sets in the real line, real plane or higher dimensional topological spaces.

References

M. Aiello, et al. (2007). 'What is Spatial Logic?'. In *Handbook of Spatial Logics*, pp. 1–11. Springer Netherlands.

B. Bennett & I. Düntsch (2007). 'Axioms, Algebras and Topology'. In *Handbook of Spatial Logics*, pp. 99–159. Springer Netherlands.

G. Dimov & D. Vakarelov (2006). 'Contact Algebras and Region-based Theory of Space: Proximity Approach - I'. *Fundamenta Mathematicae* **74**(2,3):209–249.

I. Düntsch & M. Winter (2005). 'A Representation Theorem for Boolean Contact Algebras'. *Theor. Comput. Sci.* **347**(3):498–512.

J. Kelley (1975). *General Topology*. Springer.

S. Koppelberg, et al. (1989). *Handbook of Boolean Algebras Vol. 1*. Elsevier.

D. Marker (2002). *Model Theory: An Introduction*. Springer.

J. C. C. McKinsey & A. Tarski (1944). 'The Algebra of Topology'. *The Annals of Mathematics* **45**(1):141–191.

I. Pratt & D. Schoop (1998). 'A Complete Axiom System for Polygonal Mereotopology of the Real Plane'. *Journal of Philosophical Logic* **27**(6):621–658.

I. Pratt & D. Schoop (2000). 'Expressivity in Polygonal, Plane Mereotopology'. *Journal of Symbolic Logic* **65**:822–838.

I. Pratt-Hartmann (2007). 'First-Order Mereotopology'. In *Handbook of Spatial Logics*, pp. 13–97. Springer Netherlands.

J. Renz & B. Nebel (2007). 'Qualitative Spatial Reasoning Using Constraint Calculi'. In *Handbook of Spatial Logics*, pp. 161–215. Springer Netherlands.

P. Roeper (1997). 'Region-Based Topology'. *Journal of Philosophical Logic* **26**(3):251–309.

A. Tarski (1956). 'Foundations of The Geometry Of Solids'. *Logic, semantics, metamathematics* pp. 24–29.

A. Tarski (1959). *What is Elementary Geometry?*, pp. 16–29. North Holland Publishing Company.

D. Vakarelov (2007). 'Region-Based Theory of Space: Algebras of Regions, Representation Theory, and Logics'. In *Mathematical Problems from Applied Logic II*, vol. 5 of *International Mathematical Series*, pp. 267–348. Springer New York.

LTL WITH A SUB-ORDER

Amaldev Manuel
Institute of Mathematical Sciences, India

Abstract. We consider the problem of specifying properties of an asynchronous transition system with unbounded-ly many processes. We introduce a temporal logic and shows that it is decidable. We then discuss various issues related to enhancing the expressiveness and model-checking. Finally we end with a discussion on data languages.

1. Introduction

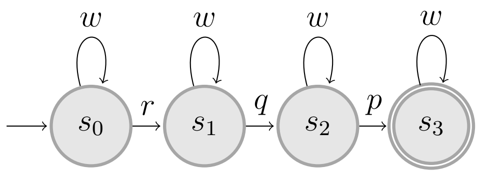
Temporal logics, which are modal logics dealing with flows of time (irreflexive and transitive frames), had been highly successful as specification languages for describing the behaviours of computing systems. The simplest of all, LTL (propositional discrete linear time temporal logic) (Pnueli 1977) formulas work on histories[a] of the form

$$w = w_1 w_2 \dots w_n, w_i \in 2^P$$

where P is the set of all propositions and is evaluated with respect to a particular time point $i \in |w|$[b]. The commonly used temporal connectives are $\diamondplus$ for FUTURE, $\diamondminus$ for PAST, $\oplus$ for TOMORROW, $\ominus$ for YESTERDAY, u for UNTIL and s for SINCE. The duals of $\diamondplus$ and $\diamondminus$ are $\boxplus$ for HENCEFORTH and $\boxminus$ for HITHERTO.

It is well-known that $\mathrm{LTL}(s, u, \oplus, \ominus)$ is expressively complete with respect to $\mathrm{FO}(<)$ (Kamp 1968), whereas $\mathrm{LTL}(\diamondplus, \diamondminus, \oplus, \ominus)$ and $\mathrm{LTL}(\diamondplus, \diamondminus)$ are complete with respect to $\mathrm{FO}^2(<, \lessdot)$ and $\mathrm{FO}^2(<)$ respectively(Etessami, et al. 1997)[cd]. The fact that the logic is propositional and the lower complexity of its decision problems[e] makes LTL a suitable specification mechanism for describing sequential behaviours of computing machineries.

A transition system describing a printer system with only one process. The actions are request, wait, queue and print abbreviated here as r, w, q, p respectively.



Every request is eventually followed by a queue and in turn by a print. A sample execution trace of the above described system is given below.

$$w \to w \to r \to w \to w \to q \to w \to w \to p \to w$$

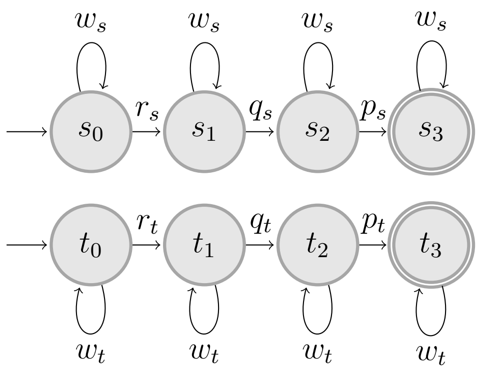
[a]In this paper we consider only LTL over finite histories.
[b]$|w|$ denote the length of w.
[c]FO^2 is the two variable fragment of first order logic.
[d]$\lessdot$ denotes the successor relation, which is the Hasse covering relation for <.
[e]LTL satisfiability problem is in PSPACE whereas FO(<) satisfiability is NON-ELEMENTARY (Stockmeyer 1974) and FO^2 satisfiability is NEXPTIME-COMPLETE(Etessami et al. 1997).

A desirable property is fairness which states that every request is eventually queued and then printed. Stated in the language of LTL it is as follows.

$$r \rightarrow \diamondplus(q \wedge \diamondplus p) \wedge \boxplus(r \rightarrow \diamondplus(q \wedge \diamondplus p))$$

If we have two asynchronous processes, we could still use LTL by adding identity to actions corresponding to each process by means of additional propositional variables, as shown below.

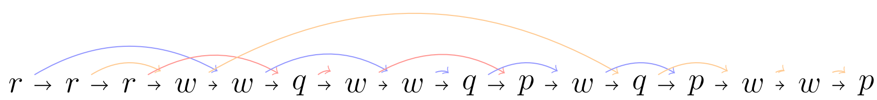
A transition system describing a printer system with two asynchronous processes. The actions request, wait, queue and print are paired with the identity of the process, a boolean value. The invariant that every request is eventually followed by a queue and in turn by a print holds individually in each of the processes. The behaviour of the whole system will be the arbitrary interleaving-s of the behaviours of the individual processes.



The fairness is stated in terms of actions of each processes, in the following way.

$$(r_s \rightarrow \diamondplus(q_s \wedge \diamondplus p_s) \wedge r_t \rightarrow \diamondplus(q_t \wedge \diamondplus p_t)) \wedge \boxplus(r_s \rightarrow \diamondplus(q_s \wedge \diamondplus p_s) \wedge r_t \rightarrow \diamondplus(q_t \wedge \diamondplus p_t))$$

This way of adding propositional variables for identities does not generalize if the system contains an unbounded number of processes[f]. In which case, one way to represent the execution trace is by annotating the sequential trace ordered by $<$, by yet another order $\lesssim$ as shown below. The following is an execution trace of printer system, the coloured edges denote the execution of each process.



The order $\lesssim$ satisfies the following properties,

1. $\lesssim$ is compatible with $<$, that means, if $i \lesssim j$ implies that $i < j$.
2. The order relation $\lesssim$ is a disjoint union of chains, this is by virtue of the fact that each chain in $\lesssim$ is a sequential trace of a process.

Henceforth $<$ denotes a total order and $\lesssim$ stands for a subset of $<$ which satisfies the above properties. One way to specify properties of such structures is to define temporal connectives which take into account both order relations. Linear orders with added relations have been studied recently, from different perspectives. CARET introduced in (Alur,

[f]Unbounded here means finite but not bounded by a constant.

et al. 2004, Alur & Madhusudan 2006) works over nested words, where the words are ornamented with a relation μ–which is a union of chains of length two, looks at program executions where the relation μ corresponds to call-return patterns which are inherently nested. Another logic one finds in the literature is the $\mathrm{LTL}\downarrow$, LTL with freeze quantifiers, introduced in (Demri & Lazic 2006, Demri & Lazic 2009), which addresses words over an alphabet $\Sigma \times \Delta$ where Σ is a finite alphabet and Δ is an infinite data domain. We define the following temporal logic,

$$\varphi ::= p \mid \Diamondplus_{\sim}\varphi \mid \Diamondplus_{\nsim}\varphi \mid \Diamondminus_{\sim}\varphi \mid \Diamondminus_{\nsim}\varphi \mid \neg\varphi \mid \varphi_1 \vee \varphi_2$$

The semantics of the logic is given with respect to histories ordered by $<$ and $\lesssim$, in the following way, boolean cases are as usual,

$$\begin{array}{lcl}
w, i \vDash p & \Leftrightarrow & w_i = p \\
w, i \vDash \Diamondplus_{\sim}\varphi & \Leftrightarrow & \exists j.\ i \lesssim j \wedge w, j \vDash \varphi \\
w, i \vDash \Diamondplus_{\nsim}\varphi & \Leftrightarrow & \exists j.\ i < j \wedge i \not\lesssim j \wedge w, j \vDash \varphi \\
w, i \vDash \Diamondminus_{\sim}\varphi & \Leftrightarrow & \exists j.\ j \lesssim i \wedge w, j \vDash \varphi \\
w, i \vDash \Diamondminus_{\nsim}\varphi & \Leftrightarrow & \exists j.\ j < i \wedge j \not\lesssim i \wedge w, j \vDash \varphi
\end{array}$$

We say that the history $w \vDash \varphi$ if $w, 1 \vDash \varphi$. We can define $\boxplus_{\sim}\varphi = \neg \Diamondplus_{\sim} \neg\varphi$, $\boxplus_{\nsim}\varphi = \neg \Diamondplus_{\nsim} \neg\varphi$, $\boxplus\varphi = \boxplus_{\sim}\varphi \wedge \boxplus_{\nsim}\varphi$, $\Diamondplus\varphi = \Diamondplus_{\sim}\varphi \vee \Diamondplus_{\nsim}\varphi$. Symmetrically we can define the past modalities as well. The fairness condition can be expressed in this logic as

$$(r \to \Diamondplus_{\sim}(q \wedge \Diamondplus_{\sim}p)) \wedge \boxplus(r \to \Diamondplus_{\sim}(q \wedge \Diamondplus_{\sim}p))$$

To give some more examples, $\neg(\Diamondplus(p \wedge \Diamondplus_{\sim}p))$ states that every process prints atmost once, $\neg \Diamondplus (r \wedge \Diamondplus_{\sim}(w \wedge \Diamondplus_{\sim}(w \wedge \Diamondplus_{\sim}(w \wedge \Diamondplus_{\sim}p))))$ says that every request has two wait only atmost twice before getting printed.

One benchmark for measuring the expressiveness of temporal logic is of course the classical first order logic. By the standard translation of modal logics we can show that,

Proposition 1..1. *Every* $\mathrm{LTL}(\Diamondplus_{\sim}, \Diamondplus_{\nsim}, \Diamondminus_{\sim}, \Diamondminus_{\nsim})$ *formula* φ *can be converted to an equivalent* $\mathrm{FO}^2(<, \lesssim)$ *formula* $\hat{\varphi}(x)$*, where* $|\hat{\varphi}(x)| \in \mathcal{O}(|\varphi|)$ *and* $qdp(\hat{\varphi}(x)) = odp(\varphi)$[g].

Proof. Use the standard translation to obtain $\hat{\varphi}(x)$ from φ, boolean cases are omitted.

$$\begin{array}{lcl}
ST_x(p) & := & p(x) \\
ST_x(\Diamondplus_{\sim}\varphi) & := & \exists y.\ x \lesssim y \wedge ST_y(\varphi) \\
ST_x(\Diamondplus_{\nsim}\varphi) & := & \exists y.\ x < y \wedge x \not\lesssim y \wedge ST_y(\varphi) \\
ST_x(\Diamondminus_{\sim}\varphi) & := & \exists y.\ y \lesssim x \wedge ST_y(\varphi) \\
ST_x(\Diamondminus_{\nsim}\varphi) & := & \exists y.\ y < x \wedge y \not\lesssim x \wedge ST_y(\varphi)
\end{array}$$

□

The other direction is the interesting one, along the lines of (Etessami et al. 1997) (we differ from the original proof only in the last step, however we reproduce the proof here.) we can show that,

[g]quantifier depth and operator depth respectively.

Theorem 1..2. *Every formula $\varphi(x)$ in* $\mathrm{FO}^2(<,\lesssim)$ *can be converted to an equivalent* $LTL(\oplus_{\sim},\oplus_{\nsim},\ominus_{\sim},\ominus_{\nsim})$ *formula φ', where $|\varphi'| \in 2^{\mathcal{O}(|\varphi|(qdp(\varphi)+1))}$ and $odp(\varphi') = qdp(\varphi)$.*

Proof. The proof is by induction on the structure of the formulas. When $\varphi(x)$ is atomic, that is $\varphi = p(x)$, $\varphi' = p_i$. When $\varphi(x)$ is composite, that is $\varphi(x) = \neg\varphi_1(x)$ (or $\varphi(x) = \varphi_1(x) \vee \varphi_2(x)$), we recursively compute φ'_1 (or φ'_1 and φ'_2) and output $\neg\varphi'_1$ (or $\varphi'_1 \vee \varphi'_2$).

The remaining cases are when $\varphi(x)$ is of the form $\exists x.\varphi_1(x)$ or $\exists y.\varphi_1(x,y)$. In the first case φ is equivalent to $\exists y.\varphi_1(y)$ (by renaming) and hence reduces to the second case (considering x as a dummy variable). In the second case we rewrite $\varphi_1(x,y)$ in the form

$$\varphi_1(x,y) = \beta\left(\chi_0(x,y),\dots,\chi_{r-1}(x,y),\xi_0(x),\dots,\xi_{s-1}(x),\zeta_0(y),\dots,\zeta_{t-1}(y)\right)$$

where β is a boolean formula, each χ_i is an order formula, ξ_i is an atomic or existential FO^2 formula with $qdp(\xi_i) < qdp(\varphi)$ and ζ_i is an atomic or existential FO^2 formula with $qdp(\zeta_i) < qdp(\varphi)$. We next pull out the ξ_i's from β by doing a case distinction on which of the sub-formulas ξ_i hold or not. Rewriting the previous expression as,

$$\bigvee_{\bar{\gamma}\in\{\top,\bot\}^s}\left(\bigwedge_{i<s}(\xi_i \leftrightarrow \gamma_i) \wedge \exists y.\beta\left(\chi_0,\dots,\chi_{r-1},\gamma_0,\dots,\gamma_{s-1},\zeta_0,\dots,\zeta_{t-1}\right)\right)$$

Next we do a case distinction on which order relation, called *order type*, holds between x and y. All possible relations which can exist between x and y which satisfy the conditions for $\lesssim$ will be an order type, namely, $x = y$, $x \lesssim y$, $x < y \wedge x \not\lesssim y$, $y \lesssim x$, $x > y \wedge y \not\lesssim x$[h]. When we assume that one of these order types is true, each atomic formula evaluates to either $\bot$ or $\top$ and in particular, each of the ξ's evaluates to either $\bot$ or $\top$; which we denote by ξ^τ. Finally, we can rewrite φ as follows, where Υ stands for the set of all order types:

$$\bigvee_{\bar{\gamma}\in\{\top,\bot\}^s}\left(\bigwedge_{i<s}(\xi_i \leftrightarrow \gamma_i) \wedge \bigvee_{\tau\in\Upsilon}\exists y\left(\tau \wedge \beta\left(\chi_0^\tau,\dots,\chi_{r-1}^\tau,\gamma_0,\dots,\gamma_{s-1},\zeta_0,\dots,\zeta_{t-1}\right)\right)\right)$$

If τ is an order type and $\psi(y)$ an FO^2 formula then for $\exists y.\tau \wedge \psi(y)$, an equivalent LTL formula $\tau\langle\psi\rangle$ can be obtained in the following way,

τ	$x = y$	$x \lesssim y$	$x < y \wedge x \not\lesssim y$	$y \lesssim x$	$x > y \wedge y \not\lesssim x$
$\tau<\psi>$	ψ	$\oplus_{\sim}\psi$	$\oplus_{\nsim}\psi$	$\ominus_{\sim}\psi$	$\ominus_{\nsim}\psi$

Now, we recursively compute $\xi'_i, i < s$ and $\zeta'_i, i < t$ and outputs,

$$\bigvee_{\bar{\gamma}\in\{\top,\bot\}^s}\left(\bigwedge_{i<s}(\xi'_i \leftrightarrow \gamma_i) \wedge \bigvee_{\tau\in\Upsilon}\tau\left\langle\beta\left(\chi_0^\tau,\dots,\chi_{r-1}^\tau,\gamma_0,\dots,\gamma_{s-1},\zeta'_0,\dots,\zeta'_{t-1}\right)\right\rangle\right)$$

The respective bounds are easily proved by an induction on the cases. ☐

Corollary. $\mathrm{LTL}(\oplus_{\sim},\oplus_{\nsim},\ominus_{\sim},\ominus_{\nsim})$ *is expressively complete w.r.t* $\mathrm{FO}^2(<,\lesssim)$.

The next interesting question about the logic is decidability, it turns out that the logic is decidable.

[h]The order types are mutually exclusive.

Proposition 1..3. $\mathrm{FO}^2(<, \lesssim)$ *is decidable in* NEXPTIME.

Proof. (Bojanczyk, et al. 2006) shows that the satisfiability of $\mathrm{FO}^2(<, \sim)$ is decidable in NEXPTIME, where $\sim$ is an equivalence relation, We interpret $\mathrm{FO}^2(<, \lesssim)$ in $\mathrm{FO}^2(<, \sim)$ by the following translation,

$$
\begin{array}{lcllcl}
\ulcorner a(x) \urcorner & := & a(x) & \ulcorner x = y \urcorner & := & x = y \\
\ulcorner x < y \urcorner & := & x < y & \ulcorner x \lesssim y \urcorner & := & x < y \wedge x \sim y \\
\ulcorner \neg\varphi \urcorner & := & \neg\ulcorner \varphi \urcorner & \ulcorner \varphi_1 \vee \varphi_2 \urcorner & := & \ulcorner \varphi_1 \urcorner \vee \ulcorner \varphi_2 \urcorner \\
\ulcorner \exists x.\varphi \urcorner & := & \exists x.\ulcorner \varphi \urcorner & & &
\end{array}
$$

This completes the proof. □

Corollary. *Satisfiability of* $\mathrm{LTL}(\diamondplus_{\sim}, \diamondplus_{\nsim}, \diamondminus_{\sim}, \diamondminus_{\nsim})$ *is in* NEXPTIME.

2. Discussion

Not only that we can get $\lesssim$ from $\sim$ in the presence of a total order $<$, we can go the other way as well. We can interpret an equivalence relation $\sim$ as $\ulcorner x \sim y \urcorner := x = y \vee x \lesssim y \vee y \lesssim x$. Below, we use this translation to import the relevant theorems to our setting.

Diamonds are not sufficient to specify properties over discrete linear time. We can enhance the expressiveness of our temporal logic by adding modalities for YESTERDAY and TOMORROW. We can redo the proofs and show that they translate to $\mathrm{FO}^2(<, \prec, \lesssim, \precsim)$. The trivial way to get expressive completeness with respect to $\mathrm{FO}^2(<, \prec, \lesssim, \precsim)$ is to add a modality for each order type definable, which in turn corresponds to taking the product of modalities definable in each oder (for instance, the order types definable by the vocabulary $(<, \prec)$ in two variables correspond to the modalities YESTERDAY, TOMORROW, PAST, FUTURE, DISTANT PAST, DISTANT FUTURE) interpreted in the obvious way. But, the satisfiability problem for $\mathrm{FO}^2(<, \prec, \lesssim, \precsim)$ is as hard as reachability in vector addition systems (Bojanczyk et al. 2006). The situation is worse when we go for binary modalities like UNTIL, because of the following.

Proposition 2..1 ((Bojanczyk et al. 2006))**.** *Satisfiability of* $\mathrm{FO}^3(<, \lesssim)$ *is undecidable.*

We can redo the above proof and show that,

Proposition 2..2. *Satisfiability of* $\mathrm{FO}^3(<, \precsim)$ *is undecidable.*

Another extension is to refine the order $\lesssim$, where we see each local process as a collection of threads. That is to say that the order $<$ stands for the global ordering of the events in each process, $\lesssim$ is the collection of orders corresponding to each local process and yet another order $\lesssim'$ is the collection of threads in each local process and hence $\lesssim'$ has to be compatible with $\lesssim$ and again a disjoint union of chains. But, defining a expressively complete decidable temporal logic is hard, since

Proposition 2..3 ((Björklund & Bojanczyk 2007))**.** *Satisfiability of* $\mathrm{FO}^2(<, \lesssim, \lesssim')$*, where* $\lesssim'$ *is compatible with* $\lesssim$ *and is a union of chains, is undecidable.*[i]

[i]Satisfiability of $\mathrm{FO}^2(<, \lesssim, \lesssim')$ is undecidable when $\lesssim'$ is not compatible with $\lesssim$ but still compatible with $<$. Since given φ we can construct the sentence $\varphi' \equiv \varphi \wedge \forall xy.\, x \lesssim' y \rightarrow x \lesssim y$.

It may be noted that in principle the temporal logic which is introduced here can be used for model-checking. The histories are of the form $(w, \lesssim)$ where $w \in \Sigma^*$ is a word and $\lesssim$ is as previously described. A collection of such histories can be recognized by the following automaton. Class memory automaton (Björklund & Schwentick 2007) was introduced in the context of data languages, we use a reformulation of the same here, the important fact is that here we do not have data elements, instead an order relation $\lesssim$. Using the interpretation mentioned earlier we can import the relevant theorems here as well. we denote by $\precsim$ the Hasse covering relation of $\lesssim$. We say that $i \in [|w|]$ is $\lesssim$-minimal if $\neg\exists j \in [|w|].\ j \lesssim i$, similarly i is $\lesssim$-maximal if $\neg\exists j \in [|w|].\ i \lesssim j$[j].

Definition 2..1. A Class memory automaton A is a six tuple $A = (Q, \Sigma, \Delta, q_0, F_l, F_g)$ where Q is a finite set of states, q_0 is the initial state, $F_l \subseteq Q$ is a set of local accepting states, $F_g \subseteq F_l$ is a set of global accepting states and $\Delta \subseteq Q \times (Q \cup \{\perp\}) \times \Sigma \times Q$ is the transition relation.

A run ρ of the automaton A on a given word $\mathbf{w} = (w, \lesssim)$, where $w = a_1 a_2 \ldots a_n$ is a sequence $q_0 q_1 \ldots q_n$ such that q_0 is the initial state and for all $i \in [n]$ there is a transition $\delta_i = (p, p', a, q)$ such that (i) $p = q_{i-1}$ and $q = q_i$ (ii) $a = a_i$ (iii) if i is $\lesssim$-minimal then p' should be $\perp$. Else there is a $j \in [n]$ such that $j \precsim i$, and $p' = q_j$. The run ρ is said to be accepting if $\{q_i \mid i \text{ is } \lesssim\text{-maximal}\} \subseteq F_l$ and $q_n \in F_g$.

Example 2..1. The printer system can be modelled by the automaton in the following way. $A = (Q, \Sigma, \Delta, q_0, F_l, F_g)$ where $\Sigma = \{w, p, q, r\}$. The states of the automaton are $Q = \{s_0, s_r, s_q, s_p\}$. The accepting states $F_l = F_g = \{s_p\}$ and the initial state is s_0. The transition relation contains the following tuples $\Delta = \{(s, s', w, s') \mid s, s' \in Q\} \cup \{(s, s_0, r, s_r) \mid s \in Q\} \cup \{(s, s_r, q, s_q) \mid s \in Q\} \cup \{(s, s_q, p, s_p) \mid s \in Q\}$.

Class memory automaton is expressively equivalent to $\mathrm{EMSO}^2(<, \prec, \lesssim, \precsim)$ – Existential Monadic Second Order logic with a first order kernel in FO^2 – and its emptiness checking is decidable (Bojanczyk et al. 2006, Björklund & Schwentick 2007). Any LTL formula can be translated to a CMA in 2-EXPTIME going via the FO^2 translation(Bojanczyk et al. 2006). The model-checking problem asks the following question: Given a collection of histories as the runs of an automaton A, do they all satisfy the specification φ, that is whether $\mathrm{A} \vDash \varphi$. Since the logic is closed under complementation, we can find the automaton $\mathrm{A}_{\neg\varphi}$ which accepts all histories which do not satisfy the formula φ. Since CMA's are closed under intersection, we can find the automaton A' such that $\mathrm{L}(\mathrm{A}') = \mathrm{L}(\mathrm{A}) \cap \mathrm{L}(\mathrm{A}_{\neg\varphi})$[k]. Now the problem reduces to checking whether the language of A' is non-empty, which is decidable. But the complexity of emptiness checking of CMA is as hard as Petri net reachability making it intractable.

3. Data Languages

Our logic and automaton work over structures of the form $(w, \lesssim)$, where $\lesssim$ is an extra-linguistic object, in the sense that it lies outside the level of abstraction of the alphabet. Even though in logic it is viable and convenient to consider such objects, in reality it is

[j] $[n]$ for $n \in \mathbb{N}$ stands for the set $\{1, \ldots n\}$.

[k] $\mathrm{L}(\mathrm{A})$ is the language accepted by the automaton A.

common that such information are represented at the level of abstraction of the alphabet[l]. Examples are time-stamps in distributed systems, nonce-s in security protocols, process id-s in schedulers, ID attributes in XML documents etc.

It is clear that since $\lesssim$ has to represent unbounded-ly many processes, any alphabet which is going to code it up, has to be unbounded. It can be done by introducing a countably infinite alphabet with suitable relations and operations on the alphabet[m]. Hence the words, called data-words, are going to be over $\Sigma \times \Delta$, where Σ is finite and Δ is countably infinite. A collection of data-words is called a data language. For instance the semantics of our logic can be given in terms of words over $\Sigma \times \Delta$ where Δ has countably infinitely many arbitrarily long chains (for example $\Delta = \mathbb{N} \times \mathbb{N}$ with the relation $(x, y) \lesssim (x', y') \Leftrightarrow x = x' \wedge y < y'$).

If there is one conclusion to draw it will be this that it seems difficult to get decidable yet sufficiently expressive logics to specify unbounded systems. We conclude by pointing to a comprehensive survey on Data languages (Segoufin 2006).

[l]It may very well be the case that such information are unavailable outside the realm of language.

[m]Equality is the most essential, prefix relation, arithmetical predicates etc. are other examples, though these have not yet been studied.

References

R. Alur, et al. (2004). 'A Temporal Logic of Nested Calls and Returns'. In K. Jensen & A. Podelski (eds.), *TACAS*, vol. 2988 of *Lecture Notes in Computer Science*, pp. 467–481. Springer.

R. Alur & P. Madhusudan (2006). 'Adding Nesting Structure to Words'. In O. H. Ibarra & Z. Dang (eds.), *Developments in Language Theory*, vol. 4036 of *Lecture Notes in Computer Science*, pp. 1–13. Springer.

H. Björklund & M. Bojanczyk (2007). 'Shuffle Expressions and Words with Nested Data'. In L. Kucera & A. Kucera (eds.), *MFCS*, vol. 4708 of *Lecture Notes in Computer Science*, pp. 750–761. Springer.

H. Björklund & T. Schwentick (2007). 'On Notions of Regularity for Data Languages'. In E. Csuhaj-Varjú & Z. Ésik (eds.), *FCT*, vol. 4639 of *Lecture Notes in Computer Science*, pp. 88–99. Springer.

M. Bojanczyk, et al. (2006). 'Two-Variable Logic on Words with Data.'. In *LICS*, pp. 7–16. IEEE Computer Society.

S. Demri & R. Lazic (2006). 'LTL with the Freeze Quantifier and Register Automata'. In *LICS '06: Proceedings of the 21st Annual IEEE Symposium on Logic in Computer Science*, pp. 17–26, Washington, DC, USA. IEEE Computer Society.

S. Demri & R. Lazic (2009). 'LTL with the freeze quantifier and register automata'. *ACM Trans. Comput. Log.* **10**(3).

K. Etessami, et al. (1997). 'First-Order Logic with Two Variables and Unary Temporal Logic'. In *Logic in Computer Science*, pp. 228–235.

H. Kamp (1968). *On tense logic and the theory of order*. Ph.D. thesis, UCLA.

A. Pnueli (1977). 'The Temporal Logic of Programs'. In *18th Annual Symposium on Foundations of Computer Science*, pp. 46–57, Providence, Rhode Island. IEEE.

L. Segoufin (2006). 'Automata and Logics for Words and Trees over an Infinite Alphabet'. In Z. Ésik (ed.), *CSL*, vol. 4207 of *Lecture Notes in Computer Science*, pp. 41–57. Springer.

L. Stockmeyer (1974). 'The complexity of decision problems in automata theory and logic'. PhD thesis, TR 133, M.I.T., Cambridge.

ON LANGUAGES THAT CONTAIN THEIR OWN UNGROUNDEDNESS PREDICATE *

Stefan Wintein
Department of Philosophy, Tilburg University

Abstract. Kripke's fixed point construction deals with paradoxical sentences such as the Liar or elements of a Yablo sequence by declaring them neither true nor false. We say that these sentences receive a third truth value, which we call *ungrounded*. We specify circumstances under which an interpreted language can—besides its truth and falsity predicate— contain its own *ungroundedness predicate*; in such a language the assertion 'σ is ungrounded' is true just in case σ is in fact ungrounded. Then, our result is applied to shed light on a dubious claim that has recently been advanced in the literature with respect the so called *Hardest Logic Puzzle Ever*.

1. Introduction

The notorious Liar paradox causes serious problems for our intuitive understanding of truth. The literature's reactions to the Liar constitute a densely populated inhomogeneous area of theories. Formally, a boundary can be drawn between *axiomatic* and *model-theoretic* theory constructions while philosophically such a boundary can be drawn between theories that study the Liar phenomenon in an *ordinary language environment* or in an *environment of mathematical language*. In the philosophical area, theories of (self-referential) truth 'range from attempts to explicate our intuitive notion of truth to assigning the truth predicate a rôle in the foundations of mathematics'[1]. The contribution of this paper is a model-theoretic result that is obtained in the attempt to make sense of a particular phenomenon of self-referentiality occurring in ordinary language and thought.

The model-theoretic result. A very important technique for the construction of theories of self-referential truth is *Kripke's fixed point construction*. Starting with a classical structure (called a *ground structure*) for the truth-free fragment[2] of a language L_T which contains a truth predicate 'T' the construction generates, upon specification of a monotonic valuation scheme V, a partial structure for L_T. The associated interpreted language $\mathcal{L}_T$[3] has the so called *fixed point property*; the truth value—true (**t**), false (**f**) or ungrounded (**u**)— of a sentence σ equals the truth value of the sentence which expresses that σ is true; σ and $T(\ulcorner\sigma\urcorner)$ are *semantically intersubstitutable*. When Kripke (1975) writes that 'Being a fixed point $\mathcal{L}_T$ contains its own truth predicate' it is arguable that his reason for calling $\mathcal{L}_T$ a language that contains its own truth predicate is precisely that it has the mentioned semantic intersubstitutability property.

In fact, Kripke's construction may also be applied to a ground structure to obtain an language $\mathcal{L}_{TF}$ which, in the sense in which $\mathcal{L}_T$ is a language that contains its own truth predicate, is a language that contains its own truth and falsity predicate. However, the reason for calling $\mathcal{L}_{TF}$ a language that contains its own falsity predicate is obviously not

*I would like to thank Reinhard Muskens for his valuable comments on this work.

[1] Halbach & Horsten (2006)

[2] In this paper, we will only consider Kripkean fixed point constructions that are carried out starting with an empty extension and anti-extension of the truth predicate; we only consider *minimal* fixed points.

[3] $\mathcal{L}_T = \langle L_T, M, V\rangle$, with V a monotonic valuation scheme and M the partial structure for L_T that is obtained by carrying out the minimal fixed point construction with V.

the semantic intersubstitutability of σ with $F(\ulcorner\sigma\urcorner)$. As '$T$' and '$F$'are both *truth value predicates*, used to express that a sentence has truth value **t** respectively **f**, one may ask for a specification of general conditions that have to be fulfilled for an interpreted language to contain its own truth value predicate(s) in the sense alluded to by Kripke. I take it that a *necessary* condition for a language to contain its own truth value predicate with respect to a certain truth value is that the language is *Truth Value Correct* (TVC) with respect to that value. A language is TVC with respect to truth value **v** just in case whenever a sentence σ has truth value **v**, the sentence which expresses that σ has truth value **v** is true. As $\mathcal{L}_{TF}$ is TVC with respect to **t** and **f**, the question arises whether a language can also be truth value correct with respect to the truth value *ungrounded*.
This paper's model-theoretic result, called the *paradoxical TVC theorem* (partially) answers this question. The theorem states any Δ*-neutral* ground structure can be expanded to a structure for a language L_{TFU}—containing a truth, falsity and Ungroundedness predicate— such that the associated interpreted language $\mathcal{L}_{TFU}$ is truth value correct with respect to **t**, **f** and **u**. In a Δ-neutral structure, we have the ability to form Liar sentences, Yablo sequences (Yablo (1993)) and, in fact, to form any self-referential construction[4] whatsoever using the predicates 'T' and 'F'. However, Δ-neutrality excludes the formation of self-referential constructions which use the ungroundedness predicate 'U'. Sentences like 'this sentence is ungrounded' or 'this sentence is ungrounded or it is false' have no formal counterpart in a Δ-neutral structure.

The paper is organized as follows. Section 2 sets up notation and Section 3 is used to review two theorems, one due to Gupta and one due to Kripke, in light of our notion of truth value correctness. Section 4 is devoted to the proof of the paradoxical TVC theorem, which involves a combination of Kripkean fixed point techniques with *revisionist techniques* and which is inspired by Gupta's proof of a theorem which can be found in Section 3 (Theorem 1) of this paper. In Section 5 we sketch an application of our theorem; we show how the theorem sheds light on the status on interesting—though obscure—argument involving self-referentiality that has recently been advanced in the literature (Rabern & Rabern (2008)) with respect to the *Hardest Logic Puzzle Ever*.

2. Preliminaries

We identify a first order language L with its set of non-logical constants and we assume that '$=$' is a logical constant, expressing the identity relation. With $n \geq 1$, $Con(L)$, $Pred^n(L)$ and $Fun^n(L)$ are used to denote the set of all constant symbols, n-ary predicate symbols and n-ary function symbols of L respectively. $Pred(L)$ and $Fun(L)$ denote the set of all predicate respectively function symbols so that $L = Con(L) \cup Pred(L) \cup Fun(L)$. The set of sentences of L (constructed in the usual manner) will be denoted as $Sen(L)$, its set of closed terms as $Cterm(L)$. A *structure* for L is a pair $M = \langle D, I\rangle$ consisting of a domain D and a function I that interprets L. With $c \in Con(L)$, $f \in Fun^n(L)$ we have $I(c) \in D$ and $I(f) \in D^{D^n}$. With $R \in Pred^n(L)$, $I(R) = (R^+, R^-) \in \mathcal{P}(D^n) \times \mathcal{P}(D^n)$ *such that* $R^+ \cap R^- = \emptyset$.[5] Whenever $R^- = D^n - R^+$

[4]Whether or not a Yablo sequence is a genuine manifestation of self-reference is a controversial issue.

[5]R^+ is called the *extension*, R^- the *anti-extension* of R. Indeed, the definition of a structure in this paper is such that the extension and anti-extension are always disjoint.

for each $n \geq 1$ and $R \in Pred^n(L)$, we say that M is *classical*, otherwise M is *non-classical*. A *valuation scheme* V assigns a function $V_M : Sen(L) \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ to each structure M for L. Here $\{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ is the set of *truth values*; σ can be true (**t**), false (**f**) or *ungrounded* (**u**). The classical (Tarskian) valuation scheme ($\mathcal{C}$), the Strong Kleene scheme (SK), the Weak Kleene scheme (WK) and the Supervaluation scheme (SV) —and only these schemes— we call *appropriate (valuation) schemes*. Note that $\mathcal{C}$ is only defined for classical structures, whereas the other appropriate schemes are defined for all structures. Any appropriate scheme V is *normal*[6] meaning that whenever M is a classical structure for some language L, $V_M(\sigma) = \mathcal{C}(\sigma)$ for all $\sigma \in Sen(L)$. We will use $den_M \subseteq Cterm(L) \times D$ for the denotation relation in structure $M = \langle D, I \rangle$; $\langle t, d \rangle \in den_M$ just in case t denotes d in M. Whenever we write '$\langle t, \sigma \rangle \in den_M$', let it be understood that t denotes a *sentence* σ of the language under consideration.

Definition 1 Quotational closure
Let L be an arbitrary first order language. We set $L^0 = L$ and define:

- $L^{n+1} = L^n \cup \{[\sigma] \mid \sigma \in Sen(L^n)\}, n \geq 0$
- $\bar{L} = \bigcup_{i=0}^{\infty} L^i$

When σ is a sentence of L^n, $[\sigma]$ is a *constant symbol* of L^{n+1}. $\bar{L}$ is the *quotational closure* obtained from L and $\{L^n\}_{n \in \mathbb{N}}$ is the *quotational hierarchy* of $\bar{L}$. Note that $m \leq n \Rightarrow L^m \subseteq L^n$. Any language $\bar{L}$, obtained as the quotational closure of some language L, is called a *quotational language*. □

Definition 2 Sentence structures
A *sentence structure* $M = \langle D, I \rangle$ is a structure for a quotational language $\bar{L}$ such that:

1. $Sen(\bar{L}) \subseteq D$
2. $I([\sigma]) = \sigma$ for all $\sigma \in Sen(\bar{L})$.

Thus the domain of a sentence structure $M = \langle D, I \rangle$ for $\bar{L}$ consists of the sentences of $\bar{L}$ and $\mathcal{O}$*ther objects*. We use $\mathcal{O}_M = D - Sen(\bar{L})$ for the set of non-sentential objects in M's domain. [7] □

When $\bar{L}$ is some quotational language, we use $\mathcal{L}$ to range over all triples $\langle \bar{L}, M, V \rangle$, where M is a sentence structure for $\bar{L}$ and where V is an appropriate scheme *that is defined for* M.[8] When $V = \mathcal{C}$, we say that $\mathcal{L}$ is classical, otherwise, we say that $\mathcal{L}$ is non-classical.[9]

Definition 3 Ground structures and their expansions
Let $\bar{L}$ be a quotational language and let $P \subseteq Pred(\bar{L})$. We say that $\hat{M} = \langle D, I \rangle$ is a *ground structure* for $\bar{L} - P$ just in case $\hat{M}$ is a *classical* structure for $\bar{L} - P$ such that:

1. $Sen(\bar{L}) \subseteq D$

[6] Throughout the paper, 'appropriate scheme' is interchangeable with 'monotonic normal scheme'.

[7] So the sentences themselves —rather than their (Gödel) codes— populate our domain. This approach is not uncommon in the literature. For instance, see Gupta (1982) or Gupta & Belnap (1993).

[8] Thus, when M is non-classical, $V \neq \mathcal{C}$.

[9] Note that a non-classical $\mathcal{L}$ may have classical M.

2. $I([\sigma]) = \sigma$ for all $\sigma \in Sen(\bar{L})$.

When $\hat{M}$ is a ground structure for $\bar{L} - P$, M is an $\bar{L}$-*expansion* of $\hat{M}$ when M is a structure for $\bar{L}$ such that the domains of $\hat{M}$ and M, as well as their respective interpretations of $\bar{L} - P$, are identical. $\square$

3. Truth value predicates

A *truth value predicate* is a predicate that is used to express the assertion that a sentence has a certain truth value; the unary predicate symbols T, F and U will be used to express that a sentence is true, false or ungrounded respectively. With $\mathbf{v} \in \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$, we use $P_\mathbf{v}$ to denote the corresponding truth value predicate, T, F or U respectively.

Definition 4 Truth value correctness
Let L be a language and let $L_{TFU} = L \cup \{T, F, U\}$. $\mathcal{L}_{TFU} = \langle \bar{L}_{TFU}, M, V \rangle$ is said to be *Truth Value Correct with respect to* $\mathbf{v} \in \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$ just in case $TVC_\mathbf{v}$ holds:

$$TVC_\mathbf{v}: \quad V_M(\sigma) = \mathbf{v} \Leftrightarrow V_M(P_\mathbf{v}(t)) = \mathbf{t}, \quad \text{for all } \langle t, \sigma \rangle \in den_M$$

$\mathcal{L}_{TFU}$ is *truth value correct* (TVC) just in case it is truth value correct with respect to each $\mathbf{v} \in \{\mathbf{t}, \mathbf{f}, \mathbf{u}\}$. When $TVC_\mathbf{t}$ ($TVC_\mathbf{f}$, $TVC_\mathbf{u}$) holds, we say that $\mathcal{L}_{TFU}$ *contains its own truth (falsity, ungroundedness) predicate*. $\square$

Note that, for classical $\mathcal{L}_{TFU}$, $TVC_\mathbf{t}$ is equivalent to (1):

$$\mathcal{C}_M(T(t) \leftrightarrow \sigma) = \mathbf{t} \qquad \text{for all } \langle t, \sigma \rangle \in den_M \tag{1}$$

So, for classical $\mathcal{L}_{TFU}$, $TVC_\mathbf{t}$ is equivalent to the truth of all instances of the notorious T-*scheme*. Not every classical $\mathcal{L}_{TFU}$ can contain its own truth predicate. For instance, let $\mathcal{L}_{TFU}$ be such that $\langle \lambda, \neg T(\lambda) \rangle \in den_M$. The sentence '$\neg T(\lambda)$', which intuitively says of itself that it is not true is a (strengthened) *Liar sentence*. Instantiating (1) with the Liar sentence gives '$\mathcal{C}_M(T(\lambda) \leftrightarrow \neg T(\lambda)) = \mathbf{t}$', which is impossible. In order to define an interesting class of classical $\mathcal{L}_{TFU}$ that do contain their own truth predicate, we need to define the notion of X-*neutrality*. Let $M = \langle D, I \rangle$ be a sentence structure—for some $\bar{L}$—and let $X \subseteq D$. With $\vec{d} = \langle d_1, \ldots, d_n \rangle \in D^n$, we say that $\vec{d'} \in D^n$ *is an* X-*swap of* $\vec{d}$ just in case for every $1 \leq i \leq n$ we have that $d_i \notin X \Rightarrow d'_i = d_i$ and that $d_i \in X \Rightarrow d'_i \in X$. We use $X(\vec{d}) \subseteq D^n$ to denote the set of all X-swaps of $\vec{d}$. We are now ready to define the notion of X-neutrality.

Definition 5 X-neutrality
Let $\bar{L}$ be a quotational language, let $M = \langle D, I \rangle$ be a sentence structure for $\bar{L}$ and let $X \subseteq Sen(\bar{L})$. We say that M *is* X-*neutral* just in case, for every $f \in Fun(\bar{L})$ and every $R \in Pred(\bar{L})$ we have that:

1. $\sigma \in X$ and $\langle t, \sigma \rangle \in den_M \Rightarrow t = [\sigma]$
2. $\vec{d} \in R^+ \Leftrightarrow X(\vec{d}) \subseteq R^+, \qquad \vec{d} \in R^- \Leftrightarrow X(\vec{d}) \subseteq R^-$
3. $I(f)(\vec{d}) = I(f)(\vec{d'}) \qquad \text{for all } \vec{d'} \in X(\vec{d})$ $\square$

Gupta (1982) showed that every ground structure $\hat{M}$ for $\bar{L}_T - \{T\}$ —where $L_T = L \cup \{T\}$— that is $Sen(\bar{L}_T)$*-neutral* can be $\bar{L}_T$-expanded, using a *revision process*, to a classical structure M such that the classical $\mathcal{L}_T$ associated with M contains its own truth predicate. However, his results are easily generalizable, delivering the following theorem.

Theorem 1 Non-paradoxical TVC theorem (Gupta)
Let V be an appropriate scheme and let $\hat{M} = \langle D, I\rangle$ be a ground structure for $\bar{L}_{TFU} - \{T, F, U\}$ that is $Sen(\bar{L}_{TFU})$-neutral. There exists a classical $\bar{L}_{TFU}$-expansion M of $\hat{M}$ such that $\mathcal{L}_{TFU} = \langle \bar{L}_{TFU}, M, V\rangle$ is truth value correct.

Proof: See Gupta (1982) or Gupta & Belnap (1993) for a proof in terms of $\mathcal{L}_T$ and carry out the necessary modifications, interpreting U with $(\emptyset, D)$ to obtain a proof for classical $\mathcal{L}_{TFU}$. As the expansion of $\hat{M}$ is classical and as any appropriate valuation scheme is normal, it follows that the theorem in fact holds for any $\mathcal{L}_{TFU}$. □

The reason that I baptized this result of Gupta the *non-paradoxical* TVC theorem is that the conditions for the theorem (i.e. $Sen(\bar{L}_{TFU})$-neutrality) explicitly forbid the formation of well-known "paradoxical" sentences such as the Liar or elements of a *Yablo sequence* (Yablo (1993)). Kripke (1975) showed that, in the presence of such paradoxical sentences, a language can still be truth value correct with respect to **t** and **f**.

Theorem 2 Paradoxical TVC theorem for $\{\mathbf{t}, \mathbf{f}\}$ (Kripke)
Let L be a language, $L_{TF} = L \cup \{T, F\}$ and let $\hat{M}$ be any ground structure for $\bar{L}_{TF} - \{T, F\}$. When V is any non-classical appropriate scheme, there exists a $\bar{L}_{TF}$-expansion M of $\hat{M}$ such that $\mathcal{L}_{TF} = \langle \bar{L}_{TF}, M, V\rangle$ is truth value correct with respect to **t** and **f**.

Proof: See Kripke (1975). □

A language $\mathcal{L}_{TF}$ that is obtained via Theorem 2 declares the Liar sentence to be ungrounded, just as the sentences that ascribe truth or falsity to the Liar sentence. However, the languages $\mathcal{L}_{TF}$ considered by Kripke do not contain an ungroundedness *predicate*, so that an assertion like 'the Liar sentence is ungrounded' has no formal representation in $\mathcal{L}_{TF}$. Thus, the question arises whether a language $\mathcal{L}_{TFU}$ can, in the presence of paradoxical sentences, be truth value correct *tout court*. In the next section, we will prove a theorem that specifies conditions under which the answer to this question is 'yes'. The proof of this "paradoxical TVC theorem" is inspired by Gupta's proof of Theorem 1.

4. The paradoxical TVC theorem

Let L be a first order language and let $\bar{L}_T$ and $\bar{L}_{TU}$ denote the quotational closure of $L \cup \{T\}$ and $L \cup \{T, U\}$ respectively. We let $\Delta = Sen(\bar{L}_{TU}) - Sen(\bar{L}_T)$ and for each $n \in \mathbb{N}$, we let $\Delta_n = \Delta \cap Sen(L^n_{TU})$. The paradoxical TVC theorem will be immediate, once we have proven the following lemma.

Lemma 1 The paradoxical TVC lemma for $\{\mathbf{t}, \mathbf{u}\}$
Let V be a non-classical appropriate valuation scheme and $\hat{M}$ a Δ-neutral ground structure for $\bar{L}_{TU} - \{T, U\}$. Then, $\hat{M}$ can be $\bar{L}_{TU}$-expanded to M such that $\mathcal{L}_{TU} = \langle \bar{L}_{TU}, M, V\rangle$ is truth value correct with respect to **t** and **u**.

The $\bar{L}_{TU}$-expansion referred to in Lemma 1 will be constructed from the ground structure $\hat{M}$ via a two stage process. The first stage, called FP, uses a fixed point construction, the second stage, called RP, uses revisionist techniques.

FP. Let $\hat{M}$ be a Δ-neutral structure and let V be a non-classical appropriate valuation scheme. For any ordinal α, let $M_\alpha = \hat{M}(T^+_\alpha, T^-_\alpha)$ denote the $\bar{L}_{TU}$-expansion of $\hat{M}$ that interprets T as (T^+_α, T^-_α) and that (classically) interprets U as $(\emptyset, D)$. Let $M_0 = \hat{M}(\emptyset, \mathcal{O}_{\hat{M}})$ and, for $\alpha > 0$, define M_α as follows.

$$SUC: \quad \alpha = \beta + 1: \begin{cases} T^+_\alpha = \{\sigma \in Sen(\bar{L}_T) \mid V_{M_\beta}(\sigma) = \mathbf{t}\} \\ T^-_\alpha = \{\sigma \in Sen(\bar{L}_T) \mid V_{M_\beta}(\sigma) = \mathbf{f}\} \cup \mathcal{O}_{\hat{M}} \end{cases}$$

$$LIM: \quad \alpha \text{ is limit}: \begin{cases} T^+_\alpha = \bigcup_{\beta<\alpha} T^+_\beta \\ T^-_\alpha = \bigcup_{\beta<\alpha} T^-_\beta \end{cases}$$

By well-known arguments, the sequence of structures $\{M_\alpha\}_{\alpha \in On}$ has a fixed point, i.e. there exists an ordinal after which further applications of SUC and LIM do not change the resulting structures anymore. We call this fixed point structure M^*.

RP. For any ordinal α, let $M^*_\alpha = \hat{M}(T^+_\alpha, T^-_\alpha, U^+_\alpha)$ denote the $\bar{L}_{TU}$-expansion of $\hat{M}$ that interprets T as (T^+_α, T^-_α) and that interprets U (classically) as $(U^+_\alpha, D - U^+_\alpha)$. We set $M^*_0 = M^*$ and define for each $\alpha > 0$, M^*_α as follows.

$$SUC': \quad \alpha = \beta + 1: \begin{cases} T^+_\alpha = \{\sigma \in Sen(\bar{L}_{TU}) \mid V_{M^*_\beta}(\sigma) = \mathbf{t}\} \\ T^-_\alpha = \{\sigma \in Sen(\bar{L}_{TU}) \mid V_{M^*_\beta}(\sigma) = \mathbf{f}\} \cup \mathcal{O}_{\hat{M}} \\ U^+_\alpha = \{\sigma \in Sen(\bar{L}_{TU}) \mid V_{M^*_\beta}(\sigma) = \mathbf{u}\} \end{cases}$$

$$LIM': \quad \alpha \text{ is limit:} \begin{cases} T^+_\alpha = \{\sigma \in Sen(\bar{L}_{TU}) \mid \exists\beta : \sigma \in \bigcap\limits_{\beta<\gamma<\alpha} T^+_{M^*_\gamma})\} \\ T^-_\alpha = \{\sigma \in Sen(\bar{L}_{TU}) \mid \exists\beta : \sigma \in \bigcap\limits_{\beta<\gamma<\alpha} T^-_{M^*_\gamma}\} \\ U^+_\alpha = \{\sigma \in Sen(\bar{L}_{TU}) \mid \exists\beta : \sigma \in \bigcap\limits_{\beta<\gamma<\alpha} U^+_{M^*_\gamma}\} \end{cases}$$

In order to prove Lemma 1, we will need the following lemma.

Lemma 2 Stabilization lemma
Let V be a non-classical appropriate valuation scheme. Let $\hat{M}$ be a Δ-neutral structure for $\bar{L}_{TU} - \{T, U\}$ and let $\{M^*_\alpha\}_{\alpha \in On}$ be the series of structures generated from $\hat{M}$ via FP and RP. Then, for all $n \in \mathbb{N}$ and $\alpha \in On$ such that $\alpha \geq n + 1$

$$\sigma \in Sen(L^n_{TU}) \Rightarrow V_{M^*_{n+1}}(\sigma) = V_{M^*_\alpha}(\sigma) \tag{2}$$

Proof: Suppose that the lemma is false. Then there has to be a least natural number, say n' for which it fails and, given n' there has to be a least ordinal $\geq n' + 1$, say α' such that $M^*_{n'+1}$ and $M^*_{\alpha'}$ disagree about the truth value of $\sigma \in Sen(L^{n'}_{TU})$. Thus from the hypothesis that the lemma is false and the minimality of n' and α' we get:

$\mathbf{C_1}$: For all $n < n', \alpha \geq n+1, \sigma \in Sen(L^n_{TU}) : V_{M^*_{n+1}}(\sigma) = V_{M^*_\alpha}(\sigma)$
$\mathbf{C_2}$: For all α s.t. $n'+1 \leq \alpha < \alpha', \sigma \in Sen(L^{n'}_{TU}) : V_{M^*_{n'+1}}(\sigma) = V_{M^*_\alpha}(\sigma)$
$\mathbf{C_3}$: $\exists \sigma \in Sen(L^{n'}_{TU}) : V_{M^*_{n'+1}}(\sigma) \neq V_{M^*_{\alpha'}}(\sigma)$

We will show that these 3 conditions can not (jointly) hold, contradicting the hypothesis of the falsity of the lemma. From the definition of RP it follows that α' has to be a successor ordinal, say $\alpha' = \beta + 1$. The structures $M^*_{n'+1}$ and $M^*_{\beta+1}$ only differ with respect to the interpretation of the predicate symbols T and U. By definition of RP, these interpretations are fully determined by the functions $V_{M^*_{n'}}(\cdot)$ and $V_{M^*_\beta}(\cdot)$ respectively. As these functions valuate $\aleph_0$ different sentences of $\Delta - \Delta_{n'-1}$ to be true (false)[10], there exists a bijection $\chi : \Delta - \Delta_{n'-1} \rightarrow \Delta - \Delta_{n'-1}$ such that —with $X = T^+, T^-$ or U^+:

$$\forall \sigma \in (\Delta - \Delta_{n'-1}) : \sigma \in X_{n'+1} \Leftrightarrow \chi(\sigma) \in X_{\alpha'} \tag{3}$$

We extend χ to a bijection[11] from D to D, by specifying that χ acts as the identity function on objects in $D - (\Delta - \Delta_{n'-1})$. We will show that χ is an isomorphism between the structures $M^*_{n'+1}$ and $M^*_{n'+\alpha'}$ in the language $L^{n'}_{TU}$. From the fact that isomorphic structures in a language are elementary equivalent w.r.t. the sentences of that language, it then follows that there cannot be a $\sigma \in Sen(L^{n'}_{TU})$ such that $M^*_{n'+1}$ and $M^*_{n'+\alpha'}$ disagree about the truth value of σ. Hence, we establish a contradiction with $\mathbf{C_3}$ and, consequently, with the hypothesis that the lemma is false. By definition of an isomorphism between structures, in order to show that χ is an isomorphism between $M^*_{n'+1}$ and $M^*_{\alpha'}$ in the language $L^{n'}_{TU}$, we need to establish that, for every $n \in \mathbb{N}$ and $\langle d_1, \ldots, d_n \rangle \in D^n$:

1. For every $R \in Pred^n(L^{n'}_{TU})$:
   $\langle d_1, \ldots, d_n \rangle \in R^+_{n'+1}$ iff $\langle \chi(d_1), \ldots, \chi(d_n) \rangle \in R^+_{\alpha'}$
   $\langle d_1, \ldots, d_n \rangle \in R^-_{n'+1}$ iff $\langle \chi(d_1), \ldots, \chi(d_n) \rangle \in R^-_{\alpha'}$

2. For every $f \in Fun^n(L^{n'}_{TU})$:
   $\chi(I(f)(d_1, \ldots, d_n)) = I(f)(\chi(d_1), \ldots, \chi(d_n))$

3. For every $c \in Con(L^{n'}_{TU})$: $\chi(I(c)) = I(c)$

ad 1. When $R \notin \{T, U\}$, the claim readily follows from the fact that χ acts as the identity function on $D - \Delta$ and that $\hat{M}$ is Δ-neutral structure. So let $R \in \{T, U\}$. Observe that $d \in D$ implies that d is either an element of $\mathcal{O}_{\hat{M}}$, Δ_{n-1}, $\Delta - \Delta_{n'-1}$ or $Sen(\bar{L}_T)$. When $d \in \mathcal{O}_{\hat{M}}$, the claim follows from the fact that $d \in T^-_{n'+1} \cap T^-_{\alpha'}$ and that χ acts as the identity function on $\mathcal{O}_{\hat{M}}$. When $d \in \Delta_{n'-1}$ the claim follows from $\mathbf{C_1}$ and the definition of RP. When $d \in \Delta - \Delta_{n'-1}$ the claim follows from (3). Finally, let $d \in Sen(\bar{L}_T)$. Observe that, as M^* results from FP, M^* is a fixed point structure "with respect to the sentences of $\bar{L}_T$", i.e.:

$$V_{M^*}(\sigma) = V_{M^*_\alpha}(\sigma) \text{ for all } \alpha \in On, \sigma \in Sen(\bar{L}_T) \tag{4}$$

[10]Let $\Delta_{-1} = \emptyset$

[11]Note that the existence of this bijection depends *only* on the fact that M^*_0 valuates $\aleph_0$ sentences as **t** and $\aleph_0$ sentences as **f**, which make $V_{M^*_{n'}}(\cdot)$ and $V_{M^*_\beta}(\cdot)$ do so too. M^*_0 either valuates $\aleph_0$ or 0 sentences as **u** and so $V_{M^*_{n'}}(\cdot)$ and $V_{M^*_\beta}(\cdot)$ respectively valuate either $\aleph_0$ or 0 sentences as **u**. In both cases, cardinality considerations show that the function χ satisfying (3) can be found and so the sought for bijection exists.

From (4) and the definition of RP it follows, —with $X = T^+, T^-$ or U^+— that:

$$\sigma \in X_{M_1^*} \Leftrightarrow \sigma \in X_{M_{1+\alpha}^*} \text{ for all } \alpha \in On, \sigma \in Sen(\bar{L}_T) \tag{5}$$

Now the claim follows from (5) and the fact that χ acts as the identity function on $Sen(\bar{L}_T)$.
ad 2. The claim follows from the fact that M is an Δ-neutral structure and that χ only permutes elements of Δ.
ad 3. When c denotes an element $d \notin \Delta$, the claim follows from the fact that $\chi(d) = d$. Whenever $c \in Con(L_{TU}^{n'})$ denotes an element of Δ, Δ-neutrality of $\hat{M}$ guarantees that it denotes an element of $\Delta_{n'-1}$, on which χ also acts as the identity function. $\square$

Lemma 1 and, in fact, the paradoxical TVC theorem now follow easily.

Theorem 3 The paradoxical TVC theorem
Let L be a first order language, $L_{TFU} = L \cup \{T, F, U\}$ and $L_{TF} = L \cup \{T, F\}$. Let $\bar{\Delta} = Sen(\bar{L}_{TFU}) - Sen(\bar{L}_{TF})$. Let V be a non-classical appropriate valuation scheme and let $\hat{M}$ be a $\bar{\Delta}$-neutral ground structure for $\bar{L}_{TFU} - \{T, F, U\}$. Then, $\hat{M}$ can be $\bar{L}_{TFU}$-expanded to a structure M such that $\mathcal{L}_{TFU} = \langle \bar{L}_{TFU}, M, V \rangle$ is truth value correct.

Proof: Apply FP and RP, modified in the obvious way, to expand $\hat{M}$ by filling the extensions and anti-extensions of T, F and U. From the (modified) Stabilization Lemma it follows that the generated series of structures $\{M_\alpha^*\}_{\alpha \in On}$ has a fixed point at ω, i.e. $M_\omega^* = M_{\omega+1}^*$. From the definition of (modified) RP, it now immediately follows that $\mathcal{L}_{TFU} = \langle \bar{L}_{TFU}, M_\omega^*, V \rangle$ is truth value correct. $\square$

5. The Tempered Liar Lemma and the paradoxical TVC theorem

In this section, the paradoxical TVC theorem will be applied to shed light on the status of a proof which appeared in Rabern & Rabern (2008). The background of their proof is the so called *Hardest Logic Puzzle Ever* ($HLPE$). See Boolos (1996), Roberts (2001) or Rabern & Rabern (2008) for a discussion of this puzzle. We will only discuss the "sub puzzle" of $HLPE$ that Rabern and Rabern (R&R) claim to solve with their proof.

The puzzle. Suppose that there is an object, o, that is either black all over, yellow all over or red all over. You do not know o's color but there is a god, True, who knows o's color and who answers all and only yes-no questions *truthfully*. What is the minimum number n of yes-no questions that you have to ask to True in order to determine o's color with certainty? One may reason as follows. First asking whether o is black and then whether o is yellow shows that $n \leq 2$ and as *obviously*, $n \neq 1$ we have $n = 2$. However, R&R give a proof, in natural language, that claims to show that this appeal to our "$n \neq 1$-intuitions" is unjustified; they claim to prove that $n = 1$. The statement that $n = 1$ will be called the *Tempered Liar Lemma* (TLL) and the question by which R&R claim to establish TLL is Q, in which 'this' refers to the question as a whole.[12]

Q : Is it the case that (your answer to this question is 'no' and o is black) or o is yellow?

[12] I will be sloppy in not distinguishing between a yes-no question and its associated declarative sentence; for instance I will speak of the truth-value of a yes-no question.

The essential idea involved is that on Liar like questions such as 'your answer to this question is 'no'?', True cannot answer with either 'yes' or 'no' without lying and shows a different reaction accordingly, say that True *explodes*. R&R, argue that if Q is answered with a) 'yes' then o is yellow, b) 'no' then o is red while c) an explosion indicates that o is black.

The proof. The three material implications a), b) and c), are established via *reductio ad absurdum* as follows.
a) Assume that True answers 'yes' and that o is not yellow. Then True says 'yes' to the left disjunct of Q and so in particular to 'your answer to this question is 'no''. This is impossible as True tells the truth.
b) Assume that True answers 'no' and that o is not red. Then, as True answered 'no' to Q, he denies the left and the right disjunct of θ, from which it respectively follows that o is not black and that o is not yellow and so o is red. Contradiction.
c) Assume that True explodes and o is not black. Then o is not yellow either, for otherwise True would answer 'yes'. Hence, as o is neither black nor yellow, True denies both disjuncts of Q and hence answers Q with 'no'. Contradiction.

The paradox. This argument of R&R is—though interesting—obscure, for nowhere in Rabern & Rabern (2008) are the principles by which True reasons specified. At first sight—at least to me—the proof looks fine. But consider the following argument to the conclusion that True does not explode on Q which is, so it seems, obtained by the same principles as those implicit in R&R's proof. Suppose that o is black and that True explodes on Q. Then, the left disjunct of Q is false (as True does not answer 'no'), and so Q is false (as o is black the second disjunct is also false) and hence True should answer Q with 'no'! Also, what would happen if we asked True: 'is the case that your answer to this question is 'no' or that you explode on this question?' Such *strengthened Liar objections* show that R&R's proof is suspect, to say the least, and that an explanation of the assumptions involved is needed.

The truth value-answer link. A possible defense against such strengthened Liar objections is that they are based on a wrong conception of "how True works". For instance, True may answer with 'yes'or 'no' iff answering 'yes' or 'no' is truthful *and* if True can do so *without contradicting himself* and otherwise, True explodes. Such an 'inferential conception' of True may be combined with the thought that if o is black, True explodes on Q and that this explosion renders Q false; the inferential conception of True then gives up the link between the truth value of a sentence and True's answer to it.[13] In contrast, the paradoxical TVC theorem can be seen as a specification of the conditions under which one can make sense of the argument of R&R when the answer of True to σ is understood as a reaction to the truth value of σ. We give two distinct ways to do so, called the metalanguage approach and the object-language approach respectively.

[13]In personal communication Brain Rabern suggested, as a reaction to my strengthened Liar objections, an inferential conception of True which gives up the link between the truth value of σ and True's answer to it. In Wintein (2009) I develop a formal approach to capture such an inferential conception of True. There an answer of True to a sentence σ is determined by the outcome of an *inferential process* of True which takes σ as input. In this formalization of an 'inferential True' one can, due to the assumption that True reacts differently to Truthtellers than to Liar sentences, also determine the color of an object which has 1 out of 4 possible colors by asking a single question to True.

The meta-language approach. Assuming a link between the truth value of σ and True's answer to it, the interpretation of '$F(x)$' as 'x is false' is extensionally equivalent to its interpretation as 'True's answer to x is 'no''. Modulo this shift of interpretation, question Q can be represented via a constant θ and an interpretation function I as follows.

$$I(\theta) = (F(\theta) \wedge B(o)) \vee Y(o) \tag{6}$$

Let L_{TFU} be a language that contains, besides the three truth value predicates (equivalently, "answering predicates") the color predicates B, Y and R and the constants θ and o. Let your ignorance about the color of the object be represented by $K \in Sen(L_{TFU})$:

$$K := (B(o) \wedge \neg Y(o) \wedge \neg R(o)) \vee (\neg B(o) \wedge Y(o) \wedge \neg R(o)) \vee (\neg B(o) \wedge \neg Y(o) \wedge R(o))$$

Any Δ-neutral ground structure for $L_{TFU} - \{T, F, U\}$ which interprets θ as (6), which interprets o with a non-sentential object and which interprets the color predicates such that K is valuated as **t** we call a *K-ground structure* and the L_{TFU}-expansion of a K-ground structure via the *Strong Kleene version* of the paradoxical TVC theorem construction, we call a *possible world*. Note that, corresponding to the three possible colors of the object, the class of possible worlds $\mathcal{M}$ allows for a tripartition. A possible way to give a valid reconstruction of R&R's argument is to understand them as reasoning in a classical meta-language about $\mathcal{M}$. We define $\models_{\mathcal{M}} \subseteq \mathcal{P}(Sen(L_{TFU})) \times Sen(L_{TFU})$ by stipulating that $\Delta \models_{\mathcal{M}} \sigma$ just in case in every $M \in \mathcal{M}$ in which all members of Δ are valuated as **t**, σ is also valuated as **t**. Neglecting parenthesis for singleton sets, the three claims of R&R may be translated as follows:

a') $T(\theta) \models_{\mathcal{M}} Y(o)$ $\quad$ b') $F(\theta) \models_{\mathcal{M}} R(o)$ $\quad$ c') $U(\theta) \models_{\mathcal{M}} B(o)$

As the reader may verify, a'), b') and c') are true, while the associated object-language counterparts of, a') and b') in terms of material implication do not hold. For instance, we do not have that $\models_{\mathcal{M}} T(\theta) \rightarrow Y(o)$; in a world in which the object is black, '$T(\theta)$' is valuated as **u**, '$Y(o)$' as **f** and hence the material implication as **u**.

The object-language approach. If we slightly alter R &R's natural language claims, we can have a correct object language representation of those claims. For observe that the fact that the ungroundedness predicate is a classical predicate gives us:

a'') $\models_{\mathcal{M}} (\neg U(\theta) \wedge T(\theta)) \rightarrow Y(o)$
b'') $\models_{\mathcal{M}} (\neg U(\theta) \wedge F(\theta)) \rightarrow R(o)$
c'') $\models_{\mathcal{M}} U(\theta) \wedge \rightarrow B(o)$

Conclusion. We used the paradoxical TVC theorem to give a rough sketch of two possible reconstructions of the reasoning of R&R. Although a lot more can be said about the details of both reconstructions, I do not think that either of them can be fruitfully converted into a genuine proof of TLL, the reason being that the condition of Δ-neutrality is too restrictive. We would like to know the principles by which True answers questions as 'do you explode on this question?' and the like, which are excluded by Δ-neutrality. Be that as it may, the paradoxical TVC theorem itself is a nice little result which can be added to our ever growing stock of truths about truth.

References

G. Boolos (1996). 'The hardest logic puzzle ever'. *The Harvard Review of Philosophy* **6**:62–65.

A. Gupta (1982). 'Truth and Paradox'. *Journal of Philosophical Logic* **11**:1–60.

A. Gupta & N. Belnap (1993). *The Revision Theory of Truth*. MIT Press, Cambridge.

V. Halbach & L. Horsten (2006). 'Axiomatizing Kripke's theory of truth'. *Journal of Symbolic Logic* **71**:677–712.

S. Kripke (1975). 'Outline of a Theory of Truth'. *Journal of Philosophical Logic* **72**:690–716.

B. Rabern & L. Rabern (2008). 'A simple solution to the hardest logic puzzle ever'. *Analysis* **68**:105–112.

T. Roberts (2001). 'Some thoughts about the hardest logic puzzle ever'. *Journal of Philosophical Logic* **30**:609–612.

S. Wintein (2009). 'The computational power of self-referential truth'. *Submitted* .

S. Yablo (1993). 'Paradox without self-reference'. *Analysis* **53**:251–252.

Characterization of Conversational Activities in a Corpus of Assistance Requests

François Bouchet
LIMSI-CNRS, Université Paris-Sud XI

Abstract. Modeling linguistic interaction is a crucial point to provide assistance to ordinary users interacting with computer-based systems and services. A first issue is more particularly the characterization of the linguistic phenomena associated with the Function of Assistance, in order to define an assisting rational agent capable of pertinent reactions to users' requests. In this paper, we present a corpus based on users' requests registered in actual assisting experimentations. First we compare it to similar task-oriented dialogical corpora through a method based on interactional profiles (speech acts oriented analysis) that assesses its specificity. Then we show through an annotation of conversational activities that the collected corpus is heterogeneous (40% of its requests are actually not task-oriented), and use again the interactional profiles approach to go towards an automatic identification of those activities.

1. Introduction

1.1. Assisting Ordinary Users

The number of novice (or ordinary) users[1] of computer applications and services has been quickly increasing for the past decades and is likely to keep growing for some time. An example of such users could be the average Net surfer swapping between websites in order to use sporadically web-services (buying an airplane ticket. . .) or producing personal web-content (sharing photos, videos. . .). Because of the complex nature of the new websites and commercial applications, these users will inevitably face difficulties to achieve their objectives. That situation can lead users into a cognitive distress, with a significant negative impact for the application providers.
Natural language (NL) has been shown to be an ideal way to provide assistance to novice users interacting with computer applications or services. First, it appears that it is a modality spontaneously used when a problem arises, particularly (but not exclusively) in the case of novice users, and which closely reflects the cognitive processes of the user (Ericsson & Simon 1993), but the use of multimodality for assistance also allows a clear cognitive separation between the application and the help system (with graphical (Morrell & Park 1993) or dialogical (Amalberti 1996) modalities). But since these situations arise mainly because of the lack of knowledge about the software, it leads to many linguistic difficulties such as: user-specific vocabulary, degraded spelling (for typed requests) or degraded prosody (for oral requests), bad categorizations, *etc.* It thus makes this type of requests really difficult to process and to interpret.

1.2. Characterizing the Function of Assistance

In order to bring an answer to this need of assistance, the DAFT project[2] from LIMSI-CNRS (Sansonnet, et al. 2005) intends to develop Assistant Conversational Agents (ACA)

[1] By ordinary user, we mean a non expert person using a software application.
[2] http://www.limsi.fr/~jps/research/daft/

able to analyze Natural Language requests expressed without constraints by ordinary users during their use of applications of increasing complexity (applets, webpages, active websites, text processor…). This choice is also motivated by the additional benefits brought by the use of an embodiment for dialogue system in terms of trust and believability, a phenomenon known as the 'Persona Effect' (Lester, et al. 1997).

The objectives of the data processing sequence of DAFT assistance system is to characterize generic components of the Function of Assistance, and to propose a rational agent engine able to assist ordinary users in the most frequent cases. The global architecture of the system is classically made of:

- a semantic analysis module of NL requests to build formal requests,
- a module of reasoning on the application model that returns a formal request,
- a production module aiming to express the agent's answer in a multimodal way (spoken or written way, action on the application, gestures from an animated virtual character…).

As a preliminary step towards the creation of this NL processing chain, we have chosen a corpus-based approach to study the phenomena we had to model, leading to three questions: is our corpus original (*i.e.* different from existing ones) and how to prove it? What kind of conversational activities does it actually contain? And how is it possible to distinguish them automatically?

In a first part of this article, we describe how that corpus has been built up and give an excerpt of collected requests. In a second part, we introduce the concept of interactional profiles to compare, in terms of speech acts, our corpus to similar ones. Finally, we show in section 4 that our corpus actually contains requests representing different conversational activities and apply again the interactional profile method to compare those subcorpora.

2. Corpus collection and building

2.1. Methodology

Currently, very few public data is actually available concerning Human-Computer dialogue. Moreover, our scope is rather different from classical dialogue systems and actually closer from Natural Languages Interfaces (Androutsopoulos & Aretoulaki 2003), since we're dealing with *isolated requests* rather than with *dialogic sessions*. On top of this, it was crucial for us to control precisely the experimental conditions for the collection of requests which had to deal specifically with the assistance. For all those reasons[3], we have chosen to collect our own specific corpus, which we'll refer to as the Daft corpus in this article.

It has been built up from three different sources (each providing roughly a third of the total corpus):

1. During two years, about 100 human subjects have been faced with several applications assisted by the Daft system (in its 1.0 version): three Java applets (modal and

[3]The validity of that difference hypothesis will be further investigated in section 3.

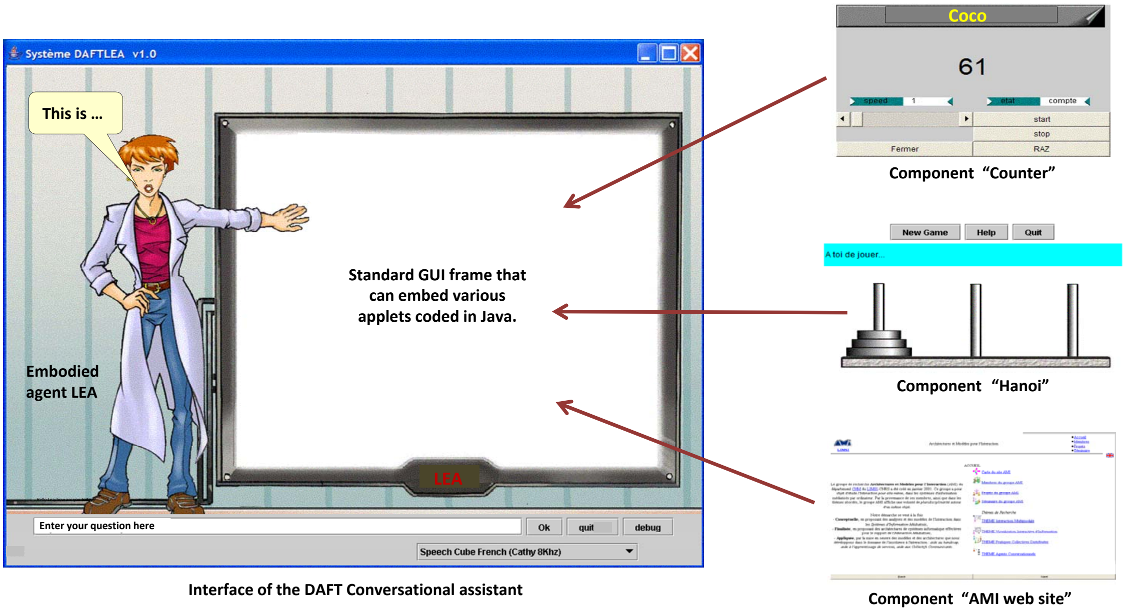


Figure 1: The conversational agent interface into which were embedded different kind of applications (Java applets or dynamic websites)

unmodal, *i.e.* with threads), two websites (one was active, *i.e.* could be dynamically edited by users – *cf.* figure 1 for an example of interface);

2. From two thesauri (Molinsky & Bliss 1994, Atkins & Lewis 1996), we have manually constructed some requests in order to provide a wider linguistic coverage of the assistance vocabulary and idioms;

3. Recently, we have added to the corpus some FAQ extracted from integrated help systems and websites concerning two widely used document creation softwares (LATEX and Microsoft Word).

2.2. General view of the Daft corpus

Table 1 shows excerpts from the Daft corpus (currently made of 11.000 requests), emphasizing some characteristics:

- more than half of the users' sentences are *not well-formed* (spoken expressions, spelling/syntactic/grammatical mistakes, acronyms from SMS and internet slang. . .), and some mistakes are not that easy to detect and correct with classical NLP tools;

- requests are *not* stored as part of a dialogue but as isolated sentences. As mentioned previously, it appeared as suggested by (Capobianco & Carbonell 2002) that in the domain of assistance, dialogic interactions are almost always limited to a single conversational turn and hence can be dealt with as *isolated requests*.

N°	Original request	Translated request
1	**appuies** sur le bouton quitter	**clicks** on the quit button
2	**clickersur** le bouton back	**clickon** the back button
3	bon, reviens à **l apage** d'accueil	ok, come back to **th ehomepage**
4	**a** quoi sert cette fenêtre**,**	what is this window for**,**
5	**c quoi** le GT ACA	**WDYM** by GT ACA
6	le bouton "fermer" et le bouton "quitter" ont exactement le même fonctionnement?	do the "close" button and the "quit" button work exactly the same way?
7	je ne **vosi** aucune page de **demso** !!	I **cna't** see any **demso** page!!
8	j'ai été surpris qu'il manque une fonction d'annulation globale	I was really surprised to see a global cancel function is missing
9	ça serait mieux si on pouvait aller directement au début	it'd be better to be able to go directly at the beginning
10	auf viedersen	**auf viedersen**
11	bon à rien !	you good-for-nothing!
12	Quel genre de musique tu aimes ?	What kind of music do you like?
13	ca marche **:-)**	works for me **:-)**
14	j'aime tes cheveux Léa	I like your hair Lea

Table 1: Original and translated examples from the Daft corpus – mistakes, in bold, have been translated as closely as possible from the original French version

3. Corpora comparison

We have seen in 2.1. that our initial hypothesis for the corpus collection was that our domain of study was specific enough to prevent us from using requests from similar existing corpora. We thus focus in this section on the validation of that hypothesis through an analysis of speech acts in different corpora annotated with different taxonomy.

3.1. Data: corpora to compare

We have chosen for this comparative study three reference corpora of annotated task-oriented dialogues:

- Switchboard (Jurafsky, et al. 1998): 200.000 manually annotated utterances from task-oriented phone talks;
- MapTask (Carletta, et al. 1996): 128 dialogues in which one person has to reproduce a route on a map, following instructions from another person with a similar map;
- Bugzilla (Ripoche 2006): 1.200.000 comments from 128.000 bug reports created during the development of the Mozilla Foundation's suite.

Switchboard and MapTask are coming from oral interactions and hence are naturally richer in words than written corpora (Kelly & Chapanis 1977), but the closeness of activities appeared more important for our comparison than this origin difference. As for the language difference (those corpora being in English whereas ours is in French), it is probably not significant in terms of speech acts. Besides, although some French (oral as well) corpora could also be relevant for this comparison, like Air France (Morel 1989), Ozkan (Ozkan 1994) or even a small French MapTask corpus (Post 2000), they aren't provided with a speech act taxonomy and annotations as it is the case the three ones aforementioned.

3.2. Methodology: interactional profiles and taxonomy mapping

Interactional profile is defined in (Ripoche 2006) as "the distribution of speech acts appearing in a given interactional unit". The interactional unit itself is nothing but a coherent set of speech acts chosen according to the analysis objective: a single utterance, a dialogue, a corpus, *etc.* Once the interactional unit has been chosen, the ratio of each speech act in this unit is calculated, and the profile itself is generally displayed as an histogram in order to have a synthetic view associated with the class of interaction. The main interest of interactional profiles is not as much their intrinsic value as the possibility they offer to allow *comparison* between two different classes of interactions.
This approach is fundamentally close from the model developed in (Twitchell, et al. 2004) but has some noticeable differences though:

- since for more accuracy we prefer to manually annotate the speech acts, our approach is discrete rather than probabilist (Twitchell et al. (2004) allowing elements of an unit to belong to many speech acts with a probability function);
- our approach is conceptually more collective (to study a global behaviour) than individual (study of one person's interactions), although both methods can certainly be used in both contexts;
- we consider the interactional profiles defined as having an absolute value, whereas Twitchell et al. (2004) subtract to it a global average profile of interactions supposed to have a normative value.

Here, the interactional unit chosen is the corpus as a whole, and the speech acts set for the shared taxonomy is made of the five classical searlian speech acts (Searle 1969), which are generic enough to allow comparison. We thus had to map the existing taxonomies used to annotate those corpora into that common one.
In the case of Switchboard, although the original annotation was done along four dimensions, it has appeared that combinations between those dimensions were rather limited, allowing to distinguish a total of 42 main categories in the DAMSL taxonomy (Jurafsky et al. 1998), which are the ones we have been using here. Some speech acts being very specific (for instance Switchboard's "self-talk") are however not trivial to convert even into a very general taxonomy as the one chosen. Similarly, although annotated at different levels, the speech acts from MapTask can be considered in a flatten way for this mapping, as displayed in table 2.

3.3. Results

The interactional profiles of those four corpora are diplayed on figure 2. Because of the impossiblity to have a perfect mapping explained in the previous section, their interpretation needs to be done cautiously. Nevertheless, some very distinct characteristics seem to distinguish the Daft corpus from the three others, namely:

Searle	Assertive	Commissive	Directive	Expressive	Declarative	Unknown
MapTask	clarify explain reply-w	-	align check instruct query-w query-yn	acknowledge ready reply-n reply-y	-	uncodable

Table 2: Speech acts mapping for the MapTask taxonomy

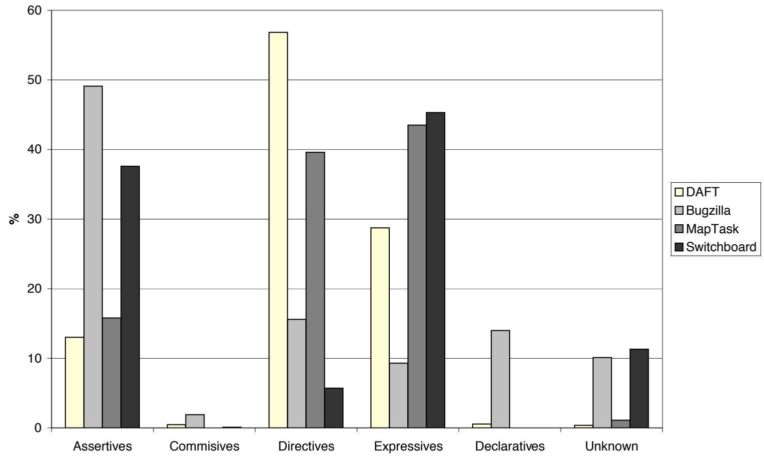


Figure 2: Interactional profiles comparison of four task-oriented dialogue corpora

- a majority of *directives* (57%), explained by the high number of orders or questions to the agent. Although other corpora are also task-oriented, interactions were only between humans, and it seems likely that talking to a computer (even through an ECA) make requests more direct, as users generally don't expect the agent to be able to make complex inferences.

- the number of *assertives* is rather low (13%), users prefering to express their feelings and states of mind (29%) concerning the situation rather than those same facts in a neutral and "objective" way as they do for example in the Bugzilla corpus, since not knowing how to do something is considered as a rather stressing situation.

- very few *commissives* (1%) are observed, which can be easily explained by the relationship user-agent: if the agent is in essence subordinate to the user, the latter rarely feel commited to do whatever the agent can suggest to him (even when the answer is perfectly relevant).

The use of conversational profiles for this comparison clearly helped to demonstrate the differences existing between those corpora. More particularly, it confirmed that the linguistic domain covered by the Function of Assistance to novice users required a dedicated corpus.

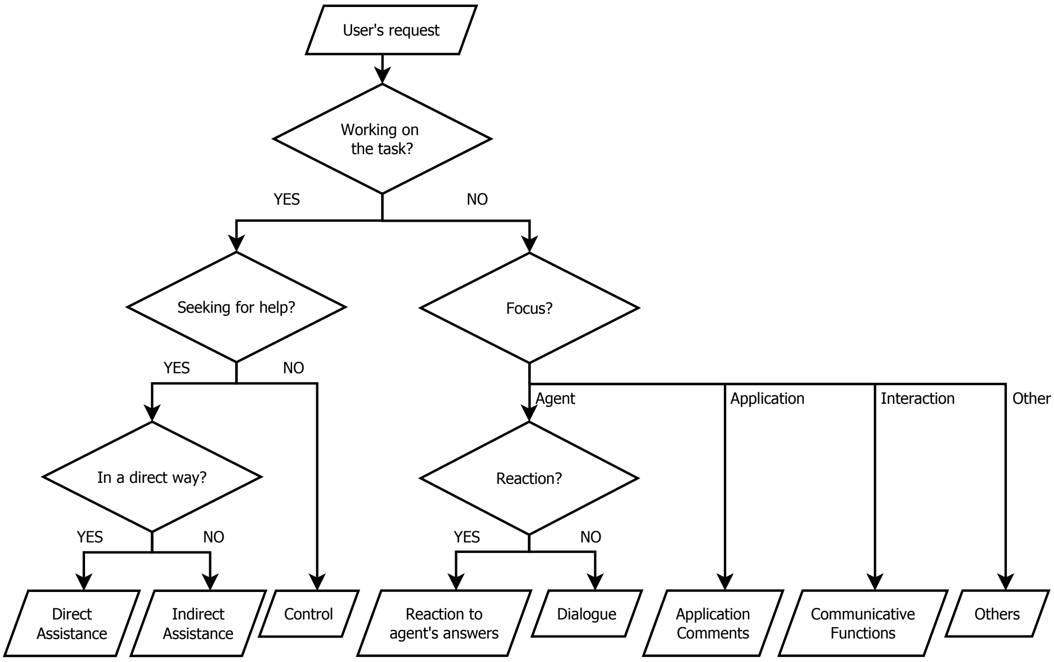


Figure 3: Requests activity classification protocol

4. Conversational activities analysis

During the corpus collection phase, human subjects were informed that they had to do some tasks for which they could ask help (if needed) from an artificial assistant agent embedded in the program to assist them. Nevertheless, subjects were completely free to act and particularly they could type what they want without any constraint, and various behaviours were observed, users having sometimes completely abandoned their original task. Eventually, it appeared that many of the collected sentences were not really linked to the assistance domain (cf. table 1). Hence we got interested in trying to identify and categorize those other conversational activities that were appearing in the corpus.

4.1. Methodology: conversational activity annotation protocol

For this purpose, we have randomly extracted from the actually collected part of the Daft corpus (*i.e.* not the manually built up parts mentioned in 2.1.) two subsets of sentences, each subset having a size equal to the tenth of the total corpus size. The two subsets have been manually annotated by a single annotator, one after another in time. The first annotation process was used to refine the protocol (described below and summarized on figure 3), whereas the second one was strictly following it. Keeping in mind our objective is to study assistance, the first step was to know if the user was seeking help to accomplish tasks through its request, thus defining a first high level granularity distinction between:

1. **task-oriented** activities (ex: sentences[4] 1 – 9): where the user is working on the application to go towards the goal he has been given (independantly from knowing if he actually succeeds in getting closer from that goal).

[4]Sentence numbers always refer to the request examples given in table 1

2. **chat-oriented** activities (ex: sentences 10 – 14): where the user is interacting with the system for a reason that is not directly relevant to accomplish the task.

In the case of task-oriented requests, the user is either working directly on accomplishing its task or trying to get help from the system in order to do so. And in that former case, his request for help appears more or less obvious, thus leading us to distinguish three distinct activities:

1. **control** (sentences 1 – 3): direct *controls*, to make the agent interact himself directly with the application software in which it is embedded.

2. **direct assistance** (sentences 4 – 6): *help requests* explicitly made by the user.

3. **indirect assistance** (sentences 7 – 9): user's judgements concerning the application that would lead a human being to interpret them as a need for assistance; it certainly requires the system to use pragmatics to detect the implicit meaning.

For chat-oriented requests, although they are less relevant to our objective, we have been categorizing them according to the element of focus of the user's in its requests, distinguishing, in the cases where it is focused on the agent itself, replies to a previous agent utterance from a chat interaction started by the user. That makes a total of five different subactivities:

1. **reactions** to an agent's answer: a set of ways to agree or disagree to the agent's answer, marks of incredulity ("I don't think so"), lack of understanding ("You lost me") or insistence ("please answer to me").

2. **communicative functions**: this set is made of forms to start or end the communication with the agent ("hello", "bye", "I don't need your help anymore"...) as well as some phatic acts ("are you there?").

3. **dialogue** with the agent: sentences where the agent becomes user's focus, from orders ("Shut up!") to questions ("do you have a soul?") and from threats ("don't force me to kill you") to compliments ("you look cute").

4. **comments** about the application: comments without any assistance value ("This page looks nice").

5. **others**: a mix of the rest of the chat requests, not easy to classify with more details ("I'm an ordinary user", "I want to do a cognitive gift"...).

4.2. Results

No significant differences have been observed between both subsets annotated, allowing us to generalize the results obtained to the rest of the collected corpus : the figures 4 and 5 show the average distribution of requests from both subsets. Focusing on assistance, we can consider our collected corpus can be divided into four "subcorpora", each corresponding to a particular activity: control, direct and indirect assistance and chat.

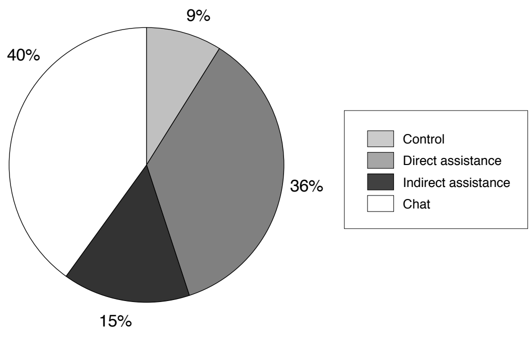


Figure 4: Daft corpus requests distribution of conversational activities

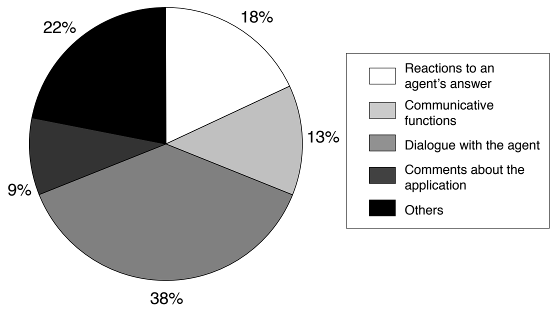


Figure 5: Chat subcorpus detailed distribution of conversational activities

The existence of the control subcorpus demonstrates that the user not only expects the agent to be able to assist him to use the application, but he also wants him to be able to act on this very application. The same way, the relative importance of chat-oriented requests, certainly related to the use of an embodied agent (*cf.* figure 1), shows that the user wants as well an agent able to react to comments not related to the task he is trying to carry out. Nonetheless, to really be able to deal properly with chat requests would require much more advanced dialogue skills: a wide range of vocabulary, personal life facts, an opinion about virtually anything, *etc.* We would thus be losing the methodological cut in complexity intended by focusing on a subdomain of natural language in the case of assistance. Finally, for the conception of our assisting agent, we take only into account control, (direct and indirect) assistance and reactions to an agent's answer activities, since that latter doesn't call into question the choice of dealing with isolated requests as it can easily be treated by only keeping in memory the previous assistance request.

4.3. Subcorpora comparison

In a similar way to what we have done in section 3., we can compare the four different subcorpora identified in section 4. within the Daft corpus by using interactional profiles methodology introduced in 3.2. (without any need for preliminary mapping though). Finding objective difference criteria like different speech acts distribution could indeed be useful for a potential automatic identification of a request activity, which would allow the agent to deal differently with control orders, assistance requests and chat utterances.

The results of that comparison are displayed on figure 6. Not only this analysis confirms the non-homogeneity of the Daft corpus (which average interactional profile is reminded in white), but it reveals a very clear difference, in terms of speech acts, between direct (mainly directives and some expressives) and indirect assistance (mainly assertives and expressives) requests. This result is particularly interesting because classical methods based on vocabulary or linguistic parameters (as described in details in (Bouchet 2007)) fail to discriminate efficiently those two kind of assistance: interactional profiles are hence perfectly complementary.

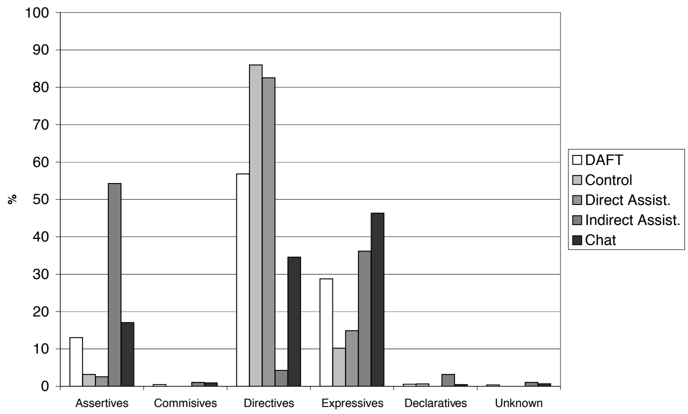


Figure 6: Interactional profiles comparison of Daft corpus' conversational activities

5. Conclusion and outlook

Using interactional profiles, we have shown that the Daft corpus was different from similar corpora in terms of speech acts distribution, certainly linked to the fact it is not human-human but human-computer interaction, thus confirming its necessity to study the Function of Assistance. Through a manual annotation of conversational activities within the Daft corpus, we have identified three assistance-related activities (control, direct and indirect assistance) representing 60% of the requests, the rest of them being chat-oriented. Finally, using again interactional profiles to compare the subcorpora defined by those activities, we managed to distinguish direct from indirect assistance requests.

Logical follow-up of this work shall focus on one side on getting a more accurate automatic identification of conversational activity (as a valuable first step analysis for the assisting agent), and on the other side on the formal modeling of those requests, particularly by taking into account the need for pragmatics in the case of indirect assistance.

Acknowledgements

The author particularly would like to thank anonymous reviewers for the methodological tips and references that have been helpful to improve the overall quality of this article.

References

R. Amalberti (1996). *La conduite des systèmes à risques*. PUF.

I. Androutsopoulos & M. Aretoulaki (2003). 'Natural Language Interaction'. In R. Mitkov (ed.), *The Oxford Handbook of Computational Linguistics*, pp. 629–649. Oxford University Press.

B. T. Atkins & H. M. A. Lewis (1996). *Language in Use*. Harper Collins Publishers.

F. Bouchet (2007). 'Caractérisation d'un Corpus de Requêtes d'Assistance'. In N. Hathout & P. Muller (eds.), *Actes de la 11e Rencontre des Etudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2007)*, vol. 2, pp. 357–366, Toulouse, France. IRIT Press.

A. Capobianco & N. Carbonell (2002). 'Contextual Online Help: a Contribution to the Implementation of Universal Access'. In S. Keates, P. J. Clarkson, & P. Robinson (eds.), *Universal Access and Assistive Technology*, pp. 131–140. Springer, London.

J. Carletta, et al. (1996). 'HCRC dialogue structure coding manual'. Tech. rep., HCRC, University of Edinburgh.

K. A. Ericsson & H. Simon (1993). *Protocol Analysis: Verbal Reports as Data*. MIT Press, Cambridge, MA, USA.

D. Jurafsky, et al. (1998). 'Switchboard Discourse Language Modeling Project Final Report'. Tech. rep., Center for Speech and Language Processing, Johns Hopkins University, Baltimore, MD, USA.

M. J. Kelly & A. Chapanis (1977). 'Limited Vocabulary Natural Language Dialogue'. *International Journal of Man-Machine Studies* **9**(4):479–501.

J. C. Lester, et al. (1997). 'The persona effect: affective impact of animated pedagogical agents'. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 359–366, Atlanta, Georgia, United States. ACM.

S. J. Molinsky & B. Bliss (1994). *Inventory of functions and conversation strategies: The Comprehensive course in functional English*, pp. 177–187. Prentice Hall.

M. Morel (1989). *Analyse linguistique de corpus*. Publications de la Sorbonne Nouvelle, Paris.

R. W. Morrell & D. C. Park (1993). 'The effects of age, illustrations, and task variables on the performance of procedural assembly tasks'. *Psychology and Aging* **8**(3):389–99. PMID: 8216959.

N. Ozkan (1994). *Vers un modèle dynamique du dialogue : analyse de dialogues finalisés dans une perspective communicationnelle*. Ph.D. thesis, INP Grenoble.

B. Post (2000). *Tonal and phrasal structures in French intonation*. Ph.D. thesis, Katholieke Universiteit Nijmegen.

G. Ripoche (2006). *Sur les traces de Bugzilla*. Ph.D. thesis, Université Paris-Sud 11.

J. Sansonnet, et al. (2005). 'Une architecture médiateur pour des agents conversationnels animés'. In *WACA'01 : Actes du Premier Workshop Francophone sur les Agents Conversationnels Animés*, pp. 31–39.

J. R. Searle (1969). *Speech Acts: An essay in the Philosophy of language*. Cambridge University Press, new edn.

D. P. Twitchell, et al. (2004). 'Using Speech Act Theory to Model Conversations for Automated Classification and Retrieval'. In *Proceedings of the 9th International Working Conference on the Language-Action Perspective on Communication Modeling*, pp. 121–129, New Brunswick, NJ. Rutgers University Press.

THE ALGEBRAIC STRUCTURE OF AMOUNTS: EVIDENCE FROM COMPARATIVES

Daniel Lassiter
Department of Linguistics, New York University

Abstract. Heim (2001) notes an empirical generalization stating that quantifiers may not intervene scopally between a comparative operator and its trace. I show that the constraint holds with some but not all quantifiiers, and that the limitations and their exceptions match quite closely the distribution of interveners in weak islands. I argue that the semantic theory of weak islands in Szabolcsi & Zwarts (1993) predicts the facts about comparatives as well. Finally, I discuss briefly the relationship between this proposal and that of Fox & Hackl (2006), suggesting that the theories are compatible but basically independent.

Several recent discussions of comparatives (Hackl 2000, Heim 2001, Bhatt & Pancheva 2004, Takahashi 2006) note an unexplained restriction on the scope of the comparative operator which has been dubbed the "Heim-Kennedy constraint":

(1) **Heim-Kennedy Constraint (HK)**:
A quantificational DP may not intervene between a degree operator and its trace.

Suppose, following Hackl (2000) and Heim (2001), that gradable adjectives are functions from individuals to sets of degrees, and that the comparative morphemes *-er* and *less* are scope-taking elements which compare the maxima of two sets of degrees. These otherwise plausible assumptions create a puzzle: certain quantifiers do not interact with degree-denoting expressions in the expected way.

(2) Every girl is less tall than Jim is.

a. Direct scope: $every\ girl > less > d\text{-}tall$
$\forall x[\mathbf{girl}(x) \rightarrow [max(\lambda d.\mathbf{tall}(x)(d))] < max(\lambda d.\mathbf{tall}(Jim)(d))]$
"For every girl x, Jim's max height is greater than x's max height."

b. Scope-splitting: $less > every\ girl > d\text{-}tall$
$max(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mathbf{tall}(x)(d)]) < max(\lambda d.\mathbf{tall}(Jim)(d)))$
"Jim's max height is greater than the max degree to which every girl is tall (i.e., he is taller than the shortest girl)."

If (2) had the"scope-splitting" reading in (2b), it would be true (on this reading) if the shortest girl is less tall than Jim. However, (2) is clearly false if any girl is taller than Jim. The Heim-Kennedy constraint (1) attempts to account for this restriction (and similar facts with different quantifiers) by stipulating that the quantificational DP *every girl* may not intervene between the degree operator *less* and its trace *d-tall*. The puzzle is what syntactic or semantic principles explain the constraint HK given that structures such as (1b) are semantically unexceptionable according to standard assumptions.

My purpose here is to show that HK follows from the theory of weak islands proposed by Szabolcsi & Zwarts (1993), henceforth S&Z. "Scope-splitting" readings of comparatives with certain quantificational DPs are semantically deviant for the same reason that how many/much questions are weak island-sensitive. One of S&Z's core claims is that amounts are weak island-sensitive because they have a more complex algebraic structure than normally assumed, and that the operations intersection and complement are not defined on this structure. As a result, the scope-splitting reading in (2b) is not available because computing it would require intersecting the heights of all the girls, and this operation is not available for purely semantic reasons.

1. Similarities between weak islands and comparative scope

As Hackl (2000) and Szabolcsi (2006) mention, there are considerable similarities between the limitations on degree operator scope summarized in HK and the core facts of weak islands disussed by Kroch (1989) and Rizzi (1990), among many others. Rullmann (1995) notes the following patterns:

(3) a. I wonder how tall Marcus is.
b. I wonder how tall this player is.
c. I wonder how tall every player is.
d. I wonder how tall most players are.
e. I wonder how tall many players are.

(4) a. *I wonder how tall Marcus isn't.
b. *I wonder how tall no player is.
c. *I wonder how tall few players are.
d. *I wonder how tall fewer than ten players are.
e. *I wonder how tall at most ten players are.

(5) a. Marcus is taller than Lou is.
b. Marcus is taller than this player is.
c. Marcus is taller than every player is.
d. Marcus is taller than most players are.
e. Marcus is taller than many players are.

(6) a. *Marcus is taller than Lou isn't.
b. *Marcus is taller than no player is.
c. *Marcus is taller than few players are.
d. *Marcus is taller than fewer than five players are.
e. *Marcus is taller than at most five players are.

These similarities are impressive enough to suggest that a theory of the weak island facts in (4) should also account for the limitations on comparatives in (6). Rullmann suggests that, in the case of how tall and taller, that the unavailability of the examples in (4) and (6) is due to semantic, rather than syntactic, facts. Specifically, both *wh*-questions and comparatives make use of a maximality operation, roughly as in (7):

(7) a. I wonder how tall Marcus is.
I wonder: what is the degree d such that $d = max(\lambda d.$Marcus is d-tall)?
b. Marcus is taller than Lou is.
$(\iota d.d = max(\lambda d.$Marcus is d-tall$)) > (\iota d.d = max(\lambda d.$Lou is d-tall$))$

With these interpretations of comparatives and questions, we predict that the sentences in (8) should be semantically ill-formed because each contains a definite description that is undefined:

(8) a. *I wonder how tall Marcus isn't.
I wonder: what is the degree d such that $d = max(\lambda d$. Marcus is not d-tall)?

b. *Marcus is taller than Lou isn't.
($\iota d.d = max(\lambda d.$Marcus is d-tall$)) > (\iota d.d = max(\lambda d.$ Lou is not d-tall))

If degrees of height are arranged on a scale from zero to infinity, there can be no maximal degree d such that Marcus or Lou is not d-tall, and so (8a) and (8b) are undefined.

Rullmann claims that similar reasoning will explain the unacceptability of the other downward entailing expressions in (4) and (6). However, the similarities between comparatives and weak island-sensitive expressions such as how tall go deeper than Rullmann's discussion would indicate. S&Z point out that several of the acceptable examples in (3) do not have all the readings predicted by the logically possible orderings of every player and how tall. As it turns out, the same scopal orders are also missing in the corresponding comparatives when we substitute *-er* for *how tall*. For example,

(9) I wonder how tall every player is.

a. *every player* > *how tall* > *d-tall*
"For every player x, I wonder: what is the max degree d s.t. x is d-tall)?"

b. *how tall* > *every player* > *d-tall*
"I wonder: what is the degree d such that d = Max(λd. every player is d-tall)?"

To satisfy the speaker's curiosity under the first reading in (9), we would have list all the players and their heights. In contrast, an appropriate response to the second reading (9b) would be to intersect the heights of all the players and give the maximum of this set, i.e. to give the height of the shortest player. This second reading is clearly not available. Similar facts hold for the corresponding comparative:

(10) Marcus is taller than every player is.

a. *every player* > *-er* > *d-tall*
"For every player, Marcus is taller than he is."

b. *-er* > *every player* > *d-tall*
"Marcus' max height is greater than the max height s.t. every player is that tall, i.e. he is taller than the shortest player."

The amount question in (9) and the amount comparative expression in (10) allow similar scopal orderings. Furthermore, Rullmann's explanation does not exclude the unacceptable readings. Unlike comparatives with an intervening negation, there *is* a maximal degree d s.t. every player is d-tall on Rullmann's assumptions, namely the height of the shortest player.

Note in addition that (10) is identical in terms of scope possibilities to our original comparative scope-splitting example in (2), although its syntax is considerably different. Like (2), (10) falls under HK, which correctly predicts the unavailability of (10b). However, HK does not address negation or *wh*-questions, and so leaves unexplained the systematic correspondences between comparatives and weak island-sensitive expressions. Rullmann addresses these correspondences, but cannot explain the missing readings in (9) and (10).

I will argue that S&Z's account of weak islands, which is designed to handle data such as those in (4), also explains the unavailability of the shortest-player' reading of the comparative in (10) as well as the corresponding gap in our original example (2). The essential insight is that the similarities between amount comparatives and amount *wh*-expressions are not due to monotonicity or to restrictions on movement of the the degree operator, but to the nature of amounts: specifically, their algebraic structure.

2. Comparative scope and the algebraic structure of amounts

In section 1 we saw that amount comparatives and amount *wh*-questions seem to have the same scope-taking abilities, despite their quite different syntax and overt word order. I will argue that the semantic theory of weak islands in S&Z extends to comparatives in a straightforward way that predicts that HK should hold. This theory leads to a clear notion of how amount comparatives and amount questions are "the same" in the relevant respects.

Like Rullmann (1995), S&Z argue that no syntactic generalization can account for the full range of weak islands, and propose to account for them in semantic terms. They formulate their basic claim as follows:

(11) Weak island violations come about when an extracted phrase should take scope over some intervener but is unable to.

S&Z explicate this claim in algebraic terms, arguing that weak islands can be understood if we pay attention to the operations that particular quantificational elements are associated with. For instance,

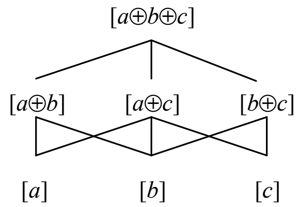
(12) Universal quantification corresponds to taking *intersections* (technically, meets).
Existential quantification corresponds to taking *unions* (technically, joins).
Negation corresponds to taking *complements.*

(12) becomes important once we assign particular algebraic structures as denotations to types of objects, since these operations are not defined for all structures. The prediction is that a sentence will be semantically unacceptable, even if it can be derived syntactically, if computing or verifying it requires performing an operation on a structure for which this operation is not defined. S&Z illustrate this claim with the verb *behave*, which induces weak islands:

(13) a. How did John behave?
b. *How didn't John behave?
c. How did everyone behave?
i. For each person, tell me: how did he behave?
ii. *What was the behavior exhibited by everyone?

Behave requires a complement that denotes a manner. S&Z argue that manners denote in a free join semilattice, the same structure which Landman (1991) suggests for masses.

(14) Free join semilattice



A noteworthy property of (14) is that it is closed under union, but not under complement or intersection. For instance, the union (technically, join) of $[a]$ with $[b \oplus c]$ is $[a \oplus b \oplus c]$, but the intersection (meet) of $[a]$ with $[b \oplus c]$ is not defined. The linguistic relevance of this observation is that it corresponds to our intuitions of appropriate answers to questions about behavior. In S&Z's example, suppose that three people displayed the following behaviors:

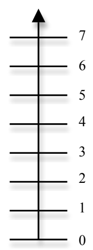
(15) John behaved *kindly and stupidly*.
Mary behaved *rudely and stupidly*.
Jim behaved *loudly and stupidly*.

If someone were to ask: "How did everyone behave?", interpreted with how taking wide scope as in (13c-ii), it would not be sufficient to answer "stupidly". The explanation for this, according to S&Z, is that computing the answer to this question on the relevant reading would require intersecting the manners in which John, Mary and Jim behaved, but intersection is not defined on (10). This, then, is a specific example of when"an extracted phrase should take scope over some intervener but is unable to". Similarly, (13b) is unacceptable because complement is not defined on (14). Extending this account to amounts is slightly trickier, since amounts seem to come in two forms. In the first, which S&Z label counting-conscious', *wh*-expressions are able to take scope over universal quantifiers. S&Z imagine a situation in which a swimming team is allowed to take a break when everyone has swum 50 laps. In this situation it would be possible to ask:

(16) [At least] How many laps has every swimmer covered by now?

In this case it seems (on the *how many*-wide interpretation) that the correct answer is the number of laps covered by the slowest swimmer. Counting-conscious amount expressions, then, had better denote in a lattice in which intersection is defined. The number line in (17) seems to be an appropriate choice.

(17) Lattice



Intersection and union are defined in this structure, though complement is not. This fact predicts that how many/much should be able to take scope over existential quantification but not negation:[1]

[1]Note that this example from S&Z constitutes a counter-example to HK stated as an LF-constraint as

(18) a. How many laps has at least one swimmer covered by now?
[Answer: the number of laps covered by the fastest swimmer.]

b. *How many laps hasn't John covered by now?

So far, then, (17) seems to be appropriate for amounts.

Many authors have assumed that the amounts that are compared in comparative constructions always denote in (17). Indeed, the problem we began this essay with — why can't *Every girl is less tall than John* mean "The shortest girl is shorter than John"? — was motivated by the assumption that it should be possible to intersect sets of degrees. The fact that intersection is defined on (17), and yet universal intervention is not available in (2), has led authors to various levels of stipulation (HK, in (Heim (2001)) or abandoning degrees in the analysis of comparatives (Schwarzchild & Wilkinson 2002).

I would like to suggest an alternative: heights and similar amounts do not denote in (17), but in a poorer structure for which intersection is not defined, as S&Z claim for island-sensitive amount *wh*-expressions. As S&Z note, such a structure is motivated already by the existence of non-counting-conscious amount *wh*-expressions which are sensitive to a wider variety of interveners than *how many* was in the examples in (16) and (18). This is clear for heights, for example:

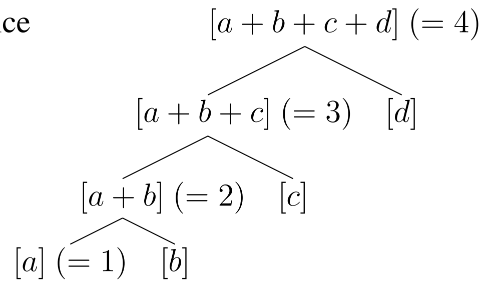
(19) How tall is every student in your class?

a. For every student in your class, how tall is he/she?'

b. * "What is the maximum height shared by all of your students, i.e. how tall is the shortest student?"

The unacceptability of (19b) is surprising given that the degree expression was able to take wide scope in the overtly similar (16). S&Z account for this difference by arguing that, unless counting is involved, amount expressions denote in a join semilattice:

(20) Join semilattice



(20) should be seen as a structure collecting arbitrary unit-sized bits of stuff, abstracting away from their real-world identity, like adding cups of milk to a recipe (S&Z pp.247-8). An important formal property of (20) is that "if p is a proper part of q, there is some part of q (the witness) that does not overlap with p" (p.247). As a result, intersection is not defined unless the objects intersected are identical. S&Z claim that this fact is sufficient to explain the unavailability of (19b), since the heights of the various students, being elements of (20), cannot be intersected.

in (1): the degree operator *how many* intervenes between the quantificational DP *every swimmer* and its trace. This constitutes further evidence for the algebraic approach advocated here, since the details of the expressions' denotations are relevant to the acceptable scopal relations, and not merely their structural configuration. It is unclear why this example is not as readily available with comparatives, however; see (Lassiter 2009) for more detailed discussion.

This explanation for (19b) relies on a quite general proposal about the structure of amounts. As a result, it predicts that amount-denoting expressions should show similar behavior wherever they appear in natural language, and not only in *wh*-expressions. The similarities between amount-denoting *wh*-expressions and comparatives, then, are explained in a most straightforward way: certain operations are not defined on amount-denoting expressions because of the algebraic structure of their denotations, regardless of the other details of the expressions they are embedded in. So, returning to (2),

(21) Every girl is less tall than Jim is.
Scope-splitting: *less* > *every girl* > *d-tall*
$max(\lambda d.\mathbf{tall}(\text{Jim})(d)) > max(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mathbf{tall}(x)(d)])$
"The max degree to which Jim is tall is greater than the max degree to which every girl is tall."

This interpretation is not available because the term $max(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mathbf{tall}(x)(d)])$ is undefined: on S&Z's theory, there can be no such degree. I conclude that the puzzle described by the Heim-Kennedy constraint was not a problem about the scope of a particular type of operator, but was generated by incorrect assumptions about the nature of amounts. Amounts are not simply points on a scale, but rather elements of (20). This proposal is independently motivated in S&Z, and it explains the restrictions captured in HK as well as other similarities between comparatives and weak islands.

At this point there are several important gaps in the account. The first is that S&Z do not work out their account compositionally, and this needs to be done. The second problem is that it remains unexplained (as it did in the original formulation of HK) why certain modals and intensional verbs are able to intervene between a degree operator and its trace both in amount comparatives and amount questions, as discussed at length in Heim (2001): (22) and (23) illustrate.

(22) (This draft is 10 pages.) The paper is required to be exactly 5 pages longer than that.

(Heim 2001, p.224)

a. *required* > *exactly 5 pages -er* > *that-long*
$\forall w \in Acc : max(\lambda d : \mathbf{long}_w(p, d)) = 15pp$
"In every world, the paper is exactly 15 pages long"

b. *exactly 5 pages -er* > *required* > *that-long*
$max(\lambda d[\forall w \in Acc : \mathbf{long}_w(p, d)) = 15pp$
"The max common length of the paper in all accessible worlds, i.e. its length in the world in which it is shortest, is 15 pages"

(23) How long is the paper required to be?

a. *required* > *how long* > *that-long*
"What is the length s.t. in every world, the paper is exactly that long?"

b. *how long* > *required* > *that-long*
"What is the max common length of the paper in all accessible worlds, i.e. its length in the world in which it is shortest?"

These data support the present theory in that comparatives and weak island-sensitive *wh*-expressions pattern similarly in yet another way. However, on the assumption that require involves universal quantification over accessible worlds, the (b) readings of these examples are problematic. S&Z suggest that these operators are acceptable interveners because they do not involve algebraic operations. This is perhaps too drastic a step, but detailed investigation is needed to account for the complicated and subtle data involved, including the fact that some modals and intensional verbs involving universal quantification (must, require) can intervene while others (should, be supposed to) cannot (cf. Heim (2001)).

3. Some notes on density and informativity

In this section I discuss very briefly the relationship between the present analysis and an influential proposal by Fox & Hackl (2006). I show that the algebraic account is not in direct competition with Fox and Hackl's theory, but that there are some complications in integrating the two approaches.

Fox and Hackl argue that amount-denoting expressions always denote on a dense scale, effectively (17) with the added stipulation that, for any two degrees, there is always a degree that falls between them. The most interesting data from the current perspective are in (24) and (25):

(24) a. How fast are we not allowed to drive?
b. *How fast are we allowed not to drive?

(25) a. How fast are we required not to drive?
b. *How fast are we not required to drive?

The contrasts in (24) and (25) are surprising from S&Z's perspective: on their assumptions, there is no maximal degree d such that you are not allowed to drive d-fast, and yet (23a) is fully acceptable. In addition, (24a) and (25a) do not ask for maxima but for minima (the least degree which is unacceptably fast, i.e. the speed limit). Fox and Hackl show that the minimality readings of (24a) and (25a), and the ungrammaticality of (24b) and (25b) follow if we assume (following Dayal (1996) and Beck & Rullmann (1999)) that *wh*-questions do not ask for a maximal answer but for a maximally informative answer, defined as follows:

(26) The maximally informative answer to a question is the true answer which entails all other true answers to the question.

Fox and Hackl show that, on this definition, upward monotonic degree questions ask for a maximum, since if John's maximum height is 6 feet, this entails that he is 5 feet tall, and so on for all other true answers. However, downward entailing degree questions ask for a minimum, since if we are not allowed to drive 70 mph, we are not allowed to drive 71 mph, etc. [2]

[2] A problem which Fox and Hackl do not note is that (i) below should be unacceptable, since their account of modal intervention assumes that a 70 mph speed limit denotes a closed interval from 70 to infinity.
(i) *How fast are we allowed to drive?* —70 mph.
(ii) *How fast are we not allowed to drive?* —70 mph.
But (i) is at least as good as (ii), and probably even more natural. This poses a problem for Fox and Hackl's explanation of modal intervention: on their account, if (ii) is accepable then (i) should be ruled out, and vice versa.

This is not as deep a problem for the present theory as it may appear. S&Z assume that *wh*-questions look for a maximal answer, but it is unproblematic simply to modify their theory so that *wh*-questions look for a maximally informative answer. Likewise, we can just as easily stipulate that a join semilattice (20) is dense as we can stipulate that a number line (17) is dense; this maneuver would replicate Fox and Hackl's result about minima in downward entailing contexts. In this way it is possible simply to combine S&Z's theory with Fox and Hackl's. In fact, this is probably independently necessary for Fox and Hackl, since their assumption that amounts always denote in (17) fails to predict the core data of the present paper: the fact that *How tall is every girl?* and *Every girl is less tall than John* lack a "shortest-girl" reading. The only real barrier to a simple marriage of these theories is the fact, already noted in the previous section, that S&Z do not have an explanation for the occasional acceptability of universal modal interveners with non-counting amount questions.

I conclude that neither theory is complete: Fox and Hackl's theory lacks an account of the unacceptability of non-modal universal interveners (i.e., HK), and of the difference between counting and non-counting amount questions; but S&Z lack an account of universal modal intervention. Nevertheless, the two theories are broadly compatible. Note that this is not an endorsement of Fox and Hackl's central thesis — it merely shows that if their theory is correct, this fact does not constitute a reason to abandon the current approach to the HK phenomena.

Finally, note that the maximal informativity hypothesis in (26), whatever its merit in *wh*-questions and other environments discussed by Fox and Hackl, is not appropriate for comparatives: here it appears that we need simple maximality.[3]

(27) a. How fast are you not allowed to drive?
b. *You're driving faster than you're not allowed to.

A simple extension of the maximal informativity hypothesis to comparatives would predict that (27b) should mean "You are exceeding the speed limit". In contrast, the maximality-based account predicts that (27b) is unacceptable, since there is no maximal speed which is not allowed. This appears to be the correct prediction.

4. Conclusion

To sum up, the traditional approach on which amounts are arranged on a scale of degrees fails to explain why the constraint HK in (1) should hold. However, S&Z's semantic account of weak islands predicts the existence of this constraint and the numerous similarities between amount comparatives and amount-denoting *wh*-expressions. To be sure, important puzzles remain.[4] Nevertheless, the algebraic approach to comparative scope offers a promising explanation for a range of phenomena that have not been previously treated in a unified fashion.

Furthermore, if S&Z's theory turns out to be wrong, all is not lost. The most important lesson of the present paper, I believe, is not that S&Z's specific theory of weak islands is

[3]Thanks to a reviewer for bringing the contrast in (27) to my attention.

[4]An additional important question which I have not discussed for space reasons is the relationship between the algebraic theory and interval-based theories such as (Abrusán 2007) and (Schwarzchild & Wilkinson 2002). The relation between these approaches is discussed in (Lassiter 2009), along with arguments in favor of the current theory.

correct — as we have seen, there are certainly empirical and technical challenges — but rather that weak island phenomena are not specific to *wh*-questions. In fact, we should probably think of the phenomena summarized by the Heim-Kennedy constraint as *comparative weak islands*. However the theory of weak islands progresses, evidence from comparatives will need to play a crucial role in its development.

Acknowledgements

I wish to thank Chris Barker and Anna Szabolcsi for their generous help, and Arnim von Stechow and Emmanuel Chemla for useful discussion.

References

M. Abrusán (2007). *Contradiction and Grammar: The case of Weak Islands*. Ph.D. diss., MIT.

S. Beck & H. Rullmann (1999). 'A flexible approach to exhaustivity in questions'. *Natural Language Semantics* **7**(3):249–298.

R. Bhatt & R. Pancheva (2004). 'Late merger of degree clauses'. *Linguistic Inquiry* **35**(1):1–45.

V. Dayal (1996). *Locality in wh-quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer.

D. Fox & M. Hackl (2006). 'The universal density of measurement'. *Linguistics and Philosophy* **29**(5):537–586.

M. Hackl (2000). *Comparative quantifiers*. Ph.D. diss., MIT.

I. Heim (2001). 'Degree operators and scope'. In Fery & Sternefeld (eds.), *Audiatur Vos Sapientiae: A Festschrift for Arnim von Stechow*. Berlin: Akademie Verlag.

A. Kroch (1989). 'Amount quantification, referentiality, and long wh-movement'. Manuscript, University of Pennsylvania .

F. Landman (1991). *Structures for Semantics*. Dordrecht: Kluwer.

D. Lassiter (2009). 'Explaining a restriction on the scope of the comparative operator'. To appear in *Proceedings of the 33rd Penn Linguistics Colloquium* .

L. Rizzi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

H. Rullmann (1995). *Maximality in the Semantics of wh-constructions*.

R. Schwarzchild & K. Wilkinson (2002). 'Quantifiers in comparatives: A semantics of degree based on intervals'. *Natural Language Semantics* **10**(1):1–41.

A. Szabolcsi (2006). 'Strong vs. weak islands'. *The Blackwell Companion to Syntax* **4**:479–531.

A. Szabolcsi & F. Zwarts (1993). ‘Weak islands and an algebraic semantics for scope taking’. *Natural language semantics* **1**(3):235–284.

S. Takahashi (2006). ‘More than two quantifiers’. *Natural Language Semantics* **14**(1):57–101.

Frame-Annotated Corpus for Extraction of the *Argument-Predicate* Relations

Ekaterina Ovchinnikova
Institute of Cognitive Science, University of Osnabrück

Abstract. Argument-predicate semantic relations are shown to have many applications in natural language processing. However, at present there exist no corresponding lexical semantic knowledge bases. This paper presents an approach to automatic acquisition of the argument-predicate relations. The relations have been extracted from SALSA, a German corpus manually annotated with FrameNet frames. The relatively small size of SALSA does not allow to estimate the semantic relatedness in the extracted argument-predicate pairs. Therefore we use a larger corpus for ranking. Since the evaluation process is still in progress, the paper describes only the evaluation strategy based on human judgments obtained experimentally.

1. Introduction

The fact that predicates ommited in a discourse because of redundancy can be predicted on the basis of the semantics of nouns was first theoretically grasped by Pustejovsky's Generative Lexicon theory, (Pustejovsky 1991). For example, the sentence *John finished the cigarette* can be most plausibly interpreted as *John finished smoking the cigarette* because the meaning of the noun *cigarette* is strongly associated with the smoking activity. It has been claimed that information about predicates associated with nouns can be helpful for a variety of tasks in natural language processing (NLP), see for example (Pustejovsky, et al. 1993, Voorhees 1994). However, currently there exist no corresponding lexical knowledge bases. The NLP practice has shown that freely available lexical resources containing semantic relations, such as WordNet (Fellbaum 1998) or FrameNet (Ruppenhofer, et al. 2006), are of great importance and have been constantly reused by different NLP systems. Therefore several approaches have been presented that aim at creating a knowledge base containing information about predicates associated with nouns. At present, there exist two main research paradigms for developing such knowledge bases. The first paradigm concerns manual development of the resource (Pustejovsky, et al. 2006), while the second one relies on automatic acquisition methods, see for example (Cimiano & Wenderoth 2007). Section 3 contains a more detailed description of the existing approaches.

In this paper we propose a procedure for automatic extraction of the argument-predicate relations from a semantically annotated corpus. We exploit SALSA (Burchardt, et al. 2006), a German newspaper corpus manually annotated with FrameNet frames based on frame semantics. Thus, the extracted argument-predicate relations are represented in terms of frames. The main contribution of the paper lies in demonstrating a possibility of learning semantic relations from annotated corpora. Using a manually annotated corpus for relation extraction has one particular advantage compared to extraction from plain text: the type of an argument-predicate relation is already annotated; there is no need to determine it by automatic means which are usually error-prone. However, the relatively small size of SALSA does not allow to make relevant predictions about the degree of semantic relatedness in the extracted argument-predicate pairs, see section 4. Therefore we use a larger corpus for estimating argument-predicate relatedness. Since the evaluation process has not been finished yet, we present only the evaluation strategy. The results will be evaluated quantitatively against human judgments obtained experimentally, see section 6. The proposed evaluation procedure is similar to that presented in (Cimiano & Wenderoth 2007). First, we create a gold standard for 30 words from the argument list and evaluate our approach with respect to this gold standard. Second, we ask the test subjects to rate extracted argument-predicate relations using a four-point scale and calculate correlation between automatic ranking and human rating.

2. Implicit Predicates

Example 1

(a) *Als ich mit diesem Buch angefangen habe...*
’When I have started this book...’

(b) *eine komplizierte Frage*
’a complicated question’

(c) *Studentenfutter*
’student food’

(d) *Nachrichtenagentur X über Beziehungen beider Seiten der Taiwan-Strasse*
’News agency X about relations of both sides of the Taiwan Strait’

(e) *Hans ist beredt*
’Hans is eloquent’

One of the most studied phenomena that (Pustejovsky 1991) has called logical metonymy is illustrated by the examples (1a) and (1b). In the case of logical metonymy an implicit predicate is inferable from particular verb-noun and adjective-noun pairs in a systematic way. The verb *anfangen* ’to start’ and the adjective *kompliziert* ’complicated’ in the mentioned examples semantically select for an event, while the nouns (*Buch* ’book’ and *Frage* ’question’ respectively) have a different semantic type. However, the set of the most probable implicit predicates is predictable from the semantics of the nouns. Thus, (1a) plausibly means *Als ich angefangen habe dieses Buch zu lesen/schreiben...* ’When I have started to read/write this book...’ and (2a) plausibly means *eine Frage die kompliziert zu beantworten ist* ’a question which is complicated to answer’.

Besides logical metonymy there are other linguistic phenomena requiring knowledge about predicates associated with an argument for their resolution. Example (1c) contains a noun compound which can be interpreted on basis of the meaning of the noun *Futter* ’food’. In general, noun compounds can be interpreted in many different ways depending on the semantics of the constituencies: *morning coffee* is a coffee which is drunk in the morning, *brick house* is a house which is made of bricks etc. In case of (1c) the relation via the predicate *essen* ’to eat’ taking *Studenten* ’students’ as a subject and *Futter* ’food’ as an object seems to be the most plausible one.

The phrase (1d) is a title of a newspaper article. As in the previous examples, a predicate is left out in (1d). The meaning of the preposition *über* ’about’ can help to narrow the set of possible predicates, but still allows an inadequately large range of interpretations. However, the semantics of the noun *Nachrichtenagentur* ’news agency’ clearly supports such interpretations as *berichten* ’to report’, *melden* ’to message’ or *informieren* ’to inform’.

Most of the literature discusses predicates inferable from nouns. However, other parts of speech can support similar inferences. In example (1e) a predicate is predictable on the basis of the meaning of the adjective *beredt* 'eloquent'. The sentence (1e) most plausibly means that Hans speaks eloquently.

The cases when a predictable predicate is left out are not rare in natural language. For example, for logical metonymy a corpus study has shown that the constructions like *begin V NP* are rare if the verb *V* corresponds to a highly plausible interpretation of *begin NP* (Briscoe, et al. 1990).

Inferring implicit predicates can be useful for a variety of NLP tasks such as language generation, information extraction, question answering or machine translation, see (Lapata & Lascarides 2003). Many NLP applications employing semantic relations are connected to paraphrasing or query expansion, see for example (Voorhees 1994). Suppose that an NLP system receives the query *schnelle Bombe* 'quick bomb'. Probably, in this case the user is interested in finding information about bombs that explode quickly rather then about bombs in general. Knowledge about predicates associated with the noun *Bombe* 'bomb' could be used for predicting a set of probable implicit predicates. However, for generation of the semantically and syntactically correct paraphrases it is sometimes not enough to guess the most probable argument-predicate pairs. Information about types of an argument-predicate relation could be helpful, i.e. which semantic and syntactic position does the argument fill in the argument structure of the predicate. For example, compare *eine Bombe explodiert schnell* 'a bomb explodes quickly' for *schnelle Bombe* and *ein Buch schnell lesen/schreiben* 'to read/write a book quickly' for *schnelles Buch* 'quick book'. In the first case the argument *Bombe* fills the subject position, while in the second case *Buch* fills the object position.

3. Related Work

At present Pustejovsky's theory of the Generative Lexicon, GL (Pustejovsky 1991), provides the most influential account of implicit predicates. According to Pustejovsky the meaning of a noun includes a *qualia structure* consisting of four roles: constitutive, agentive, formal and telic. The telic role describes purpose and function of the object denoted by the noun and the agentive role describes factors involved in the origin of the object. Thus, the lexical meaning of the noun *book* includes *read* as a telic role and *write* as an agentive role. In the framework of GL Pustejovsky et al. (2006) are manually developing the Brandeis Semantic Ontology which is a large generative lexicon ontology and dictionary.

There are several approaches to automatic acquisition of qualia structures from text corpora which aim at supporting the time-consuming manual work. For example, Pustejovsky et al. (1993) use generalized syntactic patterns for extracting qualia structures from a partially parsed corpus. Cimiano and Wenderoth (2007) suggest a pattern based method for automatic extraction of qualia structures from the Web. There also exist approaches to learning qualia structures from corpora using machine learning techniques. (Claveau & Sébillot 2004) for example suggests a symbolic machine learning method which allows to infer morpho-syntactic and semantic patterns of semantic relations between verbs and nouns. Using this method the system introduced in (Claveau & Sébillot 2004) can learn whether a given verb is a qualia element for a given noun. However, it can not distinguish between different qualia roles, i.e. it does not account for types of noun-verb relations. The results of the human judgment experiment reported in (Cimiano & Wenderoth 2007) suggest that the automatic acquisition of qualia structures is a difficult task. Human test subjects have shown a very low agreement (11,8% average agreement) in providing qualia structures for given nouns.

Another line of research on inferring implicit predicates concerns using information about collocations derived from corpora. For example, Lapata and Lascarides (2003) resolve logical metonymy on the basis of the distribution of paraphrases like *finish the cigarette – finish smoking the cigarette* and *easy problem – problem which is easy to solve* in a corpus. This approach shows promising results, but it is limited to logical metonymy. Similarly, Nastase et al. (2006) use grammatical collocations for defining semantic relations between constituents in noun compounds.

In contrast to previous approaches, in our study we aim at extracting argument-predicate relations from a semantically annotated corpus. Using an annotated corpus we avoid problems of defining types of these relations by automatic means which are usually error-prone. Moreover, we represent argument-predicate relations in terms of FrameNet frames which allow for a fine-grained and grounded representation supporting paraphrasing, see next sections. The proposed approach is not restricted to nouns. We also concern relations where argument positions are filled by adjectives, adverbs or even verbs.

4. Resources

For relation extraction we use the SALSA corpus (Burchardt et al. 2006) developed at Saarland University. SALSA is a German newspaper corpus manually annotated with role-semantic information. The 2006 SALSA release contains about 20 000 predicate instances annotated with the set of FrameNet frames (Ruppenhofer et al. 2006). The FrameNet (FN) lexical resource is based on frame semantics (Fillmore 1976). The lexical meaning of predicates in FN is expressed in terms of frames (approx. 800 frames) which are supposed to describe prototypical situations spoken about in natural language. Every frame contains a set of roles (or frame elements, FEs) corresponding to the participants of the described situation. Predicates with similar semantics evoke the same frame, e.g. *to give* and *to hand over* evoke the GIVING frame. Consider a FN annotation for the sentence (a) below. In this annotation DONOR, RECIPIENT and THEME are roles in the frame GIVING and *John*, *Mary* and *a book* are fillers of these roles. FN annotation generalizes across near meaning-preserving transformations, see (b).

Example 2

(a) [*John*]$_{\text{DONOR}}$ [*gave*]$_{\text{GIVING}}$ [*Mary*]$_{\text{RECIPIENT}}$ [*a book*]$_{\text{THEME}}$.

(b) [*John*]$_{\text{DONOR}}$ [*gave*]$_{\text{GIVING}}$ [*a book*]$_{\text{THEME}}$ [*to Mary*]$_{\text{RECIPIENT}}$.

In FN information about syntactic realization patterns of frame elements as well as information about frequency of occurrences of these patterns in corpora is provided. For example, the role DONOR in the frame GIVING is most frequently filled by a noun phrase in the subject position or by a prepositional phrase with the preposition *by* as the head in the complement position.

FrameNet project originally aimed at developing a frame-semantic lexical database for English. Later on FN frames turned out to be to large extent language independent, see (Burchardt et al. 2006). In most cases German predicates could be successfully described by the FN frames. However, some of the frames required adaptation to the German data, e.g. new FEs were introduced. Since FN does not cover all possible word senses, new frames needed to be introduced for some of the predicates.

We have chosen the SALSA corpus for our experiments because to our knowledge it is the only freely available corpus which contains both syntactic and role-semantic annotation. However, we are aware of the fact that SALSA (approx. 700 000 tokens) is too small to compute a reliable co-occurrence model, though it is relatively large for a manually annotated corpus. As it was shown in (Bullinaria & Levy 2007), co-occurrence-based approaches need very large training corpora in order to reliably compute semantic relatedness. The SALSA corpus comprising less than 1 million tokens is too small for this purpose. Moreover, a considerable number of predicates in SALSA appeared to be unannotated. We have tried to overcome the size problems by using a larger unannotated corpus for recomputing the weights of the extracted argument-predicate relations.

5. Our Approach

As already stated above, we aim at extracting argument-predicate relations using a role-semantic annotation. Our goal is to extract from SALSA tuples of the form ⟨*Argument,* ROLE, FRAME, *Predicate*⟩ such that the *Argument* plausibly fills the ROLE in the FRAME evoked by the *Predicate*. Since FrameNet contains information about syntactic realization patterns for frame elements, representation of argument-predicate relations in terms of frames directly supports generation of semantically and syntactically correct paraphrases, cf. *schnelle Bombe* and *schnelles Buch* example in section 2.

The proposed relation extraction procedure works as follows. First, we extract from the corpus all annotated frames. Then, arguments of every frame are extracted and a relation between arguments and the frame evoking predicate is defined in terms of the roles which these arguments fill in the frame. Finally, the semantic relatedness in the extracted argument-predicate pairs is estimated. In example 3 two different tuples representing argument-predicate relations have been extracted from the given sentence annotated with the frame ARREST.

Example 3
[*Fünf Oppositionelle*]$_{\text{SUSPECT}}$ *sind in Ebebiyin* [*von der Polizei*]$_{\text{AUTHORITIES}}$ [*festgenommen*]$_{\text{ARREST}}$ *worden.*
'Five members of the opposition have been arrested by the police in Ebebiyin.'

Extracted tuples:

Argument	**Role**	**Frame**	**Predicate**
Oppositionell	SUSPECT	ARREST	*festnehmen*
Polizei	AUTHORITIES	ARREST	*festnehmen*

Since SALSA is also annotated syntactically, every frame role is filled by some syntactic constituent. Therefore, in order to find proper semantic arguments one needs to extract a content head from every constituent filling a frame role. If a role filler is represented by an anaphoric expression it should be resolved. The task of finding role fillers proved to be relatively easy.[1] On the contrary, anaphora resolution is well-known to be

[1]For verb phrases with auxiliary or modal verbs as heads the main verb was taken as a corresponding role filler.

one of most challenging NLP tasks, see (Mitkov 2002). Since we do not focus on it, we treat only pronominal anaphora using a straightforward resolution algorithm: given a pronoun, the first noun which agrees in number and gender with the pronoun is supposed to be its antecedent. In order to evaluate this resolution procedure we have inspected 100 anaphoric cases. In approximately three fourths of the cases anaphora was resolved correctly. Therefore, we have assigned the confidence rate of 0,75 to the tuples resulting from a resolved anaphora. In non-anaphoric cases the confidence rate of 1 was assigned.

For every tuple we have summed up the corresponding confidence rates. Only lexemes which belong to an open word class, i.e. nouns, verbs, adjectives and adverbs, have been considered. Finally, we have obtained around 30 000 tuples with confidence rates ranging from 0,75 to 88. It is not surprising that most of the arguments appear to be nouns, while most of the predicates are expressed by verbs. Since SALSA has been annotated manually, there are almost no mistakes in defining types of the semantic relations between arguments and predicates.[2] However, as mentioned in section 4 the size of SALSA does not allow to make relevant predictions about the distribution of frames and role fillers. In order to overcome this problem we have developed a measure of semantic relatedness between the extracted arguments and predicates which takes into account their co-occurrence in a larger and more representative corpus.

For computing semantic relatedness we have used a lemmatized newspaper corpus (Süddeutsche Zeitung, SZ) of 145 million words. Given a tuple t with a confidence rate c containing an argument a and a predicate p, the relatedness measure *rm* of t was computed as follows:

$$rm(t) = lsa(a, p) + c/max(c),$$

where the $lsa(a, p)$ is based on Latent Semantic Analysis (LSA), (Deerwester, et al. 1990). LSA is a vector-based technique that has been shown to give reliable estimates on semantic relatedness. It makes use of distributional similarities of words in text and constructs a semantic space (or word space) in which every word of a given vocabulary is represented as a vector. Such vectors can then be compared to one another by the usual vector similarity measures (e.g. cosine). We calculated the LSA word space using the Infomap toolkit10 v. 0.8.6 (*http://infomap-nlp.sourceforge.net*). The co-occurrence matrix (window size: 5 words) comprised 80 000×3 000 terms and was reduced by SVD to 300 dimensions. For the vector comparisons the cosine measure was applied. To those cue arguments which did not occur in the analyzed SZ corpus (approx. 3500 words) a *lsa* measure of 0 was assigned. To provide a comparable contribution to rm, the confidence rates c extracted from SALSA are divided by the maximal confidence rate.

Table 1 contains the 5 most semantically related predicates for three example argument.[3] The table shows that more than one relation can exist between arguments and predicates. For example, the pair (*Haft*, *sitzen*) 'imprisonment', 'to sit' was annotated in SALSA both with the BEING_LOCATED and with the POSTURE frames. In this case ambiguity is due to the annotation disagreements. In some other cases, ambiguity of the semantic relation between an argument and a predicate adequately reflects variety of roles which the argument can fill. For example, two different tuples have been extracted for the

[2]Mistakes can arise only because of the annotation errors and errors in the anaphora resolution procedure.

[3]The complete list of the extracted tuples can be found online at *http://www.ikw.uni-osnabrueck.de/∼eovchinn/APrels/*.

Table 1: Examples of the extracted argument-predicate tuples

Argument	Role	Frame	Predicate	rm
Haft 'imprisonment'	FINDING	VERDICT	*verurteilen* 'to sentence'	0,939
	LOCATION	BEING_LOCATED	*sitzen* 'to sit'	0,237
	LOCATION	POSTURE	*sitzen* 'to sit'	0,226
	MESSAGE	REQUEST	*fordern* 'to ask'	0,153
	BAD_OUTCOME	RUN_RISK-FNSALSA	*drohen* 'to threaten'	0,144
Polizei 'police'	AUTHORITIES	ARREST	*festnehmen* 'to arrest'	0,782
	AUTHORITIES	ARREST	*verhaften* 'to arrest'	0,386
	INVESTIGATOR	CRIMINAL_INVESTIGATION	*ermitteln* 'to detect'	0,291
	PHENOMENON	PERCEPTION_ACTIVE	*beobachten* 'to observe'	0,24
	COGNIZER_AGENT	SEEKING	*suchen* 'to seek'	0,227
Demonstrant 'demonstrator'	AGENT	PROTESTIEREN1-SALSA	*protestieren* 'to protest'	0,739
	ADDRESSEE	REQUEST	*aufrufen* 'to call up'	0,359
	SPEAKER	REQUEST	*fordern* 'to ask'	0,215
	VICTIM	KILLING	*töten* 'to kill'	0,079
	INTERLOCUTOR_2	DISCUSSION	*sprechen* 'to speak'	0,045

pair (*Buch, schreiben*) 'book, 'to write': ⟨*Buch,* TEXT, TEXT_CREATION, *schreiben*⟩ and ⟨*Buch,* MEDIUM, STATEMENT, *schreiben*⟩. The first tuple corresponds to phrases like *ein Buch schreiben* 'to write a book', while the second one abstracts from the expressions like *in einem Buch schreiben* 'to write in a book'.

As one could expect, being a newspaper corpus SALSA appeared to be thematically unbalanced. The most top-ranked argument-predicate relations occurring in SALSA reflect common topics discussed in newspapers: economics (e.g. (*Prozent*, *steigen*), 'percent', 'to increase'), criminality (e.g. (*Haft*, *verurteilen*) 'imprisonment', 'to sentence'), catastrophes (e.g. (*Mensch*, *töten*) 'human', 'to kill') etc.

6. Evaluation Strategy

Since the evaluation process is still in progress, in this section we describe only the evaluation strategy that we apply. The extracted argument-predicate relations are intended to be used for inferring intuitively obvious predicates, therefore we aim at checking to which extent they correspond to human intuition. Similar to (Cimiano & Wenderoth 2007) we provide a gold standard for 30 arguments occurring in the SALSA corpus. The test arguments are selected randomly from the set of those arguments that have more than one predicate associated with them such that a value of argument-predicate relatedness exceeds the average one. These words are nearly uniformly distributed among 30 participants of the experiment, who are all non-linguists, making sure that each word is treated by three different subjects. We ask our subjects to write phrases that contain a predicate taking the given word as an argument, e.g. *book – to read a book*. Beside the task description and an informal introduction of the notion of predicate the participants are shown the following examples:

(a) *Aktie* 'stock' : *Kauf der Aktien* 'buying of stocks', *Aktien kaufen* 'to buy stocks', *Aktien an der Börse* 'stocks on the bourse' (is inappropriate because the word "bourse" describes a place and not an event)

(b) *beredt* 'eloquent': *beredt sprechen* 'to speak eloquently', *ein beredter Sprecher* 'an eloquent speaker' (is inappropriate because the word "speaker" describes a person and not an event)

For each of the given arguments the subjects are asked to provide between 5 and 10 phrases. After the test will be completed, we will manually annotate obtained phrases with frames. Tuples of the form ⟨*Argument,* ROLE, FRAME, *Predicate*⟩ extracted from annotated phrases will constitute the gold standard. For every test word only those tuples which follow from phrases provided by all three subjects treating this word will be included into the gold standard. We will evaluate our results with respect to the gold standard using the precision/recall characteristics, cf. (Cimiano & Wenderoth 2007). The precision characterizes the procedure exactness, i.e. how many redundant tuples not contained in the gold standard are retrieved by our procedure. The recall measures the completeness, i.e. how many tuples of the gold standard are extracted automatically. In order to check whether the developed relatedness measure gives advantaged in selecting the most plausible argument-predicate associations, we perform another experiment. On this second step of evaluation we generate phrases from tuples extracted by our procedure and ask the participants to rate these phrases with respect to their naturalness using a four-point scale. Then we will calculate the correlation between human judgments and the ranking obtained automatically.

7. Conclusion and Discussion

In this paper we have presented an approach to automatic extraction of argument-predicate relations from a semantically annotated corpus. We have combined the advantages offered by annotated and unannotated corpora. Besides extracting argument-predicate pairs the proposed method allows us to define types of semantic relations in terms of FrameNet frames. Such representation of the relations is promising with respect to the paraphrasing task, because it supports generation of syntactically and semantically correct phrases. The proposed procedure is not restricted to arguments expressed by nouns and treats also other content parts of speech.

Since the evaluation process has not been finished yet, we have presented only the evaluation strategy which we are using. It consists in comparing automatically obtained results with human judgments obtained experimentally. The participants of the experiment are asked to provide short phrases containing given cue words and predicates associated with these words as well as to rate phrases generated from the automatically extracted tuples. The evaluation results will be available in the nearest future. The complete list of the extracted relations as well as the results of the experiment will be available online at *http://www.ikw.uni-osnabrueck.de/ eovchinn/APrels/.*

This study presents only first steps towards using semantically annotated corpora for automatic relation extraction. There are several ways to improve the proposed procedure. First, an implementation of a more advanced anaphora resolution algorithm treating pronominal as well as nominal anaphora should significantly raise the performance of the procedure. Concerning relatedness measure, additional corpus-based measures such as pointwise mutual information (Church & Hanks 1991), measures based on syntactic relations (Claveau & Sébillot 2004) or Web-based measures (Cimiano & Wenderoth 2007) could appear to be useful for improving the ranking of the extracted relations.

An obvious limitation of the presented approach is that it is bounded to manual annotations which are hard to obtain. However, since semantic annotations are useful for many different goals in linguistics and NLP, the number of reliable annotated corpora constantly grows.[4] Moreover, recently several tools have been developed which perform role annotation automatically, for example see (Erk & Pado 2006). Therefore we believe that approaches using semantic annotation are valid and promising. In the future we plan to experiment with large role-annotated corpora for English such as PropBank (approx. 300 000 words, (Palmer, et al. 2005)) and the FrameNet-annotated corpus provided by the FN project (more than 135 000 annotated sentences, (Ruppenhofer et al. 2006)). Since these corpora do not contain syntactic annotation, for extracting argument-predicate relations we will need to parse annotated sentences.

References

T. Briscoe, et al. (1990). 'Enjoy the paper: Lexical semantics via lexicology'. In *Proceedings of the 13th International Conference on Computational Linguistics*, pp. 42–47.

J. Bullinaria & J. Levy (2007). 'Extracting semantic representations from word co-occurrence statistics: A computational study'. *Behavior Research Methods* **39**(3):510–526.

A. Burchardt, et al. (2006). 'The SALSA corpus: A German corpus resource for lexical semantics'. In *Proceedings of LREC 2006*, pp. 969–974, Genoa, Italy.

K. Church & P. Hanks (1991). 'Word Association Norms, Mutual Information and Lexicography'. *Computational Linguistics* **16**(1):22–29.

P. Cimiano & J. Wenderoth (2007). 'Automatic Acquisition of Ranked Qualia Structures from the Web'. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 888–895.

V. Claveau & P. Sébillot (2004). 'From efficiency to portability: acquisition of semantic relations by semi-supervised machine learning'. In *Proceedings of COLING 2004*, pp. 261–267.

S. Deerwester, et al. (1990). 'Indexing by Latent Semantic Analysis'. *American Society of Information Science* **41**(6):391–407.

K. Erk & S. Pado (2006). 'Shalmaneser - a flexible toolbox for semantic role assignment'. In *Proceedings of LREC 2006*, Genoa, Italy.

C. Fellbaum (ed.) (1998). *WordNet: An Electronic Lexical Database.* MIT Press, first edn.

C. J. Fillmore (1976). 'Frame semantics and the nature of language'. In *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, vol. 280, pp. 20–32.

M. Lapata & A. Lascarides (2003). 'A Probabilistic Account of Logical Metonymy'. *Computational Linguistics* **29(2)**:261–316.

[4]At present FrameNet annotated corpora are available for English, German and Spanish, see *framenet.icsi.berkeley.edu*.

R. Mitkov (ed.) (2002). *Anaphora resolution.* Longman.

M. Palmer, et al. (2005). 'The Proposition Bank: A Corpus Annotated with Semantic Roles'. *Computational Linguistics* **31**(1):71–106.

J. Pustejovsky (1991). 'The Generative Lexicon'. *Computational Linguistics* **17**(4):409–441.

J. Pustejovsky, et al. (1993). 'Lexical semantic techniques for corpus analysis'. *Computational Linguistics* **19**(2):331–358.

J. Pustejovsky, et al. (2006). 'The Brandeis Semantic Ontology: Towards a Generative Lexical Resource'.

J. Ruppenhofer, et al. (2006). 'FrameNet II: Extended Theory and Practice'. *International Computer Science Institute* .

E. M. Voorhees (1994). 'Query expansion using lexical-semantic relations'. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 61–69.

Towards a Factorization of String-Based Phonology

Thomas Graf
Department of Linguistics, UCLA

Abstract. Inspired by the model-theoretic treatment of phonology in Potts & Pullum (2002) and Kracht (2003), we develop an extendable modal logic for the investigation of string-based phonology. In contrast to previous research in this vein (Russell 1993, Kaplan & Kay 1994, Mohri & Sproat 1996), we ultimately strive to study the entire class of such theories rather than merely one particular incarnation thereof. To this end, we first provide a formalization of classic Government Phonology in a restricted variant of temporal logic, whose generative capacity is then subsequently increased by the addition of further operators, moving us along the subregular hierarchy until we reach the regular stringsets. We then identify several other axes along which Government Phonology might be generalized, moving us towards a parametric metatheory of phonology.

Like any other subfield of linguistics, phonology is home to a multitude of competing theories that differ vastly in their conceptual and technical assumptions. Contentious issues are, among others, the relation between phonology and phonetics (and if it is an interesting research question to begin with), if features are privative, binary or attribute valued, if phonological structures are strings or trees, if features can move from one position to another (i.e. if they are *autosegments*), and what role optimality requirements play in determining well-formedness. Meticulous empirical comparisons carried out by linguists have so far failed to yield conclusive results; it seems that for every phenomenon that lends support to certain assumptions, there is another one that refutes them. We do not think that this constitutes a problem to phonological research. Unless we assume that scientific theories can indeed reflect reality as it is rather than merely approximate it, it is to be expected that one theory may fail where another one succeeds and *vice versa*. A similar situation arises in physics, where depending on the circumstances light is thought to exhibit particle-like or wave-like properties.

But given this apparent indeterminacy of theory choice, it is only natural to ask if we can identify classes of interchangeable theories, i.e. proposals which look different superficially but are the same in any other respect. On a bigger scale, this requires developing a metatheory of phonology that uses a finite set of parameters to conclusively determine the equivalence class to which a given phonological theory belongs. This paper aims to lay the basis for such a metatheory using techniques originating in model-theoretic syntax (Blackburn & Meyer-Viol 1994, Kracht 1995, Rogers 2003). We feel obliged to point out in advance that we doubt that a linguistically adequate formal theory of phonology is attainable. However, we also think that in attempting to construct such a metatheory, one gains crucial insights into the core claims about language that are embodied by different phonological assumptions (e.g. computational complexity and memory usage) and how one may translate those claims from one theory into another. Moreover, the explicit logical formalization of linguistic theories allows us to investigate various problems in an algorithmic way using techniques from proof theory and model checking. These insights are relevant to linguists and computer scientists alike. Linguists get a better understanding of how their claims relate to the psychological reality of language, how the different modules of a given theory interact to yield generalizations and how they increase the expressivity of a theory (see Potts & Pullum (2002) for such results on optimality theory). To a limited degree, they also get the freedom to switch to a different theory for specific phenomena without jeopardizing the validity of their framework of choice. Computer scientists, on the other hand, will find that the model-theoretic perspective on phonology eases the computational implementation of linguistic proposals and allows them to gauge their runtime-behavior in advance. Furthermore, they may use the connection between finite model theory and formal language theory to increase the efficiency of their programs by picking the weakest phonological theory that is expressive enough for the task at hand.

This paper is divided into two parts as follows. First, we introduce Government Phonology as an example of a weak theory of phonology and show how it can be axiomatized as a theory of richly annotated string structures using modal logic. In the second part, we analyze several parameters that might have an effect on the generative capacity of our formalization of GP. We show that increasing the power of the spreading operation moves us along the subregular hierarchy and that different types of feature systems have no effect on expressivity in general. We close with a short discussion of two important areas of future research, the impact of the syllable template on generative capacity and the relation between derivational and representational theories.

The reader is expected to have some basic familiarity with formal language theory, non-classical logics and model-theoretic syntax. There is an abundance of introductory material for the former two, while the latter is cogently summarized in Rogers (1996) and Pullum (2007).

1. A Weak Theory of Phonology — Government Phonology

1.1. Informal Overview

Due to space restrictions, we offer but a sketch of the main ideas of Government Phonology (GP), and the reader is advised to check the exposition against the examples in figure 1 on the following page. First, though, a note on our sources is in order. Just like Government-and-Binding theory, GP has changed a lot since its inception and practitioners hardly ever fully specify the details of the version of GP they use. However, there seems to be a consensus that a GP-variant is considered canonical if it incorporates the following modules: government, the syllable template, coda licensing and the ECP from Kaye, et al. (1990), magic licensing from Kaye (1992), and licensing constraints and the revised theory of elements from Kaye (2000). Our general strategy is to follow the definitions in Kaye (2000) as closely as possible and fill in any gaps using the relevant literature. The interested reader might also want to consult Graf (2009) for an in-depth discussion of GP.

In GP, the carrier of all phonological structure is the *skeleton*, a finite, linearly ordered sequence of nodes to which phonological expressions (PEs) can be attached in order to form the *melody* of the structure. A PE is built from a set E of privative features called *elements*, yielding a pair $\langle O, H \rangle$, $O \subseteq E$ a set of *operators*, $H \in E \cup \{\emptyset\}$ the *head*, and $H \notin O$. It is an open empirical question how many features are needed for an adequate account of phonological behavior (Jensen 1994, Harris & Lindsey

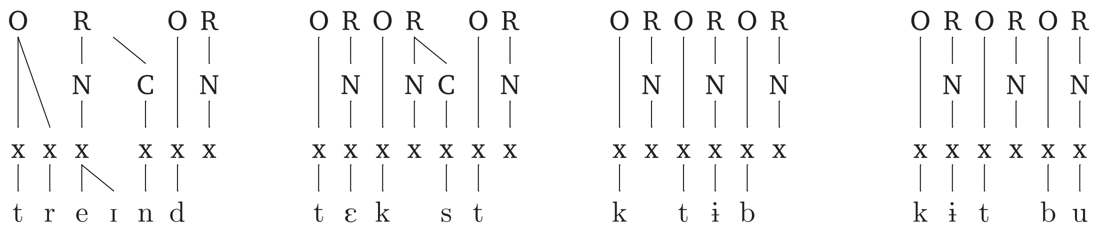


Figure 1: Some phonological structures in GP (with IPA notation)

1995) — Kaye (2000) fixes $E := \{\mathrm{A}, \mathrm{I}, \mathrm{U}, \mathrm{H}, \mathrm{L}, ʔ\}$, but for our axiomatization the only requirement is for E to be finite. The set of licit PEs is further restricted by language-specific *licensing constraints*, i.e. restrictions on the coocurrence of features and their position in the PE. Some examples of PEs are [s] $= \langle \{A, H\}, \emptyset \rangle$, [n] $= \langle \{L, ʔ\}, A \rangle$, [ɨ] $=$ $\langle \emptyset, \emptyset \rangle$, [ɪ] $= \langle \{I\}, \emptyset \rangle$, [i] $= \langle \emptyset, I \rangle$, and [j] $= \langle \emptyset, I \rangle$.

As the last two examples show, every PE is inherently underspecified; whether it is realized as a consonant or a vowel depends on its position in the structure, which is annotated with constituency information. An expression is realized as a vowel if it is associated to a node contained by a *nucleus* (N), but as a consonant if the node is contained by an *onset* (O) or a *coda* (C). Every N constitutes a *rhyme* (R), with C an optional subconstituent of R. All O, N and R may *branch*, that is be associated to up to two nodes (by transitivity of containment, a branching R cannot contain a branching N). Furthermore, word initial O can be *floated*, i.e. be associated to no node at all. The number of PEs per node is limited to one, with the exception of unary branching N, where the limit is two (to model light diphthongs).

All phonological structures are obtained from concatenating $\langle \mathrm{O}, \mathrm{R} \rangle$ pairs according to constraints imposed by two government relations. *Constituent government* restricts the distribution of elements within a constituent, requiring that the leftmost PE licenses all other constituent-internal PEs. *Transconstituent government* enforces dependencies between the constituents themselves. In particular, every branching O has to be licensed by the N immediately following it, and every C has to be licensed by the PE contained in the immediately following O. Even though the precise licensing conditions are not fully worked out for either government relation, the general hypothesis is that PE_i licenses PE_j iff PE_i is leftmost and contained by N, or leftmost and composed from at most as many elements as PE_j and licenses no $PE_k \neq PE_j$ (hence any C has to be followed by a non-branching O, but a branching O might be followed by a branching N or R).

GP also features *empty categories*: a non-coda segment associated solely to the PE $\langle \emptyset, \emptyset \rangle$ can optionally remain unpronounced. For O, this is lexically specified. For N, on the other hand, it is determined by the phonological ECP, which allows only non-branching *p-licensed* N to be mapped to the empty string. N is licensed if it is followed by a coda containing a sibilant (*magic licensing*), or in certain languages if it is the rightmost segment of the string (*final empty nucleus*, abbreviated FEN), or if it is *properly governed* (Kaye 1990). N is properly governed if the first N following it is not p-licensed and no government relations hold between or within any Cs or Os in-between the two Ns.

Finally, GP allows elements to *spread*, just as in fully autosegmental theories (Goldsmith 1976). All elements, though, are assumed to share a single tier, and association lines are allowed to cross. The properties of spreading have not been explicitly spelled out in the literature, but it is safe to assume that it can proceed in either direction and might be optional or obligatory, depending on the element, its position in the string and the language in question. While there seem to be restrictions on the set of viable targets given a specific source, the only canonical one is a ban against spreading within a branching O.

1.2. Formalization in Modal Logic

For our formalization, we use a very weak modal logic that can be thought of as the result of removing the "sometime in the future" and "sometime in the past" modalities from restricted temporal logic (Cohen, et al. 1993, Etessami, et al. 1997).

Let E be some non-empty finite set of *basic elements* different from the neutral element v, which represents the empty set of GP's feature calculus. We define the set of *elements* $\mathscr{E} := (E \times \{1,2\} \times \{head, onset\} \times \{local, spread\}) \cup (\{v\} \times \{1,2\} \times \{head, onset\} \times \{local\})$. The set of melodic features $\mathscr{M} := \mathscr{E} \cup \{\mu, \mathtt{fake}, \checkmark\}$ will be our set of propositional variables. We employ μ (mnemonic for *mute*) and $\checkmark$ to mark unpronounced and licensed segments, respectively, and `fake` for unassociated onsets. For the sake of increased readability, the set of propositional variables is "sorted" such that $x \in \mathscr{M}$ is represented by m, $m \in \mathscr{E}$ by e, heads by h, operators by o. The variable e_n is taken to stand for any element such that $\pi_2(e) = n$, where $\pi_i(x)$ returns the i^{th} projection of x. In rare occasions, we will write $\underline{e}$ and $\overline{e}$ for a specific element e in head and operator position, respectively.

We furthermore use three nullary modalities[1], N, O, C, the set of which we designate by $\mathscr{S}$, read *skeleton*. In addition, we have two unary diamond operators $\lhd$ and $\rhd$, whose respective duals are denoted by $\blacktriangleleft$ and $\blacktriangleright$. The set of well-formed formulas is built up in the usual way from $\mathscr{M}$, $\mathscr{S}$, $\lhd$, $\rhd$, $\rightarrow$ and $\bot$.

Our models $\mathfrak{M} := \langle \mathfrak{F}, V \rangle$ are built over bidirectional frames $\mathfrak{F} := \langle D, R_i, R_\lhd \rangle_{i \in \mathscr{S}}$, where $D \subseteq \mathbb{N}$, and $R_i \subseteq D$ for each $i \in \mathscr{S}$, and $R_\lhd$ is the successor function over $\mathbb{N}$. The valuation function $V : \mathscr{M} \rightarrow \wp(D)$ maps propositional variables to subsets of D. The definition of satisfaction is standard.

$$
\begin{array}{lll}
\mathfrak{M}, w \models \bot & & \text{never} \\
\mathfrak{M}, w \models p & \text{iff} & w \in V(p) \\
\mathfrak{M}, w \models \neg\phi & \text{iff} & \mathfrak{M}, w \not\models \phi \\
\mathfrak{M}, w \models \phi \wedge \psi & \text{iff} & \mathfrak{M}, w \models \phi \text{ and } \mathfrak{M}, w \models \psi \\
\mathfrak{M}, w \models N & \text{iff} & w \in R_N \\
\mathfrak{M}, w \models O & \text{iff} & w \in R_O \\
\mathfrak{M}, w \models C & \text{iff} & w \in R_C \\
\mathfrak{M}, w \models \lhd \phi & \text{iff} & \mathfrak{M}, w+1 \models \phi \\
\mathfrak{M}, w \models \rhd \phi & \text{iff} & \mathfrak{M}, w-1 \models \phi
\end{array}
$$

The formalization of the skeleton is straightforward if we model binary branching constituents as two adjacent unary branching ones and view rhymes as mere nota-

[1] We follow the terminology of Blackburn, et al. (2002) here. Nullary modalities correspond to unary relations and can hence be thought of as propositional constants.

tional devices. Observe that we implement Ns containing diphthongs as single N with both e_1 and e_2 elements associated to it.

S1	$\bigwedge_{i \in \mathscr{S}} (i \leftrightarrow \bigwedge_{i \neq j \in \mathscr{S}} \neg j)$	Unique constituency
S2	$(\blacktriangleleft \bot \rightarrow O) \wedge (\blacktriangleright \bot \rightarrow N)$	Word edges
S3	$R \leftrightarrow (N \vee C)$	Definition of rhyme
S4	$N \rightarrow \lhd O \vee \lhd N$	Nucleus placement
S5	$O \rightarrow \neg \lhd O \vee \neg \rhd O$	Binary branching onsets
S6	$R \rightarrow \neg \lhd R \vee \neg \rhd R$	Binary branching rhymes
S7	$C \rightarrow \lhd N \wedge \rhd O$	Coda placement

GP's feature calculus is also easy to capture. A propositional formula ϕ over a set of variables $x_1, \dots, x_k$ is called *exhaustive* iff $\phi := \bigwedge_{1 \leq i \leq k} \psi_i$, where for every i, ψ_i is either x_i or $\neg x_i$. A PE ϕ is an exhaustive propositional formula over $\mathscr{E}$ such that $\phi \cup \{\mathbf{F1}, \mathbf{F2}, \mathbf{F3}, \mathbf{F4}, \bigvee h\}$ is consistent.

F1	$\bigwedge (h_n \rightarrow \bigwedge_{h_n \neq h'_n} \neg h'_n)$	Exactly one head
F2	$\neg \underline{v} \rightarrow \bigwedge (h_n \rightarrow \bigwedge_{\pi_1(h) = \pi_1(o)} \neg o_n)$	No basic element (except v) twice
F3	$\overline{v} \rightarrow \bigwedge_{o \neq \overline{v}} \neg o$	$\overline{v}$ excludes other operators
F4	$\bigwedge (e_2 \rightarrow \bigvee h_1 \wedge \bigvee o_1)$	Pseudo branching implies first branch

Let *PH* be the least set containing all such ϕ, and let $lic : PH \rightarrow \wp(PH)$ map every ϕ to its set of melodic licensors. By $S \subseteq PH$ we designate the set of PEs occurring in magic licensing configurations (the letter S is mnemonic for "sibilants"). The following five axioms, then, sufficiently restrict the melody.

M1	$\bigwedge_{i \in \mathscr{S}} \left(i \rightarrow \left(\bigvee h_1 \wedge \bigvee o_1\right) \vee \mu \vee \texttt{fake}\right)$	Universal annotation
M2	$((O \vee \lhd N \vee \rhd N) \rightarrow \bigwedge \neg e_2)$	No pseudo branching for O, C & branching N
M3	$O \wedge \lhd O \rightarrow \bigwedge_{\phi \in PH} (\phi \rightarrow \bigvee_{\psi \in lic(\phi)} \lhd \psi)$	Licensing within branching onsets
M4	$C \wedge \bigwedge_{i \in S} \neg i \rightarrow \lhd \neg \mu \wedge \bigwedge_{\phi \in PH} (\phi \rightarrow \bigvee_{\psi \in lic(\phi)} \rhd \psi)$	Melodic coda licensing
M5	$\texttt{fake} \rightarrow O \wedge \bigwedge_{m \neq \texttt{fake}} \neg m$	Fake onsets

Remember that GP allows languages to impose further restrictions on the melody by recourse to licensing constraints. It is easy to see that licensing constraints operating on single PEs can be captured by propositional formulas. The licensing constraint "A must be head", for instance, corresponds to the propositional formula $\neg \overline{A}$. Licensing constraints that extend beyond a single segment can be modeled using $\lhd$ and $\rhd$, provided their domain of application is finitely bounded. See Graf (2009) and the discussion on spreading below for further details.

As mentioned above, we use μ to mark "mute" segments that will be realized as the empty string. The distribution of μ is simple for O and C — the former always allows it, the latter never does. For N, we first need to distribute ✓ in a principled manner across the string to mark the licensed nuclei, which may remain unpronounced. Note that $\underline{v} \wedge \overline{v}$ by itself does not designate unpronounced segments (remember the PE for [ə]), and that unpronounced segments may not contain any other elements (which would affect spreading).

L1	$\mu \rightarrow \neg C \wedge (N \rightarrow \checkmark) \wedge \underline{v} \wedge \overline{v}$	Empty categories
L2	$N \wedge \triangleleft N \rightarrow (\mu \leftrightarrow \triangleleft \mu)$	Licensing of branching nuclei
L3	$O \wedge \triangleleft O \rightarrow \neg \triangleleft \mu \wedge \neg \mu \wedge \neg \triangleright \mu$	Licensing of branching onsets
L4	$N \wedge \checkmark \leftrightarrow \underbrace{\triangleright (C \wedge \bigvee_{i \in S} i)}_{\text{Magic Licensing}} \vee \underbrace{(\neg \triangleleft N \wedge \blacktriangleright \bot)}_{\text{FEN}} \vee$ $\underbrace{((\neg \triangleleft N \rightarrow \triangleleft (\triangleleft N \vee \blacktriangleleft \bot)) \wedge (\neg \triangleright N \rightarrow \triangleright\triangleright (N \wedge \neg \mu)))}_{\text{Proper Government}}$	P-licensing

Axiom **L4** looks daunting at first, but it is easy to unravel. The magic licensing conditions tells us that N is licensed if it is followed by a sibilant in coda position.[2] The FEN condition ensures that wordfinal N are licensed if they are non-branching. The proper government condition is the most complex one, though it is actually simpler than the original GP definition. Remember that N is properly governed if the first N following it is pronounced and neither of the two licenses a branching onset. Also keep in mind that we treat a binary branching constituent as two adjacent unary branching constituents. The proper government condition then enforces a structural requirement such that N (or the first N is we are talking about two adjacent N) may not be preceded by two constituents that are not N and (the second N) may not be followed by two constituents that are not N or not pronounced. Given axioms **S1**–**S7**, this gives the same results as the original constraint.

The last module, spreading, is also the most difficult to accommodate. Most properties of spreading are language specific — only the set of spreadable features and the ban against onset internal spreading are universal. To capture this variability, we define a general spreading scheme σ with six parameters i, j, ω, ϖ, *min* and *max*.

$$\sigma := \bigwedge_{\pi_1(i)=\pi_1(j)} (i \wedge \omega \rightarrow \bigvee_{min \leq n \leq max} \Diamond^n (j \wedge \varpi) \wedge (O \wedge \Diamond O \rightarrow \bigvee_{min+1 \leq n \leq max} \Diamond^n (j \wedge \varpi)))$$

The variables $i, j \in \mathscr{E}$, coupled with judicious use of the formulas ω and ϖ regulate the optionality of spreading. If spreading is optional, i is a spread element and ω, ϖ are formulas describing, respectively, the structural configuration of the target of spreading and the set of licit sources for spreading operations to said target. If spreading is mandatory, then i is a local element and ω, ϖ describe the source and the set of targets. If we want spreading to be mandatory in only those cases where a target is actually available, ω has to contain the subformula $\bigvee_{min \leq n \leq max} \Diamond^n \varpi$. Observe moreover that we need to make sure that every structural configuration is covered by some ω, so that unwanted spreading can be blocked by making ϖ not satisfiable. As further parameters, the finite values $min, max > 0$ encode the minimum and maximum

[2]Note that we can easily restrict the context, if this appears to be necessary for empirical reasons. Strengthening the condition to $\triangleright (C \wedge \bigvee_{i \in S} i) \wedge \triangleleft\blacktriangleleft \bot$, for example, restricts magic licensing to the N occupying the second position in the string.

distance of spreading, respectively. Finally, the operator $\Diamond \in \{\lhd, \rhd\}$ fixes the direction of spreading for the entire formula ($\Diamond^n$ is the n-fold iteration of $\Diamond$). With optional spreading, the direction of the operator is opposite to the direction of spreading, otherwise they are identical.

As the astute reader has probably noticed by now, nothing in our logic prevents us from defining alternative versions of GP. Whether this is a welcome state of affairs is a matter of perspective. On the one hand, the flexibility of our logic ensures its applicability to a wide range of different variants of GP, e.g. to versions where spreading is allowed within onsets or where the details of proper government and the restrictions on branching vary. On the other hand, it begs the question if there isn't an even weaker modal logic that is still expressive enough to formalize GP. The basic feature calculus of GP already requires the logical symbols $\neg$ and $\wedge$, giving us the complete set of logical connectives, and we need $\lhd$ and $\rhd$ to move us along the phonological string. Hence, imposing any further syntactic restrictions on formulas requires advanced technical concepts such as the number of quantifier alternations. However, we doubt that such a move would have interesting ramifications given our goals; we do not strive to find the logic that provides the best fit for a specific theory but to study entire classes of string-based phonological theories from a model-theoretic perspective. In the next section, we try to get closer to this goal.

2. The Parameters of Phonological Theories

2.1. Elaborate Spreading — Increasing the Generative Capacity

It is easy to see that our logic is powerful enough to account for all finitely bounded phonological phenomena (note that this does not imply that GP itself can account for all of them, since certain phenomena might be ruled out by, say, the syllable template or the ECP). In fact, it is even possible to accommodate many long-distance phenomena in a straight-forward way, provided that they can be reinterpreted as arising from iterated application of finitely bounded processes or conditions. Consider for example a stress rule for language L that assigns primary stress to the last syllable that is preceded by an even number of syllables. Assume furthermore that secondary stress in L is trochaic, that is to say it falls on every odd syllable but the last one. Let 1 and 2 stand for primary and secondary stress, respectively. Unstressed syllables are assigned the feature 0. Then the following formula will ensure the correct assignment of primary stress (for the sake of simplicity, we assume that every node in the string represents a syllable; it is an easy but unenlightening exercise to rewrite the formula for our GP syllable template).

$$\bigvee_{i \in \{0,1,2\}} i \wedge \bigwedge_{i \neq j \in \{0,1,2\}} (i \rightarrow \neg j) \wedge (\blacktriangleleft \bot \rightarrow 1 \vee 2) \wedge (2 \rightarrow \rhd\, 0) \wedge$$
$$(0 \rightarrow \rhd\, (1 \vee 2) \vee \blacktriangleright \bot) \wedge (1 \rightarrow \neg \lhd 1 \wedge (\blacktriangleright \bot \vee \rhd \blacktriangleright \bot))$$

Other seemingly unbounded phenomena arising from iteration of local processes, most importantly vowel harmony (see Charette & Göksel (1996) for a GP analysis), can be captured in a similar way. However, there are several unbounded phonological phenomena that require increased expressivity (see Graf (2009) for details). As we are only concerned with string structures, it is a natural move to try to enhance our language with operators from more powerful string logics, in particular, linear temporal logic.

The first step is the addition of two operators $\triangleleft^+$ and $\triangleright^+$ with the corresponding relation $R^+_{\triangleleft}$, the transitive closure of $R_{\triangleleft}$. This new logic is exactly as powerful as restricted temporal logic (Cohen et al. 1993), which in turn has been shown in Etessami et al. (1997) to exactly match the expressivity of the two-variable fragment of first-order logic (see Weil (2004) for further equivalence results). Among other things, OCP effects (Leben 1973, Goldsmith 1976) can now be captured in an elegant way. The formula $O \wedge \underline{A} \wedge \overline{L} \wedge \overline{ʔ} \rightarrow \triangleright^+ \neg(O \wedge \underline{A} \wedge \overline{ʔ})$, for example, disallows alveolar nasals to be followed by another alveolar stop, no matter how far the two are apart.

But $\triangleleft^+$ and $\triangleright^+$ are too coarse for faithful renditions of unbounded spreading. For example, it is not possible to define all intervals of arbitrary size within which a certain condition has to hold (e.g. no b may appear between a and c). As a remedy, we add the until and since operators U and S familiar from linear temporal logic, granting us the power of full first-order logic. This enables us to define all star-free languages (McNaughton & Pappert 1971, Thomas 1979, Cohen 1991, Cohen et al. 1993). These feature a plethora of properties that make them very attractive for purposes of natural language processing. Moreover, the only phenomenon known to the author that exceeds their confines is stress assignment in Cairene Arabic and Creek, which basically works like the stress assignment system outlined above — with the one exception that secondary stress is not marked overtly (Mitchell 1960, Haas 1977). Under these conditions, assigning primary stress involves counting modulo 2, which is undefinable in first-order logic, whence a more powerful logic is needed. The next step up from the star-free stringsets are the regular languages, which can count modulo n. From previous research, we know that the regular stringsets are identical to the set of finite strings definable in monadic second order logic (MSO) (Büchi 1960), linear temporal logic with modal fixed point operators (Vardi 1988) or regular linear temporal logic (Leucker & Sánchez 2005). In linguistic terms, this corresponds to spreading being capable of picking its target based on more elaborate patterns.

A caveat is in order, though. Thatcher (1967) proved that every recognizable set is a projection of some local set. Thus the hierarchy outlined above collapses if we grant ourselves an arbitrary number of additional features to encode all the structural properties our logic cannot express. In the case of primary stress in Cairene Arabic and Creek, for instance, we could just use the feature for secondary stress assignment even though secondary stress seems to be absent in these languages. Generally speaking, we can reinterpret any unbounded dependency as a result of iterated local processes by using "invisible" features. Therefore, all claims about generative capacity hold only under the proviso that all such spurious coding-features are being eschewed.

We have just seen that the power of GP can be extended along the subregular hierarchy, up to the power of regular languages, and that there seems to be empirical motivation to do so. Interestingly, it has been observed that SPE yields regular languages, too (Johnson 1972, Kaplan & Kay 1994). But even the most powerful rendition of GP defines only a proper subset of the stringsets derivable in SPE, apparently due to its restrictions on the feature system, the syllable template and its government requirements. The question we face, then, is whether we can generalize GP in these regards, too, to push it to the full power of SPE and obtain a multidimensional vector space of phonological theories.

2.2. Feature Systems

The restriction to privative features is immaterial. A set of PEs is denoted by some propositional formula over $\mathscr{E}$, and the boolean closure of $\mathscr{E}$ is isomorphic to $\wp(\mathscr{E})$. But Keenan (2008, 81–109) shows that a binary feature system using a set of features $\mathscr{F}$ can be modeled by the powerset algebra $\wp(\mathscr{F})$, too. So if $|\mathscr{E}| = |\mathscr{F}|$, then $\wp(\mathscr{E}) \cong \wp(\mathscr{F})$, whence the two feature systems are isomorphic. The same result holds for systems using more than two feature values, provided their number is finitely bounded, since multivalued features can be replaced by a collection of binary valued features given sufficient co-occurrence restrictions on feature values (which can easily be formalized in propositional logic).

One might argue, though, that the core restriction of privative feature systems does not arise from the feature system itself but from the methodological principle that absent features, i.e. negative feature values, behave like constituency information and cannot spread. In general, though, this is not a substantial restriction either, as for every privative feature system $\mathscr{E}$ we can easily design a privative feature system $\mathscr{F} := \{e^+, e^- \mid e \in \mathscr{E}\}$ such that $\mathfrak{M}, w \models e^+$ iff $\mathfrak{M}, w \models e$ and $\mathfrak{M}, w \models e^-$ iff $\mathfrak{M}, w \models \neg e$. Crucially, though, this does not entail that the methodological principle described above has no impact on expressivity when the set of features is fixed across all theories, which is an interesting issue for future research.

2.3. Syllable Template

While GP's syllable template could in principle be generalized to arbitrary numbers and sizes of constituents, a look at competing theories such as SPE and Strict CV (Lowenstamm 1996, Scheer 2004) shows that the number of different constituents is already more than sufficient. This is hardly surprising, because GP's syllable template is modeled after the canonical syllable template, which in general is thought not to be in need of further refinement. Consequently, we only need to lift the restriction on the branching factor and allow theories not to use all three constituent types. SPE then operates with a single N constituent of unbounded size, whereas Strict CV uses N and O constituents of size 1. Regarding the government relations, the idea is to let every theory fix the branching factor b for each constituent and the maximum number l of licensees per head. Every node within some constituent has to be constituent licensed by the *head*, i.e. the leftmost node of said constituent. Similarly, all nodes in a coda or non-head position have to be transconstituent licensed by the head of the following constituent. For every head the number of constituent licensees and transconstituent licensees, taken together, may not exceed l.

Even from this basic sketch it should already be clear that the syllable template can have a negative impact on expressivity, but only under the right conditions. For instance, if our feature system is set up in a way such that every symbol of our alphabet is to be represented by a PE in N (as happens to be the case for SPE), restrictions on b and l are without effect. Thus one of the next stages in this project will revolve around determining under which conditions the syllable template has a monotonic effect on generative capacity.

2.4. Representations versus Derivations

One of the most striking differences between phonological theories is the distinction between representational and derivational ones, which begs the question how we can ensure comparability between these two classes. Representational theories are naturally captured by our declarative, model-theoretic approach, whereas derivational theories are usually formalized as regular relations (Kaplan & Kay 1994, Mohri & Sproat 1996), which resist being recast in logical terms due to their closure properties. For SPE, one can use a coding trick from two-level phonology (Koskenniemi 1983) and use an unpronounced feature like μ to ensure that all derivationally related strings have the same length. SPE can be then be interpreted as language over pairs and hence cast in MSO terms, which was successfully done by Vaillette (2003). Unfortunately, it is unclear how this method could be extended to subregular grammars. At the same time, no other open issue is of greater importance to the success of this project.

3. Conclusion

The purpose of this paper was to lay the foundation for a general framework in which string-based phonological theories can be matched against each other. We started with a modal logic which despite its restrictions was still perfectly capable of defining a rather advanced and intricate phonological theory. We then tried to generalize the theory along several axes, some of which readily lent themselves to conclusive results while others didn't. We saw that the power of spreading, by virtue of being an indicator of the necessary power of the description language, has an immediate and monotonic effect on generative capacity. Feature systems, on the other hand, were shown to be a negligible factor in theory comparisons; it remains an open question if the privativity assumption might affect generative capacity when the set of features is fixed. A detailled study of the effects of the syllable template also had to be deferred to later work. The most pressing issue in our opinion, though, is the translation from representational to derivational theories. Not only will it enable us to reconcile two supposedly orthogonal perspectives on phonology, but it also allows us to harvest results on finite-state OT (Frank & Satta 1998) to extend the framework to optimality theory. Even though a lot of work remains to be done and not all of our goals may turn out be achievable, we are confident that a model-theoretic approach provides an interesting new perspective on long-standing issues in phonology.

Acknowledgements

The following people had a positive influence on this paper in one way or the other: Anonymous reviewer 1 & 2, Bruce Hayes, Ed Keenan, Marcus Kracht, Ed Stabler, Kie Zuraw, and the members of the UCLA phonology seminar during the winter quarter 2009.

References

P. Blackburn, et al. (2002). *Modal Logic*. Cambridge University Press.

P. Blackburn & W. Meyer-Viol (1994). 'Linguistics, logics, and finite trees.'. *Logic Journal of the IGPL* **2**:3–29.

J. R. Büchi (1960). 'Weak second-order arithmetic and finite automata'. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* **6**:66–92.

M. Charette & A. Göksel (1996). 'Licensing Constraints and Vowel Harmony in Turkic languages'. *SOAS Working Papers In Linguistics and Phonetics* **6**:1–25.

J. Cohen (1991). 'On the Expressive Power of Temporal Logic for Infinite Words'. *Theoretical Computer Science* **83**:301–312.

J. Cohen, et al. (1993). 'On the expressive power of temporal logic'. *Journal of Computer and System Sciences* **46**:271–294.

K. Etessami, et al. (1997). 'First-order logic with two variables and unary temporal logic'. In *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science*, pp. 228–235.

R. Frank & G. Satta (1998). 'Optimality Theory and the Generative Complexity of Constraint Violability'. *Computational Linguistics* **24**:307–315.

J. Goldsmith (1976). *Autosegmental Phonology*. Ph.D. thesis, MIT.

T. Graf (2009). 'Comparing Incomparable Frameworks — A Model Theoretic Approach to Phonology'. To appear in the proceedings of the 33rd Penn Linguistics Colloquium. Handout version available online: http://tgraf.bol.ucla.edu/doc/talks/PLC33_mtp.pdf. Last time checked: June 14, 2009.

M. R. Haas (1977). 'Tonal Accent in Creek'. In L. M. Hyman (ed.), *Southern California Occasional Papers in Linguistics*, vol. 4, pp. 195–208. University of Southern California, Los Angeles. Reprinted in **?**).

J. Harris & G. Lindsey (1995). 'The Elements of Phonological Representation'. In J. Durand & F. Katamba (eds.), *Frontiers of Phonology*, pp. 34–79. Longman, Harlow, Essex.

S. Jensen (1994). 'Is ʔ an Element? Towards a Non-Segmental Phonology'. *SOAS Working Papers In Linguistics and Phonetics* **4**:71–78.

C. D. Johnson (1972). *Formal Aspects of Phonological Description*. Mouton, The Hague.

R. M. Kaplan & M. Kay (1994). 'Regular Models of Phonological Rule Systems'. *Computational Linguistics* **20**(3):331–378.

J. Kaye (1990). 'Government in Phonology: the case of Moroccan Arabic'. *The Linguistic Review* **6**:131–159.

J. Kaye (1992). 'Do you Believe in Magic? The story of s+C sequences'. *Working Papers in Linguistics and Phonetics* **2**:293–313.

J. Kaye (2000). 'A User's Guide to Government Phonology'. Unpublished manuscript. Available online: http://134.59.31.7/∼scheer/scan/Kaye00guideGP.pdf. Last time checked: June 14, 2009.

J. Kaye, et al. (1990). 'Constituent structure and government in phonology'. *Phonology Yearbook* **7**:193–231.

E. Keenan (2008). 'Mathematical Structures in Language'. Ms., University of California, Los Angeles.

K. Koskenniemi (1983). 'Two-level Morphology: A General Computational Model For Word-Form Recognition and Production'. Publication 11.

M. Kracht (1995). 'Syntactic Codes and Grammar Refinement'. *Journal of Logic, Language and Information* **4**:41–60.

M. Kracht (2003). 'Features in Phonological Theory'. In B. Löwe, W. Malzkorn, & T. Räsch (eds.), *Foundations of the Formal Sciences II, Applications of Mathematical Logic in Philosophy and Linguistics*, no. 17 in Trends in Logic, pp. 123–149. Kluwer, Dordrecht. Papers of a conference held in Bonn, November 11–13, 2000.

W. Leben (1973). *Suprasegmental Phonology*. Ph.D. thesis, MIT.

M. Leucker & C. Sánchez (2005). 'Regular Linear Temporal Logic'. In *Proceedings of The 4th International Colloquium on Theoretical Aspects of Computing (ICTAC'07)*, no. 4711 in Lecture Notes in Computer Science, pp. 291–305.

J. Lowenstamm (1996). 'CV as the Only Syllable Type'. In J. Durand & B. Laks (eds.), *Current Trends in Phonology: Models and Methods*, pp. 419–421. European Studies Research Institute, University of Salford.

R. McNaughton & S. Pappert (1971). *Counter-Free Automata*. MIT Press, Cambridge, Mass.

T. F. Mitchell (1960). 'Prominence and Syllabification in Arabic'. *Bulletin of the School of Oriental and African Studies* **23**(2):369–389.

M. Mohri & R. Sproat (1996). 'An Efficient Compiler for Weighted Rewrite Rules'. In *In 34th Annual Meeting of the Association for Computational Linguistics*, pp. 231–238.

C. Potts & G. K. Pullum (2002). 'Model Theory and the Content of OT Constraints'. *Phonology* **19**(4):361–393.

G. K. Pullum (2007). 'The Evolution of Model-Theoretic Frameworks in Linguistics'. In J. Rogers & S. Kepser (eds.), *Model-Theoretic Syntax @ 10*, pp. 1–10.

J. Rogers (1996). 'A Model-Theoretic Framework for Theories of Syntax'. In *Proceedings of the 34th Annual Meeting of the ACL*, pp. 10–16, Santa Cruz, USA.

J. Rogers (2003). 'Syntactic Structures as Multi-Dimensional Trees'. *Research on Language and Computation* **1**(1):265–305.

K. Russell (1993). *A Constraint-Based Approach to Phonology*. Ph.D. thesis, University of Southern California.

T. Scheer (2004). *A Lateral Theory of Phonology: What is CVCV and Why Should it be?* Mouton de Gruyter, Berlin.

J. W. Thatcher (1967). 'Characterizing Derivation Trees for Context-Free Grammars Through a Generalization of Finite Automata Theory'. *Journal of Computer and System Sciences* **1**:317–322.

W. Thomas (1979). 'Star-Free Regular Sets of ω-sequences'. *Information and Control* **42**:148–156.

N. Vaillette (2003). 'Logical Specification of Regular Relations for NLP'. *Natural Language Engineering* **9**(1):65–85.

M. Y. Vardi (1988). 'A Temporal Fixpoint Calculus'. In *Proceedings of the 15th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pp. 250–259.

P. Weil (2004). 'Algebraic Recognizability of Languages'. In J. Fiala, V. Koubek, & J. Kratochvíl (eds.), *Mathematical Foundations of Computer Science 2004*, vol. 3153 of *Lecture Notes in Computer Science*, pp. 149–175. Springer, Berlin.

VARIABLE SELECTION IN LOGISTIC REGRESSION: THE BRITISH ENGLISH DATIVE ALTERNATION

Daphne Theijssen
Department of Linguistics, Radboud University Nijmegen

Abstract. In this paper, we address the problem of selecting the 'optimal' variable subset in a logistic regression model for a medium-sized data set. As a case study, we take the British English dative alternation, where speakers and writers can choose between two (equally grammatical) syntactic constructions to express the same meaning. With the help of 29 explanatory variables taken from the literature, we build two types of models: (1) with the verb sense included as a random effect (verb senses often have a bias towards one of the two variants), and (2) without a random effect. For each type, we build three different models by including all variables and keeping the significant ones, by sequentially adding the most predictive variable (forward regression), and by sequentially removing the least predictive variable (backward regression). Seeing that the six approaches lead to five different models, we advise researchers to be careful to base their conclusions solely on the one 'optimal' model they found.

1. Introduction

There are many linguistic phenomena that researchers have tried to explain on the basis of different partially explanatory features. Probabilistic modelling techniques can help in combining these explanatory features and testing the combination on corpus data. A popular – and rather successful – technique for this purpose is logistic regression modelling. However, *how* exactly the technique is best employed for this type of research remains an open question.

Statistical models built using corpus data do precisely what they are designed to do: find the 'best possible' model for a specific data set given a specific set of explanatory features. The issue that probabilistic techniques model data (while one would actually want to model underlying processes) is only aggravated by the fact that the variables are usually not mutually independent. As a consequence, one set of data and explanatory features can result in different models, depending on the details of the model building process.

Building a regression model consists of three main steps: (1) deciding which of the explanatory features should actually be included as variables in the model formula, (2) establishing the coefficients (weights) for the variables, and (3) evaluating the model. The first step is generally referred to as *variable selection* and is the topic of the current paper.

Researchers have employed at least three different approaches to variable selection: (1) first building a model on all available explanatory features and then keeping/reporting those that have a significant contribution (e.g. Bresnan, et al. (2007)), (2) sequentially adding the most explanatory feature (forward), until no significant gain is obtained anymore (e.g. Grondelaers & Speelman (2007)), and (3) starting with a model containing all available features, and (backward) sequentially removing those that yield the lowest contribution (e.g. Blackwell (2005)). In general, researchers report on only one (optimal) model without giving clear motivations for their choices.

In this paper, we compare the three approaches in a case study: we apply them to a set of 915 instances of the British English dative alternation, taken from the British component of the ICE Corpus. In the dative alternation, speakers choose between the double object (1) and the prepositional dative variant (2).

1. She handed the student the book.
2. She handed the book to the student.

The variables (explanations suggested in the literature) are taken from Bresnan et al's work on the dative alternation in American English.

Previous research (for example, Gries & Stefanowitsch (2004)) has indicated that the verb sense often predicts a preference for one of the two constructions. However, contrary to the fourteen explanatory features suggested by Bresnan et al, which can be treated as fixed variables because of their small number of values (often only two), *verb sense* has so many different values that it cannot be treated as a variable in a regression model. Recently developed logistic regression models can handle these lexical biases by treating verb sense as a random effect (e.g. Bresnan et al. (2007)). In order to examine the effect of building such mixed models, we create models with and without a random effect in each of the three approaches. This leads to a total of six different models.

Our goal is to investigate the role of a random effect in a model of syntactic variation built with a medium-sized set of observations. In addition, we want to investigate whether it is justified to report only one 'optimal' regression model, if models can be built in three different ways. The case of the British English dative alternation is used to illustrate the issues and results.

The structure of this paper is as follows: A short overview of the related work can be found in Section 2. The data is described in 3. In Section 4, we explain the method applied. The results are shown and discussed in Section 5. In the final Section 6, we present our conclusions.

2. Related work

2.1. The dative alternation

Bresnan et al. (2007) built various logistic regression models for the dative alternation based on 2360 instances they extracted from the three-million word Switchboard Corpus of transcribed American English telephone dialogues (Godfrey, et al. 1992). With the help of a logistic mixed-effect regression model with verb as a random effect, they were able to explain 95% of the variation. To test how well the model generalizes to previously unseen data, they built a model on 2000 instances randomly selected from the total set, and tested on the remaining 360 cases. Repeating this 100 times, 94% of the test cases on average were predicted correctly.

Many of the variables in the model concern the two objects in the construction (*the student* and *the book* in example 1 and 2). In prepositional dative constructions, the object first mentioned is the theme (*the book*), and the second object the recipient (*the student*). In double object constructions, the recipient precedes the theme. Bresnan et al. found that the first object is typically (headed by) a pronoun, mentioned previously in the discourse (*given*), animate, definite and longer than the second object. The characteristics of the second object are generally the opposite: nongiven, nonpronominal, inanimate and indefinite.

According to Haspelmath (2007), there is a slight difference between the dative alternation as it occurs in British English and in American English. When the theme is a pronoun, speakers of American English tend to allow only the prepositional dative construction. In British English, clauses such as *She gave me it* and even *She gave it me* are also acceptable.

Gries (2003) performed analyses with multiple variables that are similar to those in Bresnan et al. (2007), but applying a different technique (linear discriminant analysis or LDA) on a notably smaller data set consisting of only 117 instances from the British National Corpus (Burnard 2007). The LDA model is trained on all instances, and is able to predict 88.9% of these cases correctly (with a majority baseline of 51.3%). There is no information on how the model performs on previously unseen data.

Gries & Stefanowitsch (2004) investigated the effect of verb biases in 1772 instances from the ICE-GB Corpus (Greenbaum 1996). When predicting the preferred dative construction for each verb, 82.2% of the constructions could be predicted correctly. It thus outperforms the majority baseline of 65.0% (always choosing the overall most frequent variant).

2.2. Variable selection in logistic regression

A recent and extensive textbook on modern statistical techniques is that by Izenman (2008). In chapter 5, he explains that variable selection is often needed to arrive at an interpretable model that reaches an acceptable prediction accuracy. Keeping too many variables may lead to overfitting, while a simpler model may suffer from underfitting. The risk of applying variable selection is that it optimizes the model for a particular data set. Using a slightly different data set may result in a completely different variable subset.

An approach to variable selection that is commonly used in linguistics is stepwise adding the most predictive variables to an empty model (e.g. Grondelaers & Speelman (2007)) or stepwise removing the least predictive variables from the full model (e.g. Blackwell (2005)). The main criticisms on these methods are (1) that the results are difficult to interpret when the variables are highly correlated, (2) that deciding which variable to remove or add is not trivial, (3) that both methods may result in two different models that may not even be optimal, and (4) that each provides a single model, while there may be more than one optimal subset (Izenman 2008).

Another approach Izenman mentiones in the same section is to build all models with each possible subset and select those with the best results. An important objection to this approach is that it is computationally expensive to carry out. For this reason, we do not employ this method.

Instead, we follow Sheather (2009), who builds a model containing all variables that he expects to contribute to the model, and removes the insignificant ones (chapter 8). These expectations are based on plots of the variables that he made beforehand. Where desirable, he transformed the variables to give them more predictive power (e.g. by taking their log). As indicated by Izenman (2008), variable selection on the basis of a data set may lead to a model that is specific for that particular set. Since we also want to be able to compare our models to those found by Bresnan et al. (2007), we refrain from such preprocessing and use all variables they used in the variable selection process.

3. Data

Since we study a syntactic phenomenon, it is convenient to employ a corpus with detailed (manually checked) syntactic annotations. We selected the one-million-word British component of the ICE Corpus, the ICE-GB, containing both written and (transcribed) spoken language (Greenbaum 1996).

We used a Perl script to automatically extract potentially relevant clauses from the ICE-GB. These were clauses with an indirect and a direct object (double object) and clauses with a direct object and a prepositional phrase with the preposition *to* (prepositional dative). Next, we manually checked the extracted sets of clauses and removed irrelevant clauses such as those where the preposition *to* had a locative function (e.g. *Fold the short edges to the centre.*).

Following Bresnan et al. (2007), we ignored constructions with a preposition other than *to*, with a clausal object, with passive voice and with reversed constructions. To further limit the influence of the syntactic environment of the construction, we decided to exclude variants in imperative and interrogative clauses, as well as those with phrasal verbs (e.g. *to hand over*). Coordinated verbs or verb phrases were also removed. The characteristics of the resulting data sets can be found in Table 1.

Table 1: Characteristics of the data sets

		nr of instances		
Type	Corpus	d.obj.	pr.dat.	Total
Spoken British English	ICE-GB	399	151	550
Written British English	ICE-GB	263	102	365
Total	ICE-GB	662	253	915

4. Method

4.1. Explanatory features

We adapt the explanatory features and their definitions from Bresnan et al. (2007) (Table 2), and manually annotate our data set following an annotation manual based on these definitions[1].

The table includes one new variable: *medium*. This tells us whether the construction was found in written or spoken text. It may well be that certain variables only play a role in one of the two mediums. In order to test this, we include the 14 (two-way) interactions between the variables taken from Bresnan et al. and the medium[2]. This leads to a total number of 29 features.

As mentioned in the Introduction, we will build models with and without including verb sense as a random effect. The verb sense is the lemma of the verb together with its semantic class, e.g. *pay_a* for *pay* with an abstract meaning and *pay_t* when *pay* is used to describe a transfer of possession. In total, our data set contains 94 different verb senses

[1]The annotation manual is available online: http://lands.let.ru.nl/∼daphne/downloads.html.

[2]We are aware of the fact that there are other ways to incorporate the medium in the regression models, for instance by building separate models for the written and the spoken data. Since the focus of this paper is on the three approaches in combination with the presence or absence of a random effect, we will limit ourselves to the method described.

Table 2: Features and their values (th=theme, rec=recipient). All nominal variables are transformed into binary variables with values 0 and 1. As a result, semantic verb class (*communication*, *abstract* or *transfer of possession*) is split into two effects: *verb=abstract* (0 or 1) and *verb=communication* (0 or 1). Cases with semantic verb class *transfer of possession* have value 0 for both variables.

Feature	Values	Description
rec = animate	1, 0	human or animal, or not
th = concrete	1, 0	with fixed form and/or space, or not
rec,th = definite	1, 0	definite pronoun, proper name or noun preceded by a definite determiner, or not
rec,th = given	1, 0	mentioned or evoked $\leq$20 clauses before, or not
length difference	-3.4-4.2	$\ln$(#words in th) $-$ $\ln$(#words in rec)
rec,th = plural	1, 0	plural in number, or not (singular)
rec = local	1, 0	first or second person (*I*, *you*), or not
rec,th = pronominal	1, 0	headed by a pronoun, or not
verb = abstract	1, 0	*give it some thought* is abstract, *tell him a story* is
verb = communication	1, 0	communication, *give him the book* is transfer
structural parallellism = present	1, 0	same variant used previously, or not
medium = written	1, 0	type of data is written, or not (spoken)

(derived from 65 different verbs). The distribution of the verb senses with 5 or more occurrences can be found in Table 3. As predicted by Gries & Stefanowitsch (2004), many verb senses show a bias towards one of the two constructions. The verb *pay* shows a clear bias towards the prepositional dative construction when it has an abstract meaning, but no bias when the literal transfer of possession is meant.

Table 3: Distribution of verb senses with 5 or more occurrences in the data set. The verb senses in the right-most list have a clear bias towards the double object (d.obj.) construction, those in the left-most for the prepositional dative (p.dat.) construction, and those in the middle show no clear preference. The *a* represents *abstract*, *c communication* and *t transfer of possession*.

# d.obj. $>$ # p.dat.			# d.obj. $\approx$ # p.dat.			# d.obj. $<$ # p.dat.		
verb sense	d.obj.	p.dat.	verb sense	d.obj.	p.dat.	verb sense	d.obj.	p.dat.
give_a	252	30	*do_a*	8	9	*pay_a*	2	12
give_c	65	10	*send_c*	9	7	*cause_a*	5	8
give_t	53	21	*lend_t*	8	7	*sell_t*	0	10
tell_c	67	1	*pay_t*	6	5	*owe_a*	2	6
send_t	41	15	*leave_a*	5	4	*explain_c*	0	6
show_c	37	9	*write_c*	4	5	*present_c*	0	6
offer_a	23	9				*read_c*	1	4
show_a	6	1						
offer_t	6	0						
tell_a	6	0						
wish_c	6	0						
bring_a	4	1						
bring_t	3	2						
hand_t	3	2						

4.2. Variable selection

Using the values of the variables (and the random effect), we establish a regression function that determines the log of the odds that the construction C in clause i (with verb sense j) is a prepositional dative. The prepositional dative is regarded a success (with value 1), while the double object construction is a failure (0). The regression function is defined as follows:

$$\ln odds(C_{ij} = 1) = \alpha + \sum_{k=1}^{29} (\beta_k V_{ijk}) \quad (+r_j). \tag{1}$$

The α is the intercept of the function. $\beta_k V_{ijk}$ are the weights β and values V_{ij} of the 29 variables k. The optional random effect r_j is normally distributed with mean zero ($r_j \sim N(0, \sigma)$). The optimal values for the function parameters α, β_k and r_j are found with the help of Maximum Likelihood Estimation[3]. The outcome of the regression enables us to use the model as a classifier: all cases with $\ln odds(C_{ij} = 1) \geq t$ are classified as prepositional dative, all with $\ln odds(C_{ij} = 1) < t$ as double object. The letter t is a threshold, which we set to 0. With this threshold, all instances for which the regression function outputs a negative log odds are classified as double object constructions, all other instances as prepositional dative.

In the first approach, we first include all 29 features in the model formula. We then remove all variables that do not have a significant effect in the model output, and build a model with the remaining (significant) variables.

For the second approach, being forward sequential regression, we start with an empty model and sequentially add the variable that is most predictive. As Izenman (2008) warns us, deciding which variable to keep is not trivial. We decide to keep the variable that yields the highest area under the ROC (Receiver Operating Characteristics) curve. This curve is a plot of the correctly classified positive instances (prepositional dative) and the incorrectly classified positive instances. The area under it (AUC) gives the probability that the regression function, when randomly selecting a positive (prepositional dative) and a negative (double object) instance, outputs a higher log odds for the positive instance than for the negative instance. The AUC is thus an evaluation measure for the quality of a model. It is calculated with:

$$\frac{average\_rank(x_{C=1}) - \frac{p+1}{2}}{n - p}, \tag{2}$$

where $average\_rank(x_{C=1})$ is the average rank of the instances x that are prepositional dative (when all instances are ranked numerically according to the log odds), p the number of prepositional dative instances, and n the total number of instances[4]. We add the most predictive variables to the model as long as it gives an improvement over the AUC of the model without the variable. An interaction of variable A with *Medium* is only included when the resulting AUC is higher than that reached after adding the single variable A[5]. Two AUC are considered different when the difference is higher than a threshold. We set the threshold to 0.002.[6]

For the third approach (backward sequential regression), we use the opposite procedure: we start with the full model, containing all 29 features, and sequentially leave out the variable A that yields the model with the highest AUC that is not lower than the AUC for the model with A. When the AUC of a model without variable A does not differ from

[3]We use the functions `glm()` and `lmer()` (Bates 2005) in `R` (R Development Core Team 2008).

[4]We use the function `somers2()` created in `R` (R Development Core Team 2008) by Frank Harrell.

[5]When including an interaction but not the main variables in it, the interaction will also partly explain variation that is caused by the main variables (Rietveld & van Hout 2008).

[6]The threshold value has been established experimentally.

the AUC of the model without the interaction of A with Medium, we remove the interaction. Again, AUCs are only considered different when the difference is at least the threshold (again set to 0.002).

We evaluate the models with and without random effects by establishing the model quality (training and testing on all 915 cases) by calculating the percentage of correctly classified instances (accuracy) and the area under the ROC curve (AUC). Also, we determine the prediction accuracy reached in 10-fold cross-validation (10 sessions of training on 90% of the data and testing on the remaining 10%) in order to establish how well the model generalizes to previously unseen data. In the 10-fold cross-validation setting, we provide the algorithms with the variables selected in the models trained on all 915 cases. The regression coefficients for these subsets of variables are then estimated for each separate training set.

The coefficients in the regression models help us understand which variables play what role in the dative alternation. We will therefore compare the coefficients of the significant effects in the models built on all 915 instances.

5. Results

5.1. Mixed models

Table 4 gives the model fit and prediction accuracy for the different regression models we built, including verb sense as a random effect. The prediction accuracy (the percentage of correctly classified cases) is significantly higher than the majority baseline (always selecting NP-NP) in all settings, also when testing on new data ($p < 0.001$ for the three models, Wilcoxon paired signed rank test).

Table 4: Model fit and prediction accuracy of the regression models with *verb sense* as a random effect

			model fit (train=test)		*10-fold cv*
selection	#variables	baseline	AUC	accuracy	aver. accuracy
1. significant	5	0.723	0.979	0.936	0.825
2. forward	4	0.723	0.980	0.938	0.832
3. backward	4	0.723	0.980	0.938	0.832

When training and testing on all 915 instances, the mixed models reach a considerable AUC and prediction accuracy (model quality). However, seeing the decrease in accuracy in a 10-fold cross-validation setting, it seems that the mixed models do not generalize well to previously unseen data.

The significant effects in the models resulting from the three approaches are presented in Table 5. The directions of the main effects are the same as those for American English (Bresnan et al. 2007), as presented in Section 2.1.

The forward (2) and backward (3) selection approaches lead to the same regression model. The differences between this model and the one obtained by keeping the significant variables (1) may be caused by the fact that the information contained in the variables shows considerable overlap. For instance, pronominal objects are also typically discourse given. A significant effect for the one variable may therefore decrease the possibility of

Table 5: Coefficients of significant effects in (mixed) regression models with verb sense as random effect, trained on all 915 instances, *** p<0.001 ** p<0.01 * p<0.05. The last column indicates the direction towards the prepositional dative (p.dat.) and double object construction (d.obj.).

Effect	1. significant		2. forward		3. backward		
th=pronominal, medium=written	-2.01	*					
length difference	-2.52	***	-2.41	***	-2.41	***	d.obj.
rec=local	-2.68	***	-1.86	***	-1.86	***	↑
rec=given			-1.48	***	-1.48	***	|
th=definite	1.67	***					|
th=pronominal	2.15	***					↓
th=given			2.32	***	2.32	***	p.dat.
(intercept)	1.27	**	2.53	***	2.53	***	

regarding the other as significant. This is exactly what we see: the model obtained through the two stepwise approaches contains a variable denoting the givenness of the theme but none describing its pronominality, while it is the other way around for the model with the significant variables from the full model.

Only the model obtained by keeping the significant variables in the full model contains an interaction, namely that between medium and a pronominal theme. The main effect (without medium) is also included, but it shows the opposite effect. When the theme is pronominal, speakers tend to use the prepositional dative construction (coefficient 2.15). This effect seems much less strong in writing (remaining coefficient 2.15 - 2.01 = 0.14). Whether there really exists a difference in the effect of the pronominality of the theme in speech and writing is not clear, since only one model shows this difference.

What also remains unclear, is which of the two models is more suitable for explaining the British English dative alternation. Seeing the differences between the significant effects found in the two models we found, and the relatively low prediction accuracy in 10-fold cross-validation, it seems that the models are modelling the specific data set rather than the phenomenon. A probable cause is that the mixed models are too complex to model a data set consisting of 915 instances. In the next section, we apply the three approaches to build simpler models, namely without the random effect.

5.2. Models without a random effect

The model fit and prediction accuracy for the models without a random effect can be found in Table 6.

Table 6: Model fit and prediction accuracy of the regression models without a random effect

			model fit (train=test)		*10-fold cv*
selection	#variables	baseline	AUC	accuracy	aver. accuracy
1. significant	5	0.723	0.934	0.882	0.882
2. forward	5	0.723	0.941	0.883	0.870
3. backward	8	0.723	0.945	0.882	0.875

The model fit figures AUC and accuracy are considerably lower than the figures reached with the mixed models. On the other hand, the models without a random effect generalize well to new data: the prediction accuracy in 10-fold cross-validation is very similar to the model fit accuracy (training and testing on all instances). The prediction accuracies reached in 10-fold cross-validation are significantly better than those reached with the best mixed model ($p < 0.001$ for the three regular models compared to the forward/backward mixed model following the Wilcoxon paired signed rank test). Apparently the simpler models (those without a random effect) outperform the mixed models when applying them to previously unseen data.

Table 7 shows the significant effects in the models without random effect. Again, the directions of the coefficients are as expected, but the three models disagree on the significance of the variables. Four variables have significant effects in two of the three models, one (the definiteness of the theme) only has an effect in the stepwise forward selection model. Only the concreteness of the theme is selected in all three approaches, as opposed to the mixed-effect approach of the previous section, where it was not selected at all. According to all three models, speakers tend to use the double object construction when the theme is longer than the recipient, and when the recipient is pronominal. The backward selection model (3), however, shows that the effect of length difference is especially strong in speech, while the effect of the pronominality of the recipient is particularly strong in writing. As in the previous section, where the one significant interaction (medium with pronominality of theme) was only found in model 1, it is not clear whether this difference really exists.

Table 7: Coefficients of significant effects in regression models (without random effect), trained on all 915 instances, *** p<0.001 ** p<0.01 * p<0.05. The last column indicates the direction towards the prepositional dative (p.dat.) and double object construction (d.obj.).

Effect	1. significant		2. forward		3. backward		
length difference, medium=spoken					-2.29	***	
rec=pronominal, medium=written					-2.07	***	d.obj.
length difference	-1.75	***	-1.97	***			↑
rec=pronominal, medium=spoken					-1.71	***	
length difference, medium=written					-1.49	***	
rec=definite			-1.18	***	-1.20	***	
rec=local	-1.13	***	-1.20	***			
rec=pronominal	-1.15	***	-1.20	***			
th=definite			1.12	***			
(intercept)					1.37	***	↓
th=given			0.95	**	1.44	***	p.dat.
th=concrete	1.50	***	1.52	***	1.27	***	

Surprisingly enough, excluding the verb sense as a random effect has not resulted in a significant effect for semantic verb class in any of the models. Given the high model fit we found for the models with verb sense as a random effect, and the predictive quality of verb sense found in previous research (Gries & Stefanowitsch 2004), one would expect that having information about the semantic verb class would be useful in the models without this random effect. Apparently, the effect is not strong enough.

6. Conclusion

In this paper, we built regular and mixed (i.e. containing a random effect) logistic regression models in order to explain the British English dative alternation. We used a data set of 915 instances taken from the ICE-GB Corpus, and took the explanatory factors suggested by Bresnan et al. (2007). The regular and the mixed models were constructed following three different approaches: (1) providing the algorithms with all 29 variables and keeping the significant ones, (2) starting with an empty model and forwardly sequentially adding the most predictive variables, and (3) starting with a model with all 29 features and backwardly sequentially removing the least predictive variables. In total, we thus have built six logistic regression models for the same data set.

The six models show some overlap in the variables that are regarded significant. These variables show the same effects as found for American English (Bresnan et al. 2007): pronominal, relatively short, local, discourse given, definite and concrete objects typically precede objects with the opposite characteristics. Contrary to the observations in Haspelmath (2007), we have no reason to believe that the dative alternation in British English differs from that in American English. We have found no clear indications of differences between the dative alternation in speech and writing either: only three variables were selected in interaction with medium, and they occurred in only one model.

As opposed to the mixed models, the models without a random effect generalize well to previously unseen data. This does not necessarily mean that the British English dative alternation is best modelled with logistic regression models without a random effect. The models fit the data better when verb sense is included as a random effect. The fact that the mixed models do not generalize well to new data could be an artefact of lack of data instances. In the near future, we therefore aim at extending our data set, employing the British National Corpus (Burnard 2007). Since manually extending the data set in a way similar to that taken to reach the current data set of 915 instances is too labour-intensive, we aim at automatically extending the data set (in an approach similar to that taken in Lapata (1999)), and automatically annotating it for the explanatory features in this paper. With the larger set, we hope to be able to model the underlying processes of the dative alternation, rather than modelling the instances that made it into our data set.

One of the drawbacks of variable selection is that different methods can lead to different models (Izenman 2008). Unsurprisingly, the six approaches we took have led to five different models. How can we decide which is the optimal model for our purpose? Of course, the approach depends on your goal. For a researcher building a machine translation system, the goal will probably be to reach the highest prediction accuracy on previously unseen data. For linguists, however, the goal is more complex. We want to combine the explanatory features suggested in previous research and test the combination on real data. We thus have hypotheses about what are the explanatory features and what kind of effect they show, but it is unclear how they behave in combination with the others. Also, we want a model that is interpretable and, ideally, reflects the processes in our brains. It is uncertain how (and if) we can evaluate a model in this sense. Still, despite these difficulties, using techniques such as logistic regression is very useful for gaining insight in the statistical characteristics that play a role in syntactic variability. But contrary to what is common in linguistics, researchers should be careful in choosing a single approach and drawing conclusions from one model only.

References

D. Bates (2005). 'Fitting linear mixed models in R'. *R News* **5**(1):27–30.

A. A. Blackwell (2005). 'Acquiring the English adjective lexicon: relationships with input properties and adjectival semantic typology'. *Child Language* (32):535–562.

J. Bresnan, et al. (2007). 'Predicting the Dative Alternation'. In G. Bouma, I. Kraemer, & J. Zwarts (eds.), *Cognitive Foundations of Interpretation*, pp. 69–94. Royal Netherlands Academy of Science, Amsterdam, the Netherlands.

L. Burnard (ed.) (2007). *Reference Guide for the British National Corpus (XML Edition)*. Research Technologies Service at Oxford University Computing Services, Published for the British National Corpus Consortium.

J. J. Godfrey, et al. (1992). 'Switchboard: Telephone speech corpus for research and development'. In *Proceedings of ICASSP-92*, pp. 517–520, San Fransisco, U.S.A.

S. Greenbaum (ed.) (1996). *Comparing English Worldwide: The International Corpus of English*. Clarendon Press, Oxford, U.K.

S. T. Gries (2003). 'Towards a corpus-based identification of prototypical instances of constructions'. *Annual Review of Cognitive Linguistics* (1):1–27.

S. T. Gries & A. Stefanowitsch (2004). 'Extending Collostructional Analysis: A Corpus-based Perspective on 'Alternations''. *International Journal of Corpus Linguistics* (9):97–129.

S. Grondelaers & D. Speelman (2007). 'A variationist account of constituent ordering in presentative sentences in Belgian Dutch'. *Corpus Linguistics and Linguistic Theory* **3**(2):161–193.

M. Haspelmath (2007). 'Ditransitive alignment splits and inverse alignment'. *Functions of Language* **14**(1):79–102.

A. J. Izenman (2008). *Modern Multivariate Statistical Techniques Regression, Classification, and Manifold Learning*. Springer, New York, USA.

M. Lapata (1999). 'Acquiring lexical generalizations from corpora: a case study for diathesis alternations'. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics (ACL)*, pp. 397–404, Morristown, USA.

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

T. Rietveld & R. van Hout (2008). *Statistical Techniques for the Study of Language and Language Behavior*. Mouton de Gruyter, Berlin, Germany.

S. J. Sheather (2009). *A Modern Approach to Regression with R*. Springer, New York, USA.

BASE BELIEF REVISION FOR FINITARY MONOTONIC LOGICS

Pere Pardo
Institut d'Investigació en Intel.ligència Artificial

Abstract. We slightly improve on characterization results already in the literature for base revision. We show that in order to axiomatically characterize revision operators in a logic the only conditions this logic is required to satisfy are: finitarity and monotony. A characterization of limiting cases of revision operators, full meet and maxichoice, is also offered. Finally, we distinguish two types of bases naturally arising in the context of graded fuzzy logics.

1. Introduction

This paper is about (multiple) base belief change, in particular our results are mainly about base revision, which is characterized for a broad class of logics. The original framework of (Alchourrón, et al. 1985) deals with belief change operators on deductively closed theories. This framework was generalized by Hansson (Hansson 1993, Hansson 1999) to deal with *bases*, i.e. arbitrary set of formulas, the original requirement of logical closure being dropped. Hansson characterized revision and contraction operators in, essentially, monotonic compact logics with the deduction theorem property. These results were improved in (Hansson & Wasserman 2002): while for contraction ((Hansson & Wasserman 2002, Theorem 3.8)) it is shown that finitarity and monotony of the underlying logic suffice, for revision ((Hansson & Wasserman 2002, Theorem 3.17)) their proof depends on a further condition, *Non-contravention*: for all sentences φ, if $\neg\varphi \in \mathrm{Cn}_{\mathcal{S}}(T \cup \{\varphi\})$, then $\neg\varphi \in \mathrm{Cn}_{\mathcal{S}}(T)$.

In this paper we provide a further improvement of Hansson and Wassermann's results by proving a characterization theorem for base revision in any finitary monotonic logic. Namely, taking inspiration from (Booth & Richter 2005), we show that *Non-contravention* can be dropped in the characterization of base revision if we replace the notion of unprovability (of some formula φ) in remainder sets by the stronger notion of consistency (with $\neg\varphi$) in the definition of selection functions and partial meet. This is the main contribution of the paper, together with its extension to the characterization of the revision operators corresponding to limiting cases of selection functions, i.e. full meet and maxichoice revision operators.

In the second part of the paper, as a particular class of finitary monotonic logics, we focus on graded fuzzy logics. We introduce there a distinction in basehood and observe some differences in the behavior of the corresponding base revision operators.

This paper is structured as follows. First we introduce in Section 2. the necessary background material on logic, with particular attention to fuzzy and graded logics, and on partial meet base belief change. Then in Section 3. we set out the main characterization results for base revision, including full meet and maxichoice revision operators. Finally in Section 4. a natural distinction upon bases in graded logics is considered (wether or not they are taken to be closed under truth-degrees).

1.0.1. Related Work.

The recent literature belief change contains several approaches to extend the classical works on (partial meet) revision and contraction (Alchourrón et al. 1985), (Hansson 1999). We mentioned a paper about revision in signed fuzzy logics (Booth & Richter 2005), in the framework of (Gerla 2001); fuzzy logics considered there are defined on top of some underlying logic by replacing its formulas φ with truth-constant labeled formulas (φ, r). Graded logics considered here, in contrast, take truth-constants to be propositional variables (with restricted semantics). Different methods for theory change, other than partial meet, have also been studied: entrenchment (Gärdenfors & Makinson 1998) and systems of spheres (Grove 1988)). Iterated belief change is another topic of research: see (Darwiche & Pearl 1997) and Spohn (Spohn 1988). A different change operation is that of update, where belief change is due to a change in the world; see (Katsuno & Mendelzon 1991) for characterization results. In addition, some partial but promising results have been obtained from dynamic approaches attempting to define change operators *within* dynamic logic (as opposed to the usual: *on top of* some logic); for an overview see (van Ditmarsch, et al. 2008) or (**?**). Finally, for fuzzy logics not based upon Łukasiewicz, we may cite the possibilistic approach to belief change of (Dubois & Prade 1997).

2. Preliminaries on theory and base belief change

We introduce in this section the concepts and results needed later. Following (Font & Jansana 1996), we define a logic as a finitary and structural consequence relation $\vdash_{\mathcal{S}} \subseteq \mathcal{P}(\mathbf{Fm}) \times \mathbf{Fm}$, for some algebra of formulas $\mathbf{Fm}$[1].

Belief change is the study of how some theory T (non-necessarily closed, as we use the term) in a given language L (containing $\rightarrow$ and $\overline{0}$) can adapt to new incoming information $\varphi \in L$ (inconsistent with T, in the interesting case). The main operations are: *revision*, where the new input must follow from the revised theory, which is to be consistent, and *contraction* where the input must not follow from the contracted theory. In the classical paper (Alchourrón et al. 1985), by Alchourrón, Gärdenfors and Makinson, partial meet revision and contraction operations were characterized for closed theories in, essentially, monotonic compact logics with the deduction property[2]. Their work put in solid grounds this newly established area of research, opening the way for other formal studies involving new objects of change, operations (see (Peppas 2007) for a comprehensive list) or logics. We follow (Alchourrón et al. 1985) and define change operators by using partial meet: *Partial meet* consists in (i) generating all logically maximal ways to adapt T to the new sentence (those subtheories of T making further information loss logically unnecessary), (ii) selecting some of these possibilities, (iii) forming their meet, and, optionally, (iv) performing additional steps (if required by the operation). Then a set of axioms is provided to capture these partial meet operators, by showing equivalence between satisfaction of these axioms and being a partial meet operator[3].In addition, new axioms may be intro-

[1]That is, $\mathcal{S}$ satisfies (1) If $\varphi \in \Gamma$ then $\Gamma \vdash_{\mathcal{S}} \varphi$, (2) If $\Gamma \vdash_{\mathcal{S}} \varphi$ and $\Gamma \subseteq \Delta$ then $\Delta \vdash_{\mathcal{S}} \varphi$, (3) If $\Gamma \vdash_{\mathcal{S}} \varphi$ and for every $\psi \in \Gamma$, $\Delta \vdash_{\mathcal{S}} \psi$ then $\Delta \vdash_{\mathcal{S}} \varphi$ (*consequence relation*); (4) If $\Gamma \mathcal{S} \varphi$ then for some finite $\Gamma_0 \subseteq \Gamma$ we have $\Gamma_0 \vdash_{\mathcal{S}} \varphi$ (*finitarity*); (5) If $\Gamma \vdash_{\mathcal{S}} \varphi$ then $e[\Gamma] \vdash_{\mathcal{S}} e(\varphi)$ for all substitutions $e \in Hom(\mathbf{Fm}, \mathbf{Fm})$ (*structurality*). We will use throughout the paper relational $\vdash_{\mathcal{S}}$ and functional $\mathrm{Cn}_{\mathcal{S}}$ notation indistinctively, where $\mathrm{Cn}_{\mathcal{S}}$ is the consequence operator induced by $\mathcal{S}$.

[2]That is, logics satisfying the Deduction Theorem: $\Gamma \cup \{\varphi\} \vdash_{\mathcal{S}} \psi$ iff $\Gamma \vdash_{\mathcal{S}} \varphi \rightarrow \psi$.

[3]Other known formal mechanisms defining change operators can be classified into two broad classes:

duced to characterize the limiting cases of selection in step (ii), full meet and maxichoice selection types. Finally, results showing the different operation types can be defined each other are usually provided too.

A *base* is an arbitrary set of formulas, the original requirement of logical closure being dropped. Base belief change, for the same logical framework than AGM, was characterized by Hansson (see (Hansson 1993), (Hansson 1999)). The results for contraction and revision were improved in (Hansson & Wasserman 2002) (by Hansson and Wassermann): for contraction ((Hansson & Wasserman 2002, Theorem 3.8)) it is shown that finitarity and monotony suffice, while for revision (Theorem (Hansson & Wasserman 2002, Theorem 3.17)) their proof depends on a further condition, *Non-contravention*: for all sentences φ, if $\neg\varphi \in \mathrm{Cn}_{\mathcal{S}}(T \cup \{\varphi\})$, then $\neg\varphi \in \mathrm{Cn}_{\mathcal{S}}(T)$. Observe this condition holds in logics having (i) the deduction property and (ii) the structural axiom of Contraction[4]. We show *Non-contravention* can be dropped in the characterization of revision if we replace unprovability (remainders) by consistency in the definition of partial meet.

3. Multiple base revision for finitary monotonic logics.

Definition 1. *((Zhang & Foo 2001), (Booth & Richter 2005)) Given some monotonic logic $\vdash_{\mathcal{S}}$ (or simply, $\mathcal{S}$), let T_0, T_1 be theories. We say T_0 is* consistent *if $T_0 \nvdash_{\mathcal{S}} \overline{0}$, and define the set of* subsets *of T_0* maximally consistent with T_1 *as follows: $X \in Con(T_0, T_1)$ iff:*

(i) $X \subseteq T_0$,
(ii) $X \cup T_1$ *is consistent, and*
(iii) *For any X' such that $X \subsetneq X' \subseteq T_0$, we have $X' \cup T_1$ is inconsistent*

Now we prove some properties[5] of $\mathrm{Con}(\cdot, \cdot)$ which will be helpful for the characterization theorems of base belief change operators for arbitrary finitary monotonic logics.

Lemma 2. *Let $\mathcal{S}$ be some finitary monotonic logic and T_0 a theory. For any $X \subseteq T_0$, if $X \cup T_1$ is consistent, then X can be extended to some Y with $Y \in \mathrm{Con}(T_0, T_1)$.*

Proof. Let $X \subseteq T_0$ with $X \cup T_1 \nvdash_{\mathcal{S}} \overline{0}$. Consider the poset $(T^*, \subseteq)$, where $T^* = \{Y \subseteq T_0 : X \subseteq Y$ and $Y \cup T_1 \nvdash_{\mathcal{S}} \overline{0}\}$. Let $\{Y_i\}_{i \in I}$ be a chain in $(T^*, \subseteq)$; that is, each Y_i is a subset of T_0 and consistent with T_1. Hence, $\bigcup_{i \in I} Y_i \subseteq T_0$; since $\mathcal{S}$ is finitary, $\bigcup_{i \in I} Y_i$ is also consistent with T_1 and hence is an upper bound for the chain. Applying Zorn's Lemma, we obtain an element Z in the poset with the next properties: $X \subseteq Z \subseteq T$ and Z maximal w.r.t. $Z \cup \{\varphi\} \nvdash_{\mathcal{S}} \overline{0}$. Thus $X \subseteq Z \in \mathrm{Con}(T, \varphi)$. □

Remark 3. *Considering $X = \emptyset$ in the preceding lemma, we infer: if T_1 is consistent, then $\mathrm{Con}(T_0, T_1) \neq \emptyset$.*

---

selection-based mechanisms include selection functions on remainder sets and incision functions on kernels; *ranking*-based mechanisms include entrenchments and systems of spheres. For the logical framework assumed in the original developments (compact -and monotonic- closure operators satisfying the deduction property), all these methods are equivalent (see (Peppas 2007) for a comparison). These equivalences between methods need not be preserved in more general class of logics.

[4]If $T \cup \{\varphi\} \vdash_{\mathcal{S}} \varphi \to \overline{0}$, then by the deduction property $T \vdash_{\mathcal{S}} \varphi \to (\varphi \to \overline{0})$; i.e. $T \vdash_{\mathcal{S}} (\varphi \& \varphi) \to \overline{0}$. Finally, by modus ponens from the axiom of contraction, we obtain $T \vdash_{\mathcal{S}} \varphi \to \overline{0}$.

[5]Note that $\mathrm{Con}(T_0, T_1)$ cannot be empty, since if input T_1 is consistency, then in the worst case, we will have $\emptyset \subseteq T_0$ to be consistent with T_1.

For simplicity, we assume that the input base T_1 (to revise T_0 by) is consistent[6].

Definition 4. *Let T_0 be a theory. A* selection function *for T_0 is a function*

$$\gamma : \mathcal{P}(\mathcal{P}(\mathbf{Fm})) \setminus \{\emptyset\} \longrightarrow \mathcal{P}(\mathcal{P}(\mathbf{Fm})) \setminus \{\emptyset\}$$

such that for all $T_1 \subseteq \mathbf{Fm}$, $\gamma(\mathrm{Con}(T_0, T_1)) \subseteq \mathrm{Con}(T_0, T_1)$ and $\gamma(\mathrm{Con}(T_0, T_1))$ is non-empty.

Thus, selection functions and revision operators are defined relative to some fixed base T_0. Although, instead of writing $\circledast^{T_0} T_1$ we use the traditional infix notation $T_0 \circledast T_1$ for the operation of revising base T_0 by T_1.

3.1. Base belief revision.

The axioms we propose (inspired by (Booth & Richter 2005)) to characterize (multiple) base revision operators for finitary monotonic logics $\mathcal{S}$ are the following, for arbitrary sets T_0, T_1:

(F1)	$T_1 \subseteq T_0 \circledast T_1$	(Success)
(F2)	If T_1 is consistent, then $T_0 \circledast T_1$ is also consistent.	(Consistency)
(F3)	$T_0 \circledast T_1 \subseteq T_0 \cup T_1$	(Inclusion)
(F4)	For all $\psi \in \mathbf{Fm}$, if $\psi \in T_0 - T_0 \circledast T_1$ then, there exists T' with $T_0 \circledast T_1 \subseteq T' \subseteq T_0 \cup T_1$ and such that $T' \nvdash_{\mathcal{S}} \overline{0}$ but $T' \cup \{\psi\} \vdash_{\mathcal{S}} \overline{0})$	(Relevance)
(F5)	If for all $T' \subseteq T_0$ $(T' \cup T_1 \nvdash_{\mathcal{S}} \overline{0} \Leftrightarrow T' \cup T_2 \nvdash_{\mathcal{S}} \overline{0})$ then $T_0 \cap (T_0 \circledast T_1) = T_0 \cap (T_0 \circledast T_2)$	(Uniformity)

These axioms express natural conditions on the operation of revision; the case of axiom (F4), which may be less obvious, gives a necessary condition -inconsistency with some candidate revision- to withdraw some sentence of T_0 during revision.

Given some theory $T_0 \subseteq \mathbf{Fm}$ and selection function γ for T_0, we define partial meet revision operator $\circledast_\gamma$ for T_0 by $T_1 \subseteq \mathbf{Fm}$ as follows:

$$T_0 \circledast_\gamma T_1 \; = \; \bigcap \gamma(\mathrm{Con}(T_0, T_1)) \cup T_1$$

Definition 5. *Let $\mathcal{S}$ be some finitary logic, and T_0 a theory. Then $\circledast : \mathcal{P}(\mathbf{Fm}) \to \mathcal{P}(\mathbf{Fm})$ is a* revision operator *for T_0 iff for any $T_1 \subseteq \mathbf{Fm}$, $T_0 \circledast T_1 = T_0 \circledast_\gamma T_1$ for some selection function γ for T_0.*

Lemma 6. *Condition* $\mathrm{Con}(T_0, T_1) = \mathrm{Con}(T_0, T_2)$ *is equivalent to the antecedent of Axiom* (F5)

$$\forall T' \subseteq T_0 \; (T' \cup T_1 \nvdash_{\mathcal{S}} \overline{0} \Leftrightarrow T' \cup T_2 \nvdash_{\mathcal{S}} \overline{0})$$

Proof. (If-then) Assume $\mathrm{Con}(T_0, T_1) = \mathrm{Con}(T_0, T_2)$ and let $T' \subseteq T_0$ with $T' \cup T_1 \nvdash_{\mathcal{S}} \overline{0}$. By Lemma 2, T' can be extended to $X \in \mathrm{Con}(T_0, T_1)$. Hence, by assumption we get $T' \subseteq X \in \mathrm{Con}(T_0, T_2)$ so that $T' \cup T_2 \nvdash_{\mathcal{S}} \overline{0}$ follows. The other direction is similar. (Only if) This direction follows from the definition of $\mathrm{Con}(T_0, \cdot)$. □

[6]Observe one could define for T_1 inconsistent: $\mathrm{Con}(T_0, T_1) = \mathrm{Con}(T_0, \{\overline{1}\})$, so in case T_0 was consistent this definition would make $\mathrm{Con}(T_0, T_1) = \{T_0\}$, and otherwise it would consist of all consistent subtheories of T_0.

Theorem 7. *Let $\mathcal{S}$ be a finitary monotonic logic. For any $T_0 \subseteq \mathbf{Fm}$ and function $\circledast : \mathcal{P}(\mathbf{Fm}) \to \mathcal{P}(\mathbf{Fm})$:*

$$\circledast \textit{ satisfies } (\mathrm{F1}) - (\mathrm{F5}) \text{ iff } \circledast \textit{ is a revision operator for } T_0$$

Proof. (Soundness) Given some partial meet revision operator $\circledast_\gamma$ for T_0, we prove $\circledast_\gamma$ satisfies (F1) $-$ (F5).

(F1) $-$ (F3) hold by definition of $\circledast_\gamma$. (F4) Let $\psi \in T_0 - T_0 \circledast_\gamma T_1$. Hence, $\psi \notin T_1$ and for some $X \in \gamma(\mathrm{Con}(T_0, T_1))$, $\psi \notin X$. Simply put $T' = X \cup T_1$: by definitions of $\circledast_\gamma$ and Con we have (i) $T_0 \circledast_\gamma T_1 \subseteq T' \subseteq T_0 \cup T_1$ and (ii) T' is consistent (since T_1 is). We also have (iii) $T' \cup \{\psi\}$ is inconsistent (otherwise $\psi \in X$ would follow from maximality of X and $\psi \in T_0$, hence contradicting our previous step $\psi \notin X$). (F5) We have to show, assuming the antecedent of(F5), that $T_0 \cap (T_0 \circledast_\gamma T_1) = T_0 \cap (T_0 \circledast_\gamma T_2)$. We prove the $\subseteq$ direction only since the other is similar. Assume, then, for all $T' \subseteq T_0$,

$$T' \cup T_1 \nvdash_{\mathcal{S}} \overline{0} \;\Leftrightarrow\; T' \cup T_2 \nvdash_{\mathcal{S}} \overline{0}$$

and let $\psi \in T_0 \cap (T_0 \circledast_\gamma T_1)$. This set is just $T_0 \cap (\bigcap \gamma(Con(T_0, T_1)) \cup T_1)$ which can be transformed into $(T_0 \cap \bigcap \gamma(Con(T_0, T_1)) \cup (T_0 \cup T_1)$, i.e. $\bigcap \gamma(Con(T_0, T_1)) \cup (T_0 \cup T_1)$ (since $\bigcap \gamma(Con(T_0, T_1)) \subseteq T_0$). Case $\psi \in \bigcap \gamma(Con(T_0, T_1))$. Then we use Lemma 6 upon the assumption to obtain $\bigcap \gamma(Con(T_0, T_1)) = \bigcap \gamma(Con(T_0, T_2))$, since γ is a function. Case $\psi \in T_0 \cap T_1$. Then $\psi \in X$ for all $X \in \gamma(Con(T_0, T_1))$, by maximality of X. Hence, $\psi \in \bigcap \gamma(Con(T_0, T_1))$. Using the same argument than in the former case, $\psi \in \bigcap \gamma(Con(T_0, T_2))$. Since we also assumed $\psi \in T_0$, we obtain $\psi \in T_0 \cap (T_0 \circledast_\gamma T_2)$.
(Completeness) Let $\circledast$ satisfy (F1) $-$ (F5). We have to show that for some selection function γ and any T_1, $T_0 \circledast T_1 = T \circledast_\gamma T_1$. We define first

$$\gamma(\mathrm{Con}(T_0, T_1)) = \{X \in \mathrm{Con}(T_0, T_1) : X \supseteq T_0 \cap (T_0 \circledast T_1)\}$$

We prove that (1) γ is well-defined, (2) γ is a selection function and (3) $T_0 \circledast T_1 = T \circledast_\gamma T_1$.

(1) Assume (i) $\mathrm{Con}(T_0, T_1) = \mathrm{Con}(T_0, T_2)$; we have to prove that $\gamma(\mathrm{Con}(T_0, T_1)) = \gamma(\mathrm{Con}(T_0, T_2))$. Applying Lemma 6 to (i) we obtain the antecedent of (F5). Since $\circledast$ satisfies this axiom, we have (ii) $T_0 \cap (T_0 \circledast T_1) = T_0 \cap (T_0 \circledast T_2)$. By the above definition of γ, $\gamma(\mathrm{Con}(T_0, T_1)) = \gamma(\mathrm{Con}(T_0, T_2))$ follows from (i) and (ii).

(2) Since T_1 is consistent, by Remark 3 we obtain $\mathrm{Con}(T_0, T_1)$ is not empty; we have to show that $\gamma(\mathrm{Con}(T_0, T_1))$ is not empty either (since the other condition $\gamma(\mathrm{Con}(T_0, T_1)) \subseteq \mathrm{Con}(T_0, T_1)$ is met by the above definition of γ). We have $T_0 \cap T_0 \circledast T_1 \subseteq T_0 \circledast T_1$; the latter is consistent and contains T_1, by (F2) and (F1), respectively; thus, $(T_0 \cap T_0 \circledast T_1) \cup T_1$ is consistent; from this and $T_0 \cap T_0 \circledast T_1 \subseteq T_0$, we deduce by Lemma 2 that $T_0 \cap T_0 \circledast T_1$ is extensible to some $X \in \mathrm{Con}(T_0, T_1)$. Thus, exists some $X \in \mathrm{Con}(T_0, T_1)$ such that $X \supseteq T_0 \cap T_0 \circledast T_1$. In consequence, $X \in \gamma(\mathrm{Con}(T_0, T_1)) \neq \emptyset$.

For (3), we prove first $T_0 \circledast T_1 \subseteq T_0 \circledast_\gamma T_1$. Let $\psi \in T_0 \circledast T_1$. By (F3), $\psi \in T_0 \cup T_1$. Case $\psi \in T_1$: then trivially $\psi \in T_0 \circledast_\gamma T_1$ Case $\psi \in T_0$. Then $\psi \in T_0 \cap T_0 \circledast T_1$. In consequence, for any $X \in \mathrm{Con}(T_0, T_1)$, if $X \supseteq T_0 \cap T_0 \circledast T_1$ then $\psi \in X$. This implies, by definition of γ above, that for all $X \in \gamma(\mathrm{Con}(T_0, T_1))$ we have $\psi \in X$, so that $\psi \in \bigcap \gamma(\mathrm{Con}(T_0, T_1)) \subseteq T_0 \circledast_\gamma T_1$. In both cases, we obtain $\psi \in T_0 \circledast_\gamma T_1$.

Now, we prove the other direction: $T_0 \circledast_\gamma T_1 \subseteq T_0 \circledast T_1$. Let $\psi \in \bigcap \gamma(\mathrm{Con}(T_0, T_1)) \cup T_1$. By (F1), we have $T_1 \in T_0 \circledast T_1$; then, in case $\psi \in T_1$ we are done. So we may assume

$\psi \in \bigcap \gamma(\mathrm{Con}(T_0, T_1))$. Now, in order to apply (F4), let X be arbitrary with $T \circledast T_1 \subseteq X \subseteq T_0 \cup T_1$ and X consistent. Consider $X \cap T_0$: since $T_1 \subseteq T_0 \circledast T_1 \subseteq X$ implies $X = X \cup T_1$ is consistent, so is $(X \cap T_0) \cup T_1$. Together with $X \cap T_0 \subseteq T_0$, by Lemma 2 there is $Y \in \mathrm{Con}(T_0, T_1)$ with $X \cap T_0 \subseteq Y$. In addition, since $T_0 \circledast T_1 \subseteq X$ implies $T_0 \circledast T_1 \cap T_0 \subseteq X \cap T_0 \subseteq Y$ we obtain $Y \in \gamma(\mathrm{Con}(T_0, T_1))$, by the definition of γ above. Condition $X \cap T_0 \subseteq Y$ also implies $(X \cap T_0) \cup T_1 \subseteq Y \cup T_1$. Observe that from $X \subseteq X \cup T_1$ and $X \subseteq T_0 \cup T_1$ we infer that $X \subseteq (X \cup T_1) \cap (T_0 \cup T_1)$. From the latter being identical to $(X \cap T_0) \cup T_1$ and the fact that $(X \cap T_0) \cup T_1 \subseteq Y \cup T_1$, we obtain that $X \subseteq Y \cup T_1$. Since $\psi \in Y \in \mathrm{Con}(T_0, T_1)$, we have $Y \cup T_1$ is consistent with ψ, so its subset X is also consistent with ψ. Finally, we may apply *modus tollens* on Axiom (F4) to obtain that $\psi \notin T_0 - T_0 \circledast T_1$, i.e. $\psi \notin T_0$ or $\psi \in T_0 \circledast T_1$. But since the former is false, the latter must be the case. $\square$

3.1.1. Full meet and maxichoice base revision operators.

The previous result can be extended to limiting cases of selection functions formally defined as follows:

Definition 8. *A revision operator for T_0 is* full meet *if it is generated by the identity selection function $\gamma_{\mathrm{fm}} = \mathsf{Id}$: $\gamma_{\mathrm{fm}}(\mathrm{Con}(T_0, T_1)) = \mathrm{Con}(T_0, T_1)$; that is,*

$$T_0 \circledast_{\mathrm{fm}} T_1 = (\bigcap \mathrm{Con}(T_0, T_1)) \cup T_1$$

A revision operator for T_0 is maxichoice *if it is generated by a selection function of type $\gamma_{\mathrm{mc}}(\mathrm{Con}(T_0, T_1)) = \{X\}$, for some $X \in \mathrm{Con}(T_0, T_1)$, and in that case $T_0 \circledast_{\gamma_{\mathrm{mc}}} T_1 = X \cup T_1$.*

To characterize *full meet* and *maxichoice* revision operators for some theory T_0 in any finitary logic, we define the next additional axioms:

$$\begin{array}{ll} \text{(FM)} & \text{For any } X \subseteq \mathbf{Fm} \text{ with } T_1 \subseteq X \subseteq T_0 \cup T_1 \\ & \quad X \nvdash_{\mathcal{S}} \overline{0} \text{ implies } X \cup (T_0 \circledast T_1) \nvdash_{\mathcal{S}} \overline{0} \\ \text{(MC)} & \text{For all } \psi \in \mathbf{Fm} \text{ with } \psi \in T_0 - T_0 \circledast T_1 \text{ we have} \\ & \quad T_0 \circledast T_1 \cup \{\psi\} \vdash_{\mathcal{S}} \overline{0} \end{array}$$

Theorem 9. *Let $T_0 \subseteq \mathbf{Fm}$ and $\circledast$ be a function $\circledast : \mathcal{P}(\mathbf{Fm})^2 \to \mathcal{P}(\mathbf{Fm})$. Then the following hold:*

$$\begin{array}{llll} \text{(fm)} & \circledast \textit{ satisfies } \text{(F1)} - \text{(F5)} \textit{ and } \text{(FM)} & \text{iff} & \circledast = \circledast_{\gamma_{\mathrm{fm}}} \\ \text{(mc)} & \circledast \textit{ satisfies } \text{(F1)} - \text{(F5)} \textit{ and } \text{(MC)} & \text{iff} & \circledast = \circledast_{\gamma_{\mathrm{mc}}} \end{array}$$

Proof. We prove (fm) first. (Soundness): We know $\circledast_{\gamma_{\mathrm{fm}}}$ satisfies (F1) − (F5) so it remains to be proved that (FM) holds. Let X be such that $T_1 \subseteq X \subseteq T_0 \cup T_1$ and $X \nvdash_{\mathcal{S}} \overline{0}$. From the latter and $X - T_1 \subseteq (T_0 \cup T_1) - T_1 \subseteq T_0$ we infer by Lemma 2 that $X - T_1 \subseteq Y \in \mathrm{Con}(T_0, T_1)$, for some Y. Notice $X = X' \cup T_1$ and that for any $X'' \in \mathrm{Con}(T_0, T_1) X'' \cup T_1$ is consistent and

$$T_0 \circledast_{\gamma_{\mathrm{fm}}} T_1 = (\bigcap \mathrm{Con}(T_0, T_1)) \cup T_1 \subseteq X' \subseteq X''$$

Hence $X \subseteq X''$, so that $T_0 \circledast_{\gamma_{\mathrm{fm}}} T_1 \cup X \subseteq X''$. Since the latter is consistent, $T_0 \circledast_{\mathrm{fm}} T_1 \cup X \nvdash_{\mathcal{S}} \overline{0}$. (Completeness) Let $\circledast$ satisfy (F1) − (F5) and (FM). It suffices to prove

that $X \in \gamma(\mathrm{Con}(T_0, T_1)) \Leftrightarrow X \in \mathrm{Con}(T_0, T_1)$; but we already know that $\circledast = \circledast_\gamma$, for selection function γ (for T_0) defined by: $X \in \gamma(\mathrm{Con}(T_0, T_1)) \Leftrightarrow T_0 \cap T_0 \circledast T_1 \subseteq X$. It is enough to prove, then, that $X \in \mathrm{Con}(T_0, T_1)$ implies $X \supseteq T_0 \cap T_0 \circledast T_1$. Let $X \in \mathrm{Con}(T_0, T_1)$ and let $\psi \in T_0 \cap T_0 \circledast T_1$. Since $\psi \in T_0$ and $X \in \mathrm{Con}(T_0, T_1)$, we have by maximality of X that either $X \cup \{\psi\} \vdash_{\mathcal{S}} \overline{0}$ or $\psi \in X$. We prove the former case to be impossible: assuming it we would have $T_1 \subseteq X \cup T_1 \subseteq T_0 \cup T_1$. By (FM), $X \cup T_1 \cup (T_0 \circledast T_1) \nvdash_{\mathcal{S}} \overline{0}$. Since $\psi \in T_0 \circledast T_1$, we would obtain $X \cup \{\psi\} \nvdash_{\mathcal{S}} \overline{0}$, hence contradicting the case assumption; since the former case is not possible, we have $\psi \in X$. Since X was arbitrary, $X \in \mathrm{Con}(T_0, T_1)$ implies $X \subseteq T_0 \cap T_0 \circledast T_1$ and we are done.
For $\underline{\mathrm{mc}}$): (Soundness) We prove (MC), since (F1) – (F5) follow from $\circledast_{\gamma_{\mathrm{mc}}}$ being a partial meet revision operator. Let $X \in \mathrm{Con}(T_0, T_1)$ be such that $T_0 \circledast_{\gamma_{\mathrm{mc}}} \varphi = X \cup T_1$ and let $\psi \in T_0 - T_0 \circledast_{\gamma_{\mathrm{mc}}} T_1$. We have $\psi \notin X \cup T_1 = T_0 \circledast T_1$. Since $\psi \in T_0$ and $X \in \mathrm{Con}(T_0, T_1)$, $X \cup \{\psi\} \vdash_{\mathcal{S}} \overline{0}$. Finally $T_0 \circledast T_1 \cup \{\psi\} \vdash_{\mathcal{S}} \overline{0}$. (Completeness) Let $\circledast$ satisfy (F1) – (F5) and (MC). We must prove $\circledast = \circledast_{\gamma_{\mathrm{mc}}}$, for some maxichoice selection function γ_{mc}. Let $X, Y \in \mathrm{Con}(T_0, T_1)$; we have to prove $X = Y$. In search of a contradiction, assume the contrary, i.e. $\psi \in X - Y$. We have $\psi \notin \bigcap \gamma(\mathrm{Con}(T_0, T_1))$ and $\psi \in X \subseteq T_0$. By MC, $T_0 \circledast T_1 \cup \{\psi\} \vdash_{\mathcal{S}} \overline{0}$. Since $T_0 \circledast T_1 \subseteq X$, we obtain $X \cup \{\psi\}$ is also inconsistent, contradicting previous $\psi \in X \nvdash_{\mathcal{S}} \overline{0}$. Thus $X = Y$ which makes $\circledast = \circledast_{\gamma_{\mathrm{mc}}}$, for some maxichoice selection function γ_{mc}. □

4. The case of graded fuzzy logics.

The characterization results for base revision operators from the previous section required weak assumptions (monotony and finitarity) upon the consequence relation $\vdash_{\mathcal{S}}$. In particular these results hold for a wide family of systems of (mathematical) fuzzy logic. The distinctive feature of these logics is that they cope with graded truth in a compositional manner (see (Hájek 1998, Section 3) for an overview). Graded truth may be dealt implicitly, by means of comparative statements, or explicitly, by introducing truth-degrees in the language . Here we will focus on a particular kind of fuzzy logical languages allowing for explicit representation of truth-degrees, that will be referred as *graded fuzzy logics*, and which are expansions of t-norm logics with countable sets of truth-constants. These logics allow for occurrences of truth-degrees, represented as new propositional atoms $\overline{r}$ (one for each $r \in \mathcal{C}$) in any part of a formula. These truth-constants and propositional variables can be combined arbitrarily using connectives to obtain new formulas. The graded language obtained in this way will be denoted as $\mathbf{Fm}(\mathcal{C})$. These expansions also require additional axioms for truth-constants, e.g. $(\overline{r} \to \overline{s}) \equiv \overline{r \Rightarrow s}$, called book-keeping axioms. A prominent example of a logic over a graded language is Hájek's Rational Pavelka Logic RPL (Hájek 1998), an extension of Łukasiewicz logic with rational truth-constants in $[0, 1]$; for other graded extensions of t-norm based fuzzy logics see e.g. (Esteva, et al. 2007). In t-norm based fuzzy logics, due to the fact that the implication is residuated, a formula $\overline{r} \to \varphi$ gets value 1 under a given interpretation e iff $r \leq e(\varphi)$. In what follows, we will also use the signed language notation (φ, r) to denote the formula $\overline{r} \to \varphi$.

If $\mathcal{S}$ denotes a given t-norm logic, let us denote by $\mathcal{S}(\mathcal{C})$ the corresponding expansion with truth-constants from a suitable countable set $\mathcal{C}$ such that $\{0, 1\} \subset \mathcal{C} \subseteq [0, 1]$. For instance if $\vdash_{\mathcal{S}}$ is Łukasiewicz logic and $\mathcal{C} = \mathbb{Q} \cap [0, 1]$, then $\vdash_{\mathcal{S}(\mathcal{C})}$ would refer to RPL.

For these graded fuzzy logics, besides the original definition of a base as simply a set of formulas, it makes sense to consider another natural notion of basehood, where bases are closed by lower bounds of truth-degrees. We call them $\mathcal{C}$-closed. bases.

Definition 10. *(Adapted from (Hansson 1993)) Given some (monotonic) t-norm fuzzy logic $\mathcal{S}$ with language $\mathbf{Fm}$ and a countable set $\mathcal{C} \subset [0,1]$ of truth-constants, let $T \subseteq \mathbf{Fm}(\mathcal{C})$ be a base in $\mathcal{S}(\mathcal{C})$. We define $\mathrm{Cn}_{\mathcal{C}}(T) = \{(\varphi, r') : (\varphi, r) \in T$, for $r, r' \in \mathcal{C}$ with $r \geq r'\}$. A base $T \subseteq \mathbf{Fm}(\mathcal{C})$ is called $\mathcal{C}$-closed when $T = \mathrm{Cn}_{\mathcal{C}}(T)$.*

Notice that in some sense, using Gerla's framework of abstract fuzzy logic (Gerla 2001), the approach by Booth and Richther (Booth & Richter 2005) defines revision operators for bases which are closed with respect to some complete lattice W of truth-values.

The following results prove $\circledast_\gamma$ operators preserve $\mathcal{C}$-closure, thus making $\mathcal{C}$-closed revision a particular case of base revision under Theorem 7.

Proposition 11. *If T_0, T_1 are $\mathcal{C}$-closed graded bases, for any partial meet revision operator $\circledast_\gamma$, $T_0 \circledast_\gamma T_1$ is also a $\mathcal{C}$-closed graded base.*

Proof. Since T_0 is $\mathcal{C}$-closed, by maximality of $X \in \gamma(\mathrm{Con}(T_0, T_1))$ we have X is also $\mathcal{C}$-closed, for any such X. Let $(\psi, s) \in \bigcap \gamma(\mathrm{Con}(T_0, T_1))$ and $s' <_{\mathcal{C}} s$ for some $s' \in \mathcal{C}$. Then $(\psi, s) \in X$ for any $X \in \gamma(\mathrm{Con}(T_0, T_1))$ implies $(\psi, s') \in X$ for any such X. Hence $\bigcap \gamma(\mathrm{Con}(T_0, T_1))$ is $\mathcal{C}$-closed. Finally, since T_1 is $\mathcal{C}$-closed, we deduce $\bigcap \gamma(\mathrm{Con}(T_0, T_1)) \cup T_1$ is also $\mathcal{C}$-closed. $\Box$

Let $\mathcal{P}_{\mathcal{C}}(\mathbf{Fm})$ be the set of $\mathcal{C}$-closed sets of $\mathbf{Fm}$ sentences. We introduce an additional axiom $(\mathrm{F0})$ for revision of $\mathcal{C}$-closed bases by $\mathcal{C}$-closed inputs:

$$(\mathrm{F0}) \qquad T_0 \circledast T_1 \text{ is } \mathcal{C}\text{-closed, if } T_0, T_1 \text{ are}$$

Corollary 12. *Assume $\mathcal{S}$ and $\mathcal{C}$ are as before and let $\circledast : \mathcal{P}_{\mathcal{C}}(\mathbf{Fm}) \to \mathcal{P}(\mathbf{Fm})$. Then, $\circledast$ satisfies $(\mathrm{F0}) - (\mathrm{F5})$ iff for some selection function γ, $T_0 \circledast T_1 = T_0 \circledast_\gamma T_1$ for every $T_1 \in \mathcal{P}_{\mathcal{C}}(\mathbf{Fm})$.*

For the case of RPL, where the negation operator $\neg$ is interpreted by the negation function on $[0,1]$ defiend as $n(x) = 1 - x$, both approaches (non-$\mathcal{C}$-closed, $\mathcal{C}$-closed) differ in the revision output. Hence, this distinction in basehood has important consequences.

Example 13. *(In RPL) Let $\mathcal{C} = \mathbb{Q} \cap [0,1]$ and define $T_0 = \{(\varphi, 0.5), (\varphi, 0.7)\}$.In each case, there in only a possible selection function, call them γ_0 and γ_1; revision results in:*

$$\begin{array}{rcl} T_0 \circledast_{\gamma_0} (\neg\varphi, 0.4) & = & \{(\varphi, 0.5), (\neg\varphi, 0.4)\}, \textit{ while} \\ \mathrm{Cn}_{\mathcal{C}}(T_0) \circledast_{\gamma_1} \mathrm{Cn}_{\mathcal{C}}(\{(\neg\varphi, 0.4)\}) & = & \mathrm{Cn}_{\mathcal{C}}(\{(\varphi, 0.6), (\neg\varphi, 0.4)\}) \end{array}$$

Observe this distinction in basehood makes sense only for logics whose negation is involutive, i.e. where $\neg\neg\varphi \to \varphi$ is valid; otherwise, the distinction collapses. For Gödel and Product logics (with non-involutive Gödel negation $\neg_G x = 0$ for $x > 0$ and $\neg_G 0 = 1$) we have that any $\mathcal{C}$-closed base revision operator outputs the same result than the deductive closure of its non-$\mathcal{C}$-closed counterpart.

5. Conclusions and future work

We improved Hansson and Wasserman characterization of the revision operator by dropping one of their conditions, implicitly characterizing revision operators for the class logics with the deduction property. Apart from the general theorem, standard results for full meet and maxichoice revision operators are also provided. Then we moved to the field of graded fuzzy logics, in contradistinction to the approach by Booth and Richter in (Booth & Richter 2005); their work inspired us to prove similar results for a more general logical framework, including t-norm based fuzzy logics from Hájek. Finally, we observed the differences between bases if they are assumed to be closed under truth-degrees.

Several problems are left open for future research: mainly, wether the present (consistency-based) results can be used to characterize contraction as well. Presumably, the standpoint adopted in this paper would lead to a definition of contraction with slightly different properties than that proposed by Hansson and Wasserman.

Acknowledgements

The author wishes to thank Pilar Dellunde, Lluís Godo and two anonymous reviewers for helpful comments, and Consolider project CSD2007-022 from the Spanish Ministry of Science and Technology for financial support.

References

C. Alchourrón, et al. (1985). 'On the Logic of Theory Change: Partial Meet Contraction and Revision Functions'. *The Journal of Symbolic Logic* **50**:510–530.

R. Booth & E. Richter (2005). 'On Revising Fuzzy Belief Bases'. *Studia Logica* **80**:29–61.

A. Darwiche & J. Pearl (1997). 'On the logic of iterated belief change'. *Artificial Intelligence* **89**:1–29.

D. Dubois & H. Prade (1997). 'A synthetic view of beleif revision with uncertain inputs in the framework of possibility theory'. *International Journal of Approximate Reasoning* **17, 2-3**:295–324.

F. Esteva, et al. (2007). 'Adding truth-constants to logics of continuous t-norms: Axiomatization and completeness results'. *Fuzzy Sets and Systems* **158**:2597–618.

J. M. Font & R. Jansana (1996). *A General Algebraic Semantics for Sentential Logics*, vol. 7. Springer-Verlag.

P. Gärdenfors & D. C. Makinson (1998). 'Revisions of knowledge systems using epistemic entrenchment'. In *Proceedings of Theoretical Aspects of Reasoning about Knowledge*, pp. 83–95.

G. Gerla (2001). *Fuzzy Logic: Mathematical Tools for Approximate Reasoning*, vol. 11 of *Trends in Logic*. Kluwer Academic Publishers.

A. J. Grove (1988). 'Two modellings for theory change'. *Journal of Philosophical Logic* **17**:157–170.

P. Hájek (1998). *Metamathematics of Fuzzy Logic*, vol. 4 of *Trends in Logic*. Kluwer Academic Publishers.

S. O. Hansson (1993). 'Reversing the Levi Identity'. *Journal of Philosophical Logic* **22**:637–699.

S. O. Hansson (1999). *A Textbook of Belief Dynamics. Theory Change and Database Updating*. Kluwer.

S. O. Hansson & R. Wasserman (2002). 'Local Change'. *Studia Logica* **70**:49–76.

H. Katsuno & A. O. Mendelzon (1991). 'On the difference between updating a knowledge base and revising it'. In *Proc. of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann Publishers.

P. Peppas (2007). 'Belief Revision'. In B. P. van Harmelen F., V. Lifschitz (ed.), *Handbook of Knowledge Representation*. Elsevier.

W. Spohn (1988). 'Ordinal Conditional Functions: A Dynamic Theory of Epistemic States'. In W. L. Harper & B. Skyrms (eds.), *Causation in Decision, Belief Change and Statistics*, vol. II, pp. 105–134. Kluwer.

H. van Ditmarsch, et al. (2008). *Dynamic Epistemic Logic*, vol. 337 of *Synthese Library*. Springer.

D. Zhang & N. Foo (2001). 'Infnitary Belief Revision'. *Journal of Philosophical Logic* **30**:525–570.

Extraction in the Lambek–Grishin calculus

Arno Bastenhof
Utrecht Institute of Linguistics OTS, Utrecht University

Abstract. In an attempt to deal with quantifier scope ambiguities in categorial type logics (CTL), Moortgat (1992) proposed a type constructor that allows abstraction over the syntactic context of an expression. Here, we describe its application to *extraction*. We extend this result to a syntactic analysis of the latter phenomenon for the *Lambek–Grishin calculus*, an instance of CTL that *derives* Moortgat's type constructor by extending the logical vocabulary with the mirror images of the usual connectives for subcategorization. Semantically, we complete the analysis by using a λ–calculus with constants for existential quantification and equality, pulling off a trick similar to Montague's PTQ take on *to be* in order to identify the extracted argument with its associated gap.

Categorial type logics (Moortgat 1997, CTL) meet with serious difficulties when confronted with extraction phenomena: as logics about strings (Lambek 1958, **L**) or trees (Lambek 1961, **NL**), they fail to explain the full range of discontinuous dependencies encountered in natural languages. In order to overcome the expressive limitations of these basic systems, Moortgat (1992) proposed a connective that allows one to abstract within a type over (the type of) the syntactic domain containing its occurrence. Though originally intended for the analysis of quantifier scope ambiguities, we here describe its application to extraction. We consider a proposal by Bernardi & Moortgat (2007) for *decomposing* Moortgat's type constructor in the *Lambek–Grishin calculus* (**LG**), an instance of CTL that conversatively extends upon **NL** by augmenting it with its own mirror–image in the derivability relation. We exploit Bernardi and Moortgat's results in extending our syntactic analysis of extraction to **LG**, and complete the picture by associating it with a compositional semantics.

We proceed as follows. In §1, we briefly review the application of CTL to linguistic analysis. §2 describes our usage of Moortgat's connective in finding a type schema for predicates figuring in extraction environments, its instantiations being meant to coexist with the usual types for cases of *non*–extraction. §3 turns to the decomposition of Moortgat's type constructor in **LG**, showing it to reduce the lexical ambiguity posited in §2. Finally, §4 reviews the *continuation semantics* for **LG** (Bernardi & Moortgat 2007), describing in terms of it the Montagovian semantics associated with our type schema for extraction.

1. Categorial analyses

We adopt a logical perspective on natural language syntax: syntactic categories are logical formulas, or *(syntactic) types* (written $A..E$), as we shall call them. Defined inductively:

$$A, B \quad ::= \quad n \quad | \quad np \quad | \quad s \quad | \quad (B\backslash A) \quad | \quad (A/B)$$

Here, n (the type of common nouns), *np* (noun phrases) and s (sentences) are *atomic* syntactic types. For the interpretations of *complex* types (derived using the *slashes*, or *implications* / and \), we turn to the proof-theoretical meanings of logical connectives (or *type constructors*). In syntax, we understand proofs to derive *sequents* $\Gamma \vdash A$, establishing Γ to be a well-formed binary-branching syntactic tree of category A:

$$\Gamma, \Delta \quad ::= \quad A \quad | \quad (\Gamma \circ \Delta)$$

An expression of type A/B ($B\backslash A$) we then understand to combine directly with an expression of type B to its right (left) into an expression of type A. Made explicit in a Natural Deduction presentation:

$$\frac{}{A \vdash A} Ax$$

$$\frac{\Gamma \vdash A/B \quad \Delta \vdash B}{\Gamma \circ \Delta \vdash A} /E \qquad \frac{\Delta \vdash B \quad \Gamma \vdash B\backslash A}{\Delta \circ \Gamma \vdash A} \backslash E \qquad \frac{\Gamma \circ B \vdash A}{\Gamma \vdash A/B} /I \qquad \frac{B \circ \Gamma \vdash A}{\Gamma \vdash B\backslash A} \backslash I$$

Here, *introduction* rules $(/I), (\backslash I)$ allow the *inference* of a type, whereas *elimination* rules $(/E), (\backslash E)$ allow the *use* of a type in an inference. Axioms (*Ax*) capture the intuition that each structure corresponding to a single leaf is a well–formed structure of the type found at that leaf.

As an example of these definitions, consider the following type assignments to the lexical items in John offered the lady a drink:[1]

John	offered	the	lady	a	drink
np_{su}	$((np_{su}\backslash s)/np_{do})/np_{io}$	np_{io}/n	n	np_{do}/n	n

As it is, the resulting calculus (**NL**, the *Non–Associative Lambek Calculus*) is no stronger than context–free formalisms, making it unsuitable for analyzing such unbounded dependencies as featured in various instances of extraction.

2. Types for extraction

In order to reach beyond the context–free boundaries of standard CTL, Moortgat (1992) introduces type constructors that allow abstraction over the type of the syntactic *context* of an expression. By the latter we understand trees $\Gamma[]$ with a *hole* $[]$. Substituting for $[]$ some Δ yields the tree $\Gamma[\Delta]$, the notation emphasizing the (distinguished) *occurrence* of Δ in Γ. Defined inductively:

$$\Gamma[], \Delta[] \quad ::= \quad [] \quad | \quad \Gamma[] \circ \Delta \quad | \quad \Delta \circ \Gamma[]$$

Any given tree Γ may be (uniquely) rewritten into a form $\Gamma[\Delta]$ for each of its subtrees Δ. From the perspective of type assignment to trees (derivations), nothing exciting happens: if Γ was found to be a well–formed constituent of type B, looking at Γ from the perspective of one of its subtrees Δ isn't going to change this. This observation is captured by the admissibility of the Cut rule:[2]

$$\frac{\Delta \vdash A \quad \Gamma[A] \vdash B}{\Gamma[\Delta] \vdash B} Cut$$

[1] Subscripts are added to occurrences of *np* to facilitate matching with the corresponding grammatical roles: *su*, *do* and *io* abbreviate *subject*, *direct object* and *indirect object* respectively.

[2] In our substitution metaphor, A in $\Gamma[A]$ may be seen as a type-assignment to the hole in $\Gamma[]$ (necessarily coinciding with the type assigned to Δ).

Looking at it from this perspective, we become interested in the possibility of allowing the embedded constituent to manipulate the embedding context. Such is achieved by a ternary type constructor, deriving types $A\left[\begin{smallmatrix} C \\ B \end{smallmatrix}\right]$ with local syntactic distribution characterized by A, but when embedded within an expression of type B resets the latter's type to C:[3]

$$\frac{\Delta \vdash A\left[\begin{smallmatrix} C \\ B \end{smallmatrix}\right] \quad \Gamma[A] \vdash B}{\Gamma[\Delta] \vdash C}\,[\;]E$$

Moortgat (1992) considered as an application the assignment of the type $np\left[\begin{smallmatrix} s \\ s \end{smallmatrix}\right]$ to quantified noun phrases, containing an abstraction over the sentential domain defining their scope (the type of the resulting constituent being again s). Here, we consider its application to *extraction*: for a predicate an argument of which is extracted, we assign the type $A\left[\begin{smallmatrix} B \div \Diamond \Box C \\ B \end{smallmatrix}\right]$ that parameterizes over: its local syntactic distribution A, the type of the extracted argument C (instantiating $B \div \Diamond\Box\, C$ by $B/\Diamond\Box C$ if it occurs in a *right* branch, and $\Diamond\Box C\backslash B$ otherwise), and the type of the extraction domain B. For instance, in the complex noun phrase lady whom John offered ϵ a drink, the object noun phrase ($C = np$) of offered is extracted (from a *right* branch) at the level of the embedded sentence ($B = s$), read itself locally selecting for a subject and direct object ($A = (np\backslash s)/np$):

whom	John	offered	a	drink
$(n\backslash n)/(s/\Diamond\Box np_{io})$	np_{su}	$((np_{su}\backslash s)/np_{do})\left[\begin{smallmatrix} s/\Diamond\Box np_{io} \\ s \end{smallmatrix}\right]$	np_{do}/n	n

Use of $([\;]E)$ allows the assignment of $s/\Diamond\Box np$ to the gapped embedded sentence John offered ϵ a drink, establishing it as a proper argument for whom. The $\Diamond$ and $\Box$ are unary type-forming operators, held subject to:

$$\Diamond\Box A \vdash A \vdash \Box\Diamond A$$

thereby preventing *overgeneration*: expressions of type np do not directly combine with gapped clauses, seeing as they do not derive $\Diamond\Box np$. We make the following observations:

1. Given that $([\;]E)$ may operate at an arbitrarily deep level within a context, our type schema for extraction places no constraints on the location of the gap.[4]

2. Our proposal relies on a limited amount of lexical ambiguity: types $A\left[\begin{smallmatrix} B \div \Diamond \Box C \\ B \end{smallmatrix}\right]$ are lexically assigned to predicates *in addition* to the usual types for *non*–extraction environments. Although the amount of ambiguity is well under control (it being finite), we show in the next section for several specific cases that both types are derivable from a common ancestor.

[3]This is essentially the type constructor proposed by Moortgat (1992), although he used the notation $q(A, B, C)$ and described its proof-theory in sequent calculus. Our notation is borrowed from Shan (2002), where it was used as an abbreviation for the lifted types $(A \to B) \to C$ in the continuation semantics of Barker (2002). The analogy is made on purpose: under the decomposition of types $A\left[\begin{smallmatrix} C \\ B \end{smallmatrix}\right]$ in **LG** in the next section, they obtain a proof-theoretic behavior close to what we find in Barker's work.

[4]Compare this to the situation in **L**, where only *peripheral* extraction is possible, at least as long as one does not wish to resort to the use of types that are highly specialized to their syntactic environments.

3. Our analysis is, of course, not limited to the case where the extraction domain is a sentence. For example, with wh-extraction in English the gap shows up at the level of a yes–no question (Vermaat 2005, Chapter 3):

Whom	did	John	offer		a	drink?
$wh/(q/\diamond\Box np)$	q/s_{inf}	np	$((np\backslash s_{inf})/np)$	$\left[\begin{array}{c} q/\diamond\Box np \\ q \end{array}\right]$	np/n	n

Here q, *wh* are the types of yes–no and wh–questions respectively.

3. The Lambek-Grishin Calculus

Various extensions of **(N)L** have been proposed in the literature that sought to decompose types $A\left[\begin{array}{c} C \\ B \end{array}\right]$ in terms of more primitive type constructors. Here, we consider in particular the efforts of Bernardi & Moortgat (2007) in the Lambek–Grishin calculus (Moortgat 2007). The latter work sought to establish an extension of **NL** exhibiting an *involutive* operation $(\cdot^{\infty})$ on types and sequents, manifesting itself at the level of derivations as an *arrow–reversing duality*:

$$\Gamma \vdash A \text{ is derivable} \quad \Leftrightarrow \quad A^{\infty} \vdash \Gamma^{\infty} \text{ is derivable}$$

More accurately, once we admit of some notion of *co*structure Π naturally occurring on the *right*hand side of the turnstile, we wish to establish:

$$\Gamma \vdash \Pi \text{ is derivable} \quad \Leftrightarrow \quad \Pi^{\infty} \vdash \Gamma^{\infty} \text{ is derivable}$$

Formally, we start by adding *co*slashes, or *subtractions*, to the basic repertoire of type-forming operations:

$$\begin{array}{lllllllllll} A..E & ::= & n & \mid & np & \mid & s & \mid & (B\backslash A) & \mid & (A/B) \\ & & & & & & & \mid & (A \oslash B) & \mid & (B \obslash A) \end{array}$$

Realizing $(\cdot^{\infty})$ as:

$$\begin{array}{lll@{\qquad}lll} (B/A)^{\infty} & =_{def} & (A^{\infty} \obslash B^{\infty}) & (B \oslash A)^{\infty} & =_{def} & (A^{\infty}\backslash B^{\infty}) \\ (A\backslash B)^{\infty} & =_{def} & (B^{\infty} \oslash A^{\infty}) & (A \obslash B)^{\infty} & =_{def} & (B^{\infty}/A^{\infty}) \end{array}$$

And $A^{\infty} =_{def} A$ if A is atomic. That $(\cdot^{\infty})$ is involutive ($A^{\infty\infty} = A$) is easily checked. At the level of sequents, we require a notion of *co*structure:

$$\Pi, \Sigma \quad ::= \quad A \quad \mid \quad (\Pi \bullet \Sigma) \qquad \left(\begin{array}{lll} (\Gamma \circ \Delta)^{\infty} & =_{def} & (\Delta^{\infty} \bullet \Gamma^{\infty}) \\ (\Pi \bullet \Sigma)^{\infty} & =_{def} & (\Sigma^{\infty} \circ \Gamma^{\infty}) \end{array}\right)$$

where, for the extension of $(\cdot^{\infty})$, we rely on its definition at the level of types for the base cases (where Γ or Π is A). *Co*contexts $\Pi[]$ are defined similarly to contexts $\Gamma[]$. Derivations are now defined through the following inference rules:

$$\frac{}{A \vdash A}\,Ax \qquad \frac{A \circ \Gamma \vdash \Pi[B]}{\Gamma \vdash \Pi[(A\backslash B)]}\,\backslash I \qquad \frac{\Gamma \circ A \vdash \Pi[B]}{\Gamma \vdash \Pi[(B/A)]}\,/I$$

$$\frac{\Delta \vdash \Pi[A] \quad \Gamma[A] \vdash \Sigma}{\Gamma[\Delta] \vdash \Pi[\Sigma]}\,Cut \qquad \frac{\Gamma[B] \vdash \Pi \bullet A}{\Gamma[(B \oslash A)] \vdash \Pi}\,\oslash I \qquad \frac{\Gamma[B] \vdash A \bullet \Pi}{\Gamma[(A \obslash B)] \vdash \Pi}\,\obslash I$$

$$\frac{\Delta \vdash B \quad \Gamma \vdash (B\backslash A)}{\Delta \circ \Gamma \vdash A}\,\backslash E \qquad \frac{\Gamma \vdash (A/B) \quad \Delta \vdash B}{\Gamma \circ \Delta \vdash A}\,/E$$

$$\frac{(A \oslash B) \vdash \Pi \quad B \vdash \Sigma}{A \vdash \Pi \bullet \Sigma}\,\oslash E \qquad \frac{B \vdash \Sigma \quad (B \obslash A) \vdash \Pi}{A \vdash \Sigma \bullet \Pi}\,\obslash E$$

where, for Cut, either $\Gamma[] = []$ or $\Pi[] = []$. Note that the latter rule is *not* eliminable, as with usual Natural Deduction. We establish $(\cdot^\infty)$ to be an arrow–reversing duality through a straightforward extension of $(\cdot^\infty)$ to the level of derivations. For instance:

$$\left(\frac{\Gamma \vdash (A/B) \quad \Delta \vdash B}{\Gamma \circ \Delta \vdash A}\,/E \right)^\infty =_{def} \frac{B^\infty \vdash \Delta^\infty \quad (B^\infty \obslash A^\infty) \vdash \Gamma^\infty}{A^\infty \vdash \Delta^\infty \bullet \Gamma^\infty}\,\obslash E$$

The calculus thus derived is essentially the Lambek-Grishin calculus, though here presented in a format more Natural Deduction-esque. Compared to **NL**, its expressivity is enhanced through the introduction rules: the inferred (co–)slash may occur arbitrarily deep inside the right–hand side (left–hand side) of the sequent. This property is exploited by Bernardi & Moortgat (2007) through the decomposition of types $A\left[\begin{smallmatrix} C \\ B \end{smallmatrix}\right]$ from §2 as $(C \oslash B) \obslash A$. Indeed, $([\,]E)$ then becomes *derivable* in **LG**:

$$\dfrac{\Delta \vdash ((B \oslash C) \obslash A) \qquad \dfrac{\dfrac{\Gamma[A] \vdash B \qquad \dfrac{\dfrac{}{(B \oslash C) \vdash B \oslash C}\,Ax \quad \dfrac{}{C \vdash C}\,Ax}{B \vdash (B \oslash C) \bullet C}\,\oslash E}{\Gamma[A] \vdash (B \oslash C) \bullet C}\,Cut}{\Gamma[((B \oslash C) \obslash A)] \vdash C}\,\obslash I}{\Gamma[\Delta] \vdash C}\,Cut$$

thereby establishing our analysis of extraction to be applicable to **LG**. However, the particular use of types $(B \oslash C) \obslash A$ allows for further improvements. Recall from the previous examples our use of two separate types for offered, depending on whether or not it was used in an extraction environment. In **LG**, we can find a type D from which both may be *derived*:

$$\begin{array}{lcll} D & \vdash & ((np_{su}\backslash s)/np_{do})/np_{io} & \textit{(No extraction)} \\ D & \vdash & ((np_{su}\backslash s)/np_{do})\left[\begin{array}{c} s/\diamond\Box np_{io} \\ s \end{array}\right] & \textit{(Extraction)} \end{array}$$

whereas $D \nvdash (np_{su}\backslash(s/np_{io}))/np_{do}$ and $D \nvdash ((np_{su}\backslash s)/np_{io})/np_{do}$, preventing overgeneration. Abbreviating $((np_{su}\backslash s)/np_{do})/np_{io}$ as *dtv*, we have the following solution for D:[5]

$$\left((s/s)\left[\begin{array}{c} dtv \\ s \end{array}\right] \otimes s\left[\begin{array}{c} s/\diamond\Box np_{io} \\ s \end{array}\right] \right) \otimes \diamond\Box np_{io}$$

[5] Here, $(\cdot \otimes \cdot)$ is the counterpart of the structural connective $(\cdot \circ \cdot)$ at the level of types:

$$\frac{\Gamma \vdash A \otimes B \quad \Delta[A \circ B] \vdash C}{\Gamma[\Delta] \vdash C}\,\otimes E \qquad \frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma \circ \Delta \vdash A \otimes B}\,\otimes I$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{s/s \vdash s/s^*}\,Ax}{s/s \circ s^* \vdash s}\,/E'}{s/s \circ s\left[\begin{array}{c} s/\Diamond\Box np_{io} \\ s \end{array}\right] \vdash s/\Diamond\Box np_{io}^*}\,[\,]E'}{(s/s^* \circ s\left[\begin{array}{c} s/\Diamond\Box np_{io} \\ s \end{array}\right]) \circ \Diamond\Box np_{io} \vdash s}\,/E'}{((s/s)\left[\begin{array}{c} dtv \\ s \end{array}\right] \circ s\left[\begin{array}{c} s/\Diamond\Box np_{io} \\ s \end{array}\right]) \circ \Diamond\Box np_{io} \vdash dtv}\,[\,]E'
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{s/s \vdash s/s^*}\,Ax}{s/s^* \circ s \vdash s}\,/E'}{(s/s)\left[\begin{array}{c} dtv \\ s \end{array}\right] \circ s \vdash dtv^*}\,[\,]E'}{((s/s)\left[\begin{array}{c} dtv \\ s \end{array}\right] \circ s) \circ \Diamond\Box np_{io} \vdash (np_{su}\backslash s)/np_{do}^*}\,/E'}{((s/s)\left[\begin{array}{c} dtv \\ s \end{array}\right] \circ s^*) \circ \Diamond\Box np_{io} \vdash s \oslash (s/\Diamond\Box np_{io}) \bullet ((np_{su}\backslash s)/np_{do})\left[\begin{array}{c} s/\Diamond\Box np_{io} \\ s \end{array}\right]}\,\obslash E'}{((s/s)\left[\begin{array}{c} dtv \\ s \end{array}\right] \circ s\left[\begin{array}{c} s/\Diamond\Box np_{io} \\ s \end{array}\right]) \circ \Diamond\Box np_{io} \vdash ((np_{su}\backslash s)/np_{do})\left[\begin{array}{c} s/\Diamond\Box np_{io} \\ s \end{array}\right]}\,\obslash I
$$

Figure 1: Explaining the intuition behind D. Active formulas are explicitly marked with a suffix * in order to faciliate understanding.

thus at least partially eliminating the need for (lexically) ambiguous type assignment. Before moving on to the more general case, we explain the intuition behind D in Figure 1, using the following *derived* rules in order to abstract away from all unnecessary details:

$$
\dfrac{\Gamma[A] \vdash B}{\Gamma[A\left[\begin{array}{c} C \\ B \end{array}\right]] \vdash C}\,[\,]E' \qquad \dfrac{\Gamma \vdash A/B \;\; (C \in \{B, \Diamond\Box B\})}{\Gamma \circ C \vdash A}\,/E' \qquad \dfrac{\Gamma \vdash B}{\Gamma \vdash C \bullet C \obslash B}\,\obslash E'
$$

In the more general case, we can find such a D for types A/B and $A\left[\begin{array}{c} C/\Diamond\Box B \\ C \end{array}\right]$ provided $C = head(A)$ (e.g., if $A = (np\backslash s)/np$, then $C = s$), the solution for D then being:[6]

$$
\left((C/C)\left[\begin{array}{c} A/B \\ C \end{array}\right] \otimes C\left[\begin{array}{c} C/\Diamond\Box B \\ C \end{array}\right] \right) \otimes \Diamond\Box B
$$

For the reader familiar with Barker and Shan's work on continuation semantics, we conclude this section with the following observation. We may understand our analysis of extraction, when translated to **LG**, to involve an application of continuations by the following reasoning: types $(B \oslash C) \obslash A$ behave like the computations $(A \to B) \to C$ of Barker (2002), in that the following lifted elimination rules become **LG**–*derivable* (and similarly for \), provided that we allow *co*structures to be associative:

$$
\dfrac{\Gamma \vdash (A/B)\left[\begin{array}{c} E \\ D \end{array}\right] \quad \Delta \vdash B\left[\begin{array}{c} D \\ C \end{array}\right]}{\Gamma \circ \Delta \vdash A\left[\begin{array}{c} E \\ C \end{array}\right]} \qquad \dfrac{\Gamma \vdash (A/B)\left[\begin{array}{c} D \\ C \end{array}\right] \quad \Delta \vdash B\left[\begin{array}{c} E \\ D \end{array}\right]}{\Gamma \circ \Delta \vdash A\left[\begin{array}{c} E \\ C \end{array}\right]}
$$

[6]The case of extraction from a left branch is similar.

In this sense, our analysis of extraction may be perceived of as an alternative to the type-logical treatment of continuations found in Barker & Shan (2006).

4. Formal Semantics

In coupling CTL with a Montagovian semantics, we first define the semantic counterparts of syntactic types (referred to as *(semantic) types* when no confusion arises) in a suitable fragment of intuitionistic logic. The latter we shall fix to the $\{\times, \rightarrow, e, \bot\}$-fragment of minimal logic, with types σ, τ given by e (for entities), $\bot$ (for truth–values), *function* types $(\sigma \rightarrow \tau)$ and *product* types $(\sigma \times \tau)$. For **NL**, this gives the familiar mapping

$$\begin{array}{llllllll} [\![A \div B]\!] & =_{def} & [\![B]\!] \rightarrow [\![A]\!] & [\![\Diamond A]\!] & =_{def} & [\![A]\!] & [\![\Box A]\!] & =_{def} \ [\![A]\!] \\ [\![s]\!] & =_{def} & \bot & [\![np]\!] & =_{def} & e & [\![n]\!] & =_{def} \ e^{\bot} \end{array}$$

Here, $A \div B$ may be either A/B or $B\backslash A$, and $\phi \rightarrow \bot$ abbreviates $\phi^{\bot}$ (writing $\phi^{\bot\bot}$ for $(\phi^{\bot})^{\bot}$). For **LG**, we shall adopt the *continuation-passing style* translation of Bernardi & Moortgat (2007), coupling each syntactic type A with its *continuation* $[\![A]\!]_K$. For atomic types, the latter may be seen as an abstraction from $\bot$ (typing the sentence denotation) over the **NL** denotation (of type) $[\![A]\!]$:

$$[\![s]\!]_K \ =_{def} \ \bot^{\bot} \qquad [\![np]\!]_K \ =_{def} \ e^{\bot} \qquad [\![n]\!]_K \ =_{def} \ e^{\bot\bot}$$

In particular, we aim to adorn lexical entries of syntactic type A with denotations of type $[\![A]\!]_K^{\bot}$ (referred to as an A *computation*), carrying abstractions over their own continuations. For derived type $(A \div B)$ $(\div \in \{\backslash, /\})$, we wish such *lifted* denotations to be of type $[\![A \div B]\!]_K^{\bot} = [\![B]\!]_K^{\bot} \rightarrow [\![A]\!]_K^{\bot}$.[7] Realizing that, by (de)currying, the latter is isomorphic to $([\![B]\!]_K^{\bot} \times [\![A]\!]_K)^{\bot}$ and $([\![A]\!]_K \times [\![B]\!]^{\bot})^{\bot}$, this means we can define:

$$[\![A/B]\!]_K \ =_{def} \ [\![B]\!]_K^{\bot} \times [\![A]\!]_K \qquad [\![B\backslash A]\!]_K \ =_{def} \ [\![A]\!]_K \times [\![B]\!]_K^{\bot}$$

To finish, continuations associated with the types $(A \oslash B)$ and $(B \obslash A)$ are dual to those carrying a slash as principal sign in the following sense:

$$\begin{array}{lll} [\![A \oslash B]\!]_K & =_{def} & [\![A\backslash B]\!]_K^{\bot} \\ & = & ([\![B]\!]_K \times [\![A]\!]_K^{\bot})^{\bot} \end{array} \qquad \begin{array}{lll} [\![B \obslash A]\!]_K & =_{def} & [\![B/A]\!]_K^{\bot} \\ & = & ([\![A]\!]_K^{\bot} \times [\![B]\!]_K)^{\bot} \end{array}$$

And, as before, we identify $[\![\Diamond A]\!]_K$ and $[\![\Box A]\!]_K$ with $[\![A]\!]_K$. In illustrating the application of these definitions to our analysis of extraction, we return to one of our previous examples:[8]

lady	whom	John	offered		a drink
n	$(n\backslash n)/(s/\Diamond\Box np)$	np	$((np\backslash s)/np)$	$\left[\ s/\underset{s}{\Diamond}\Box np\ \right]$	np

[7]This follows the *call–by–name* evaluation paradigm. Normally, the latter postulates an additional double negation 'layer', as in $([\![B]\!]_K^{\bot} \rightarrow [\![A]\!]_K^{\bot})^{\bot\bot}$, identifying $[\![A \div B]\!]_K$ with $([\![B]\!]_K^{\bot} \rightarrow [\![A]\!]_K^{\bot})^{\bot}$. The use of product types, however, allows for the simpler treatment pursued here, as explained in the main text (see also the work of Yves Lafont & Streicher (1993)).

[8]In order to focus on the semantics of offered, we simplify our analysis by taking a drink to constitute a single lexical entry.

We analyze, in turn, its *derivational* and *lexical* semantics. As for the former, our task is to inductively extract a λ-term $\llbracket \mathscr{D} \rrbracket^{\mathbf{LG}}$ (coding a *proof* in minimal logic) from the derivation $\mathscr{D}$ of lady whom John offered a drink, telling us how the latter's denotation is to be obtained from those of the words it contains. In doing so, we understand the derivation of a sequent $\Gamma \vdash A$ (the only kind that we need to consider for the current example) to be mapped into a function from computations of the types found in Γ to a computation of A.

In addition to the usual *applications* $(M^{\sigma\to\tau}\ N^{\sigma})^{\tau}$ and *abstractions* $(\lambda x^{\sigma} M^{\tau})^{\sigma\to\tau}$, we allow $\llbracket \mathscr{D} \rrbracket^{\mathbf{LG}}$ to be derived using *pairings* $\langle M^{\sigma}, N^{\tau}\rangle^{\sigma\times\tau}$ and *projections* $(\pi^{l} M^{\sigma\times\tau})^{\sigma}$ and $(\pi^{r} M^{\sigma\times\tau})^{\tau}$, although in practice we shall usually mention explicitly only the types of *bound* variables. Initially, we have only the elimination rules $(/E), (\backslash E)$ to deal with:[9]

$$\mathscr{D}_1 \quad = \quad \cfrac{\cfrac{\text{John}}{np} \qquad \cfrac{\cfrac{}{(np\backslash s)/np \vdash (np\backslash s)/np}\,Ax \qquad \cfrac{\text{a\_drink}}{np}}{(np\backslash s)/np \circ \text{a\_drink} \vdash np\backslash s}\,/E}{\text{John} \circ ((np\backslash s)/np \circ \text{a\_drink})) \vdash s}\,\backslash E$$

The associated meaning recipe is obtained by decurried application:

$$\begin{array}{lcl} \llbracket \mathscr{D}_1 \rrbracket^{\mathbf{LG}} & = & \lambda\gamma^{\llbracket s \rrbracket_K}(q\ \langle \llbracket \text{a\_drink} \rrbracket^{\mathbf{LG}}, \langle \gamma, \llbracket \text{John} \rrbracket^{\mathbf{LG}}\rangle\rangle) \\ & \simeq & ((q\ \llbracket \text{a\_drink} \rrbracket^{\mathbf{LG}})\ \llbracket \text{John} \rrbracket^{\mathbf{LG}}) \end{array}$$

parameterizing over lexical denotations $\llbracket w \rrbracket^{\mathbf{LG}}$ of semantic type $\llbracket A \rrbracket^{\perp}_{K}$ for words w of syntactic type A, and a free variable q of type $\llbracket (np\backslash s)/np \rrbracket^{\perp}_{K}$. Proceeding:

$$\mathscr{D}_2 \quad = \quad \cfrac{\cfrac{\text{offered}}{((np\backslash s)/np)\left[\begin{matrix} s/\Diamond\Box np \\ s \end{matrix}\right]} \qquad \begin{matrix}\mathscr{D}_1 \\ \text{John} \circ ((np\backslash s)/np \circ \text{a\_drink}) \vdash s\end{matrix}}{\text{John} \circ (\text{offered} \circ \text{a\_drink}) \vdash s/\Diamond\Box np}\,[\ ]E$$

In the decomposition of $([\ ]E)$ in **LG**, this gives us the term $\llbracket \mathscr{D}_2 \rrbracket^{\mathbf{LG}}$:

$$\lambda\delta^{\llbracket s/\Diamond\Box np \rrbracket_K}(\llbracket \text{offered} \rrbracket^{\mathbf{LG}}\ \lambda\langle q^{\llbracket (np\backslash s)/np \rrbracket^{\perp}_{K}}, y^{\llbracket s/\Diamond\Box np \rrbracket_K \times \llbracket s \rrbracket^{\perp}_{K}}\rangle(y\ \langle \delta, \llbracket \mathscr{D}_1 \rrbracket^{\mathbf{LG}}\rangle))$$

This term reminds us of Montague's Quantify In: $\llbracket \text{offered} \rrbracket^{\mathbf{LG}}$ is essentially applied to an abstraction over the occurrence of q in M_1. Concluding our derivation $\mathscr{D}$, we have:

$$\cfrac{\cfrac{\text{lady}}{n} \qquad \cfrac{\cfrac{\text{whom}}{(n\backslash n)/(s/\Diamond\Box np)} \qquad \begin{matrix}\mathscr{D}_2 \\ \text{John} \circ (\text{offered} \circ \text{a\_drink}) \vdash s/\Diamond\Box np\end{matrix}}{\text{whom} \circ (\text{John} \circ (\text{offered} \circ \text{a\_drink})) \vdash n\backslash n}\,/E}{\text{lady} \circ (\text{whom} \circ (\text{John} \circ (\text{offered} \circ \text{a\_drink}))) \vdash n}\,\backslash E$$

$$\llbracket \mathscr{D} \rrbracket^{\mathbf{LG}} \quad = \quad \lambda\epsilon^{\llbracket n \rrbracket_K}(\llbracket \text{whom} \rrbracket^{\mathbf{LG}}\ \langle \llbracket \mathscr{D}_2 \rrbracket^{\mathbf{LG}}, \langle \epsilon, \llbracket \text{lady} \rrbracket^{\mathbf{LG}}\rangle\rangle)$$

With the *lexical* semantics, our task is to substitute appropriate terms in $\llbracket \mathscr{D} \rrbracket^{\mathbf{LG}}$ for

$$\begin{array}{lcllcl} \llbracket \text{lady} \rrbracket^{\mathbf{LG}} & :: & \llbracket n \rrbracket^{\perp}_{K} & \llbracket \text{that} \rrbracket^{\mathbf{LG}} & :: & \llbracket (n\backslash n)/(s/\Diamond\Box np) \rrbracket^{\perp}_{K} \\ \llbracket \text{John} \rrbracket^{\mathbf{LG}} & :: & \llbracket np \rrbracket^{\perp}_{K} & \llbracket \text{offered} \rrbracket^{\mathbf{LG}} & :: & \llbracket (s \oslash (s/\Diamond\Box np)) \obslash (np\backslash s) \rrbracket^{\perp}_{K} \\ \llbracket \text{a\_drink} \rrbracket^{\mathbf{LG}} & :: & \llbracket np \rrbracket^{\perp}_{K} & & & \end{array}$$

---

[9] We directly incorporate lexical substitutions in our derivations by writing $\frac{w}{A}$ for axioms when w is a word that is lexically assigned the type A.

such that the result becomes logically equivalent to:

$$\lambda\epsilon^{[\![n]\!]}.(\epsilon\ \lambda z^e.((\text{LADY}\ z) \wedge (\exists\ \lambda y^e((\text{DRINK}\ y) \wedge (\text{OFFER}\ z\ y\ \text{JOHN})))))$$

We discuss for each word w how its usual denotation $[\![w]\!]^{\mathbf{NL}}$ in **NL** may be lifted to $[\![w]\!]^{\mathbf{LG}}$. For this task, we assume to have at our disposal the constants in the following table:

Constant(s)	**Type**	**Description** (denotation)
$\exists$	$e^{\perp\perp}$	Existential quantification
$=$	$e \to e^{\perp}$	Equality of entities
$\wedge$	$\perp \to \perp^{\perp}$	Conjunction
JOHN	e	Entity
LADY, DRINK	$e^{\perp}$	(First–order) Properties
OFFER	$e \to e \to e^{\perp}$	Ternary relation

Table 1: Additional constants assumed for the construction of lexical denotations.

The cases $[\![\text{John}]\!]^{\mathbf{LG}}$ and $[\![\text{lady}]\!]^{\mathbf{LG}}$ offer little challenge: we simply 'lift' JOHN (= $[\![\text{John}]\!]^{\mathbf{NL}}$) and LADY (= $[\![\text{lady}]\!]^{\mathbf{NL}}$):

$$\begin{array}{lcl} [\![\text{John}]\!]^{\mathbf{LG}} & = & (\mathtt{LIFT}^{\perp}_{e}\ \text{JOHN}) \\ [\![\text{lady}]\!]^{\mathbf{LG}} & = & (\mathtt{LIFT}^{\perp}_{e^{\perp}}\ \text{LADY}) \end{array} \qquad \left(\begin{array}{ll} & \mathtt{LIFT}^{\tau}_{\sigma} \\ =_{def} & \lambda x^{\sigma}\lambda\alpha^{\sigma\to\tau}(\alpha\ x) \end{array} \right)$$

Whereas for $[\![\text{a\_drink}]\!]^{\mathbf{LG}}$ we can take $\lambda\alpha^{e^{\perp}}\exists y^e((\text{DRINK}\ y) \wedge (\alpha\ y))$. With whom, things start to get more interesting. Our starting point is again a denotation in **NL**:

$$\begin{array}{lcl} [\![\text{whom}]\!]^{\mathbf{NL}} & = & \lambda q^{[\![s/\Diamond\Box np]\!]}\lambda r^{[\![n]\!]}\lambda x^e((r\ x) \wedge (q\ x)) \\ & :: & [\![(n\backslash n)/(s/\Diamond\Box np)]\!] \end{array}$$

$[\![\text{whom}]\!]^{\mathbf{LG}}$, on the other hand, is to be of the form of a paired abstraction

$$\begin{array}{lcl} \lambda\langle Q^{[\![s/\Diamond\Box np]\!]^{\perp}_{K}}, \langle\alpha^{[\![n]\!]_K}, R^{[\![n]\!]^{\perp}_{K}}\rangle\rangle M & :: & ([\![s/\Diamond\Box np]\!]^{\perp}_{K} \times ([\![n]\!]_K \times [\![n]\!]^{\perp}_{K}))^{\perp} \\ & (= & [\![(n\backslash n)/(s/\Diamond\Box np)]\!]^{\perp}_{K}) \end{array}$$

For M, we shall now explain the following to be a suitable instantiation:

$$(R\ \lambda r^{e^{\perp}}(\alpha\ (([\![\text{whom}]\!]^{\mathbf{NL}}\ \lambda x^e(Q\ \langle\mathtt{LIFT}^{\perp}_{e}\ x, \lambda p^{\perp}p\rangle))\ r)))$$

This term is dissected step by step in the following series of equivalences:

$$\begin{array}{lcl} M_1 & =_{def} & ([\![\text{whom}]\!]^{\mathbf{NL}}\ \lambda x(Q\ \langle\mathtt{LIFT}^{\perp}_{e}\ x, \lambda p^{\perp}p\rangle)) \\ & \equiv_{\beta} & \lambda r^{e^{\perp}}\lambda x^e((r\ x) \wedge (Q\ \langle\mathtt{LIFT}^{\perp}_{e}\ x, \lambda p^{\perp}p\rangle)) \\ & :: & e^{\perp} \to e^{\perp} \\ M_2 & =_{def} & \lambda r^{e^{\perp}}(\alpha\ (M_1\ r)) \\ & \equiv_{\beta} & \lambda r^{e^{\perp}}(\alpha\ \lambda x^e((r\ x) \wedge (Q\ \langle\mathtt{LIFT}^{\perp}_{e}\ x, \lambda p^{\perp}p\rangle))) \\ & :: & e^{\perp\perp} \\ M & =_{def} & (R\ M_2) \\ & :: & \perp \end{array} \qquad \begin{array}{c} \left(\begin{array}{lcl} Q & :: & [\![s/\Diamond\Box np]\!]_K \\ & = & [\![np]\!]^{\perp}_{K} \times [\![s]\!]_K \end{array} \right) \\ \left(\begin{array}{lcl} \alpha & :: & [\![n]\!]_K \\ & = & e^{\perp\perp} \end{array} \right) \\ \left(\begin{array}{lcl} R & :: & [\![n]\!]^{\perp}_{K} \\ & = & e^{\perp\perp\perp} \end{array} \right) \end{array}$$

The real challenge we find with $[\![\text{offered}]\!]^{\perp}_{K}$. It is again to be an abstraction

$$\lambda Q^{([\![(np\backslash s)/np]\!]_K^\perp \times [\![s \oslash (s/\Diamond\Box np)]\!]_K)^\perp} N \quad :: \quad ([\![(np\backslash s)/np]\!]_K^\perp \times [\![s \oslash (s/\Diamond\Box np)]\!]_K)^{\perp\perp}$$
$$(= \quad [\![(s \oslash (s/\Diamond\Box np)) \oslash ((np\backslash s)/np)]\!]_K^\perp)$$

Q takes terms of type $[\![(np\backslash s)/np]\!]_K^\perp$ and $[\![s \oslash (s/\Diamond\Box np)]\!]_K$ into the result type $\perp$. For the first of these arguments we take

$$\begin{array}{lll} N_1 & =_{def} & \lambda\langle Y^{[\![np]\!]_K^\perp}, \langle \gamma^{[\![s]\!]_K}, X^{[\![np]\!]_K^\perp}\rangle\rangle (Y\ \lambda y^e(X\ \lambda x^e(\gamma\ (\text{OFFER}\ u\ y\ x)))) \\ & :: & ([\![np]\!]_K^\perp \times ([\![s]\!]_K \times [\![np]\!]_K^\perp))^\perp = [\![(np\backslash s)/np]\!]_K^\perp \end{array}$$

featuring a *free* variable u (type e) as a place-holder for the extracted argument (i.e., the indirect object). For the second argument of Q, we construe

$$\begin{array}{lll} N_2 & =_{def} & \lambda\langle\langle Z^{[\![np]\!]_K^\perp}, \beta^{[\![s]\!]_K}\rangle, w^{[\![s]\!]_K^\perp}\rangle (w\ \lambda p^\perp((\beta\ p) \wedge (Z\ \lambda u^e(u = z)))) \\ & :: & (([\![np]\!]_K^\perp \wedge [\![s]\!]_K) \wedge [\![s]\!]_K^\perp)^\perp = [\![s \oslash (s/\Diamond\Box np)]\!]_K \end{array}$$

In essence, this term contains the recipe for construing the denotation of the gapped embedded clause John offered ϵ a drink. Crucially, it attributes to the denotation of the extracted argument (the bound variable Z of type $[\![np]\!]_K^\perp$) the property of being equal to u (the free variable standing in for the indirect object in N_1). We conclude our definition of $[\![\text{offered}]\!]^{\mathbf{LG}}$ with an existential binding of u:

$$N \quad =_{def} \quad (\exists\ \lambda u^e(Q\ \langle N_1, N_2\rangle)) :: \perp$$

With the denotations for each of the lexical items fixed, extensive use of β-reduction provides us with the following term of type $[\![n]\!]_K^\perp$ as the final denotation:

$$(\mathtt{LIFT}^\perp_{[\![n]\!]_K}\ \lambda z^e((\text{LADY}\ z) \wedge \exists u^e \exists y^e((\text{DRINK}\ y) \wedge (\text{OFFER}\ u\ y\ \text{JOHN}) \wedge (u = z))))$$

Allowing us to obtain the desired result with the following equivalence (licensed by the laws of predicate logic)[10]

$$\begin{array}{ll} & \exists u^e \exists y^e((\text{DRINK}\ y) \wedge (\text{OFFER}\ u\ y\ \text{JOHN}) \wedge (u = z)) \\ \equiv & \exists y^e((\text{DRINK}\ y) \wedge (\text{OFFER}\ z\ y\ \text{JOHN})) \end{array}$$

5. Conclusion

The previous sections saw the formulation of an analysis of extraction in the Lambek–Grishin calculus. Its syntactic component was inspired by Moortgat's type constructor for discontinuity (quantifier scope ambiguities constituting its original domain of application), with the semantic component phrased in terms of a continuation semantics. In the latter case, we relied on a λ–calculus with equality in order to identify the extracted argument with its gap position. An obvious direction for future research is as follows. Our compositional semantics of §4 targeted our type schema for extraction in §2. However, we have also considered in §3 a type D that can derive both (specific) instantiations of this schema as well as types for *non*–extraction environments. Can our semantic analysis of §4 be generalized so as to be applicable to D?

Acknowledgements

I am grateful to the following people for their comments on earlier drafts of this paper: Michael Moortgat, Raffaella Bernardi, Jan van Eijck, Andres Löh, Gianluca Giorgolo, Christina Unger, as well as two anonymous referees.

[10]Compare this to Montague's PTQ treatment of *to be*.

References

C. Barker (2002). 'Continuations and the Nature of Quantification'. *Natural Language Semantics* **10**(3):211–242.

C. Barker & C.-C. Shan (2006). 'Types as Graphs: Continuations in Type Logical Grammar'. *Journal of Logic, Language and Information* **15**(4):331–370.

R. Bernardi & M. Moortgat (2007). 'Continuation Semantics for Symmetric Categorial Grammar.'. In D. Leivant & R. J. G. B. de Queiroz (eds.), *WoLLIC*, vol. 4576 of *Lecture Notes in Computer Science*, pp. 53–71. Springer.

J. Lambek (1958). 'The mathematics of sentence structure'. *American Mathematical Monthly* **65**:154–169.

J. Lambek (1961). 'On the calculus of syntactic types'. In R. Jakobson (ed.), *Structure of Language and its Mathematical Aspects, Proceedings of the Twelfth Symposium in Applied Mathematics*, pp. 166–178.

M. Moortgat (1992). 'Generalized quantifiers and discontinuous type constructors'. In A. Horck & W. Sijtsma (eds.), *Discontinuous Constituency*, pp. 181–208, Berlin. Mouton de Gruyter.

M. Moortgat (1997). 'Categorial Type Logics'. In J. F. A. K. van Benthem & G. B. A. ter Meulen (eds.), *Handbook of Logic and Language*, pp. 93–177. Elsevier.

M. Moortgat (2007). 'Symmetries in Natural Language Syntax and Semantics: The Lambek-Grishin Calculus.'. In D. Leivant & R. J. G. B. de Queiroz (eds.), *WoLLIC*, vol. 4576 of *Lecture Notes in Computer Science*, pp. 264–284. Springer.

C. Shan (2002). 'A continuation semantics of interrogatives that accounts for Baker's ambiguity'. In B. Jackson (ed.), *Proceedings of SALT XII: semantics and linguistic theory*, pp. 246–265.

W. Vermaat (2005). *The logic of variation. A cross-linguistic account of wh-question formation*. Ph.D. thesis, Utrecht Institute of Linguistics OTS, Utrecht University.

B. R. Yves Lafont & T. Streicher (1993). 'Continuation semantics or expressing implication by negation'. Tech. rep., Ludwig-Maximilians-Universität, München.

TOWARDS MORPHOLOGICALLY ENHANCED AUTOMATED LEXICAL ACQUISITION

Kostadin Cholakov
Center for Language and Cognition Groningen

Abstract. In this paper we tackle the lexical coverage problem for large-scale hand-crafted grammars. Since such grammars usually rely heavily on manually created lexicons, their coverage and quality are crucial for the grammar performance. To this extent, we propose accurate machine learning methods for automated lexical acquisition and prove their high quality by applying them on a large-scale grammar for Dutch. We also emphasise on the use of morphology in the acquisition process and illustrate its importance.

1. Introduction

1.1. Motivation

At present, various wide-coverage symbolic parsing systems for different languages exist and have been integrated into real-world NLP applications, such as IE, QA, grammar checking, MT and intelligent IR. This integration, though, has reminded us of the shortcomings of symbolic systems, in particular lack of lexical coverage. The hand-crafted grammars which usually lie at the heart of symbolic parsing systems are often strongly *lexicalised*– the lexicon encodes various kinds of sophisticated linguistic information and constraints. Naturally, such hand-crafted lexicons cannot cover all words in a language which could cause lower coverage and accuracy results when, for example, the grammar is applied on domains which contain many unknown words (i.e. words that are not listed in the lexicon). Thus, it is crucial to find accurate ways to extend them and improve their quality.

In this paper we propose machine learning methods for the automated acquisition of linguistic information for missing and incomplete lexical entries. We ensure the general applicability and the quality of the algorithms we propose by applying them to a large-scale grammar for Dutch.

1.2. Previous Work

There have been various approaches to the problem of lexical acquisition (LA). (Cussens & Pulman 2000) describe a symbolic approach towards grammar extension with *inductive logic programming* and (Erbach 1990, Barg & Walther 1998, Fouvry 2003) have followed a unification-based approach. However, these approaches suffer from the problem that the generated lexical entries might be both too general or too specific. It is also doubtful if these methods can be applied to large-scale grammars.

(Zhang & Kordoni 2006, van de Cruys 2006, Cholakov, et al. 2008) have taken a machine learning approach towards the problem of LA. They treat the problem as a classification task and, by using the linguistic knowledge that is already encoded in the grammar, they assign a given word one or more categories from the lexicon. Such approaches are considered to be formalism- and language-independent and can be applied to large-scale grammars.

That is why, we choose to follow the method described in (van de Cruys 2006). However, we make some modifications in order to improve it and thus, to achieve better results. The most important one is the incorporation of more morphological knowledge into the LA process. We use more morphological features during the classification and generate the paradigm of a given unknown word, so we can use the morphological information it provides to enhance our learning algorithm.

The remainder of the paper is organised as follows. Section 2 presents the general learning algorithm. Section 3 describes the experiments and the results. Section 4 illustrates the initial steps towards building a more sophisticated method for LA. Section 5 concludes the paper and presents some ideas for future work.

2. General Algorithm

2.1. The Alpino Grammar and Parser

In our experiments we use the Alpino wide-coverage dependency parser for Dutch which is based on a large HPSG grammar (van Noord 2006). The grammar takes a 'constructional' approach, with rich lexical representations stored in the lexicon and a large number of detailed, construction specific rules (about 800). As an HPSG grammar, both the lexicon and the rule component of the Alpino grammar are organised in a multiple type inheritance system.

At the moment, the lexicon contains about 100K lexical entries and a list of about 200K named entities. Each word is assigned one or more lexical types. The lexical types encode various kinds of detailed morphosyntactic information:

(1) a. kind (child) *noun(het,count,sg)*
b. amuseert (to amuse) *verb(hebben,sg3,intransitive)*, *verb(hebben,sg3,transitive)*

The lexical type for *kind* shows that the word is a singular countable noun that goes with the *het* definite article which is used only with neutral nouns in Dutch. The verb *amuseert* is assigned two lexical types because it can be used both in a transitive and an intransitive way. Furthermore, the other features show that it is a present third person singular verb form and it forms perfect tense with the auxiliary verb *hebben*. Due to the detailed morphosyntactic information encoded in the lexical types, their number is very large– there are about 20K of them in the Alpino grammar.

Since we follow the statistical lexical acquisition approach, the goal of our methods is, given a *target type inventory*, to *assign* the correct lexical type(s) to a given unknown word.

2.2. Parsing with Universal Typeset

The target type inventory for our experiments contains only open-class lexical types: nouns, adjectives and verbs, under the assumption that the grammar is already able to handle all other cases. Furthermore, we ignore *infrequent types*. A type is considered only if there are at least 15 distinct words occurring in large Dutch newspaper corpora (∼16M sentences) which belong to it. This boils down to 641 types. For the purposes of our experiment, we call them *universal types*.

The first step in our algorithm is to extract 100 sentences from large corpora or Internet for each unknown word. These sentences are parsed with a different version of the Alpino parser where the given unknown word is assigned all universal types. The parser runs in 'best-only' mode– for each of these 100 sentences only the parse which is considered the best by the disambiguation model of the parser is preserved. Then, the lexical type that has been assigned to the unknown word in the best parse is stored. When all 100 sentences have been parsed, a list can be drawn up for each unknown word with the types that have been used and their frequency:

(2) *borstbeen* (breast bone, sternum)
noun(het,count,sg) 77
noun(de,count,pl) 12
proper_name(sg,'ORG') 4
v_noun(intransitive) 4
adjective(e) 2
tmp_noun(het,count,sg) 1

In this case, the most frequently used type, *noun(het,count,sg)* is also the correct one. By doing this first step, we allow the grammar itself to decide which lexical type is best suited for a given unknown word. This is an efficient way to incorporate the linguistic knowledge that is already encoded in the grammar in the LA process.

It is important to note that Alpino is very robust– essentially, it always produces a parse. If there is no analysis covering the whole sentence, the parser finds all parses for each substring and returns the best sequence of non-overlapping parses. During parsing, Alpino's POS tagger (Prins & van Noord 2001) keeps filtering implausible type combinations. For example, if a determiner occurs before the unknown word, all verb types are not taken into consideration. This heavily reduces the computational overload and makes parsing with universal types computationally feasible.

2.3. Adaptation of the Disambiguation Model

For the parsing method to work properly, the disambiguation model of the parser needs to be adapted. The model heavily relies on the lexicon and, based on training data, it has preferences how to parse certain phrases. For example, it has a preference to parse prepositional phrases as verb complements, if there is a verb with such subcategorization frame. However, this does not make sense when parsing with universal types because every prepositional phrase would get analysed as a verb complement and thus, if the unknown word occurs with a PP, it would always be analysed as a verb which subcategorizes for a PP.

To avoid this, (van de Cruys 2006) proposed to weight each universal type with the actual frequency it occurs in the training data. We propose an alternative solution where the disambiguation model is retrained on a specific set of sentences which is meant to fine-tune it in a way that allows it to handle an input containing a large number of unknown words. 560 words which have between 50 and 100 occurrences in the newspaper corpora have been selected and temporarily removed from the Alpino lexicon, i.e. made unknown to the grammar. We make the assumption that low frequency words are typically not listed in the lexicon and the words we chose are meant to simulate their behaviour. Then, all sentences from the Alpino treebank which contain these words are extracted and used to retrain the disambiguation model.

2.4. Statistical Type Predictor

We use both morphological features and features extracted from the parsing stage in a statistical classifier in order to predict the correct lexical type(s)of a given unknown word. These features are given in Table 1.

Features
prefix of the unknown word (length is less or equal 4)
suffix of the unknown word (length is less or equal 4)
the unknown word starts with a separable particle
the unknown word contains a hyphen
the most frequently used tags (80% threshold)
each attribute of these tags
the tags the Alpino unknown word guesser assigns to the unknown word

Table 1: Features used in the classifier

Many Dutch verbs contain separable particles which are an important morphosyntactic property and should be considered as a separate feature during the classification process:

(3) Ik *ruim* mijn kamer *op*.
I clean my room SEP PART
'I am cleaning my room.'

We use a list of common separable particles to determine if the unknown word starts with such a particle. Naturally, such an approach overgenerates but nevertheless, we found it to be helpful for guessing the correct lexical types.

The lexical types which have been used in at least 80% of the parses in the parsing stage are also used in the classifier as separate features. For example, for the word in (2) only the first two types, *noun(het,count,sg)* and *noun(de,count,pl)* are considered. Further, each attribitue of the considered types is also taken as a separate feature in an attempt to enable the classifier to make broader generalizations:

(4) a. noun(het,count,sg)– *noun<het>, noun<count>, noun<sg>*
b. noun(de,count,pl)– *noun<de>, noun<count>, noun<pl>*

Finally, Alpino employs an unknown word guesser which, based on the morphological form of the unknown word, uses various heuristics to guess the right lexical type(s) for it. It performs quite well in certain cases (e.g. compounds) and we decided to use the type(s) predicted by the guesser as features in the classifier.

A maximum entropy (ME) classifier is used to predict the type(s) of a given unknown word. The probability of a lexical type *t*, given an unknown word and its context *c* is:

$$(5) \qquad p(t|c) = \frac{exp(\sum_i \Theta_i f_i(t,c))}{\sum_{t' \in T} exp(\sum_i \Theta_i f_i(t',c))}$$

where $f_i(t, c)$ may encode arbitrary characteristics of the context and Θ is a weighting factor, estimated on training data, which maximises the entropy. Input to the classifier are features explained above and output are lexical types. It is important to note that if a certain word has more than one lexical type in the lexicon (i.e. the word is ambiguous according to the lexicon), during training, all its types are taken into account for each context. This is an attempt to enable the classifier to deal with natural ambiguity.

To train and test the classifier, 2600 words with 50 to 100 occurrences in the newspaper corpora are chosen[1] and temporarily removed from the lexicon of the grammar. This is again done under the assumption that low frequency words tend to be unknown or problematic for the grammar. 2000 words are used to train the classifier and the remaining 600 form the test set.

3. Experiments and Evaluation

3.1. Initial Experiments and Results

The classifier yields a probability score for each predicted type. In our experiments, for each test word, only the types that sum up to 95% of the total sum of the probability scores are preserved. Since one word can have more than one correct type, we evaluate the results in terms of precision and recall. Precision indicates how many types found by our method are correct and recall indicates how many of the lexical types of a given word are actually found. The presented results are the average precision and recall for the 600 test words.

Additionally, we develop three baseline methods:

- *naive baseline*– each unknown word is assigned the most frequent type in the lexicon, namely *noun(de,count,sg)*
- *POS tagger baseline*– the unknown word is given the type most frequently assigned by the Alpino POS tagger in the parsing stage
- *Alpino baseline*– the unknown word is assigned the most frequently used type in the parsing stage

The overall results are given in Table 2. Table 3 shows the results for each POS in the ME-based model.

Model	Precision(%)	Recall(%)	F-measure(%)
Naive baseline	19.60	18.77	19.17
POS tagger baseline	30	26.21	27.98
Alpino baseline	45.40	38.72	41.79
ME-based model	79.71	81.58	80.63

Table 2: Overall experiment results

POS	Precision(%)	Recall(%)	F-measure(%)
Nouns	85.93	88.49	87.19
Adjectives	69.18	79.93	74.17
Verbs	62.54	59.05	60.75

Table 3: ME-based model detailed results

The performance of the baseline methods is pretty low which confirms that the task of learning unknown words is not trivial. Though it is often argued that a POS tagger might be sufficient for such tasks, we see clearly that this is not the case here. The low performance of the tagger baseline is due to the large tagset (641 tags) and the fact that the

[1]Naturally, we have not selected words which had been used for the retraining of the disambiguation model

Alpino types also contain sophisticated syntactic features which cannot be captured by a POS tagger. Our method, on the other hand, does not suffer from such shortcomings and it has the best performance. It also outperforms the F-measure result reported in (van de Cruys 2006) by 6% despite that it uses a much larger target type inventory (641 types against 340).

The results in Table 3 show that nouns are already easy to learn. Adjectives, on the other hand, seem to be difficult. This is partly due to the fact that Alpino employs a rather complicated adjective system. Different distinctions exist for adjectives that can be used predicatively, attributively and so on. These features are pretty difficult to be captured by the classifier mainly because they are based on semantic properties and our features are purely morphosyntactic. However, verbs cause most difficulties for the classifier. Most of the problems are due to the fact that the classifier has difficulties capturing the right subcategorization frames.

Here are some examples of the typical problems for the ME-based model:

- Too many wrong analyses for morphologically ambiguous words

(6) *OESO-landen* (countries of the OESO organisation) has 1 correct type but receives 25 predictions because *landen* is also the Dutch verb for 'to land'; even the unknown word guesser guessed only verb types and thus, the word receives many verb predictions

- Often verbs and adjectives get predictions that differ only in the subcat frame:

(7) *onderschrijf* (to support) has 2 possible subcat frames but receives 8 predictions which differ only in the subcat features

- *De/het* distinction for nouns– depending on their gender and number, Dutch nouns are used either with the *de* definite article (for masculine and feminine, and also plural) or the *het* definite article (for neuter)

3.2. Use of Word Paradigms to Enhance LA

Many of the wrong predictions could be avoided by enhancing the LA process with more morphological information. The paradigm of the unknown word is one very good source of such information. For example, since the definite article is distinguishable only in the singular noun form, we can determine the correct article of a word, assigned a noun plural type, if we know its singular form.

The knowledge we extract from the word paradigm could be then incorporated back into the lexical type predictor in the form of additional features which is likely to improve its performance. Another important advantage of using word paradigms is the fact that it would allow the LA process to assign lexical types to all forms of the unknown word, i.e. to learn the *whole paradigm* and not only a particular form.

We employ the method described in (Cholakov & van Noord 2009) where a simple finite state morphology is applied to generate paradigms for each unknown word. Since the morphology does not have access to any additional linguistic information, it generates all possible paradigms allowed by the word structure.

The paradigm for verbs includes 6 morphologically distinguishable forms– the *infinitive* which is also the form for *present plural* in Dutch, *1st person singular present*, *2nd/3rd person singular present*, *past singular*, *past plural* and *past participle*. The paradigm for the nouns includes the singular and the plural noun forms and the one for adjectives consists also of 6 forms– the *base*, *comparative* and *superlative* forms plus their inflected counterparts. When adjectives are used attributively in Dutch, they get an *-e* suffix. The only exception is when they precede a neutral noun which is not used with a definite article or a pronoun: ***een*** *duur hotel* (an expensive hotel) but ***het*** *dure hotel* (the expensive hotel).

Further, the web is used to validate the non-deterministic output of the finite state morphology. We use Yahoo to search for all forms in a paradigm and we also apply some simple heuristics to determine if there are enough search hits for all forms in a given paradigm. If so, the paradigm is accepted as valid. For nouns, we are also able to determine the correct definite article by comparing the number of occurrences of the singular form with *de* and *het*. Some words have more than one valid paradigm. For example, past participles can often be adjectives in Dutch.

After finding all valid paradigms for a given unknown word, we use the word form(s) (e.g. plural *de* noun, inflected comparative adjective) as additional features in the type predictor. The results are shown in Table 4.

POS		Precision(%)	Recall(%)	F-measure(%)
Nouns	old	85.93	88.49	87.19
	new	91.19	89.89	90.54
Adjectives	old	69.18	79.93	74.17
	new	75.67	84.62	79.70
Verbs	old	62.54	59.05	60.75
	new	66.68	62.86	64.71
Overall	old	79.71	81.58	80.63
	new	85	84.10	84.55

Table 4: ME-based model improved results

The incorporation of additional morphological information significantly improved the performance of the classifier– by 4% in terms of F-measure. Moreover, the number of errors like the one given in (6) and errors due to a wrong noun article is extremely small. The remaining wrong predictions represent the really hard cases, like wrong subcategorization frames, for example, where morphology cannot help.

Since our goal is to improve the performance of the Alpino grammar and parser, we have conducted an experiment with a test set of about 400 sentences which contain 405 unique unknown words, i.e. words that are not listed in the Alpino lexicon. The sentences had already been linguistically annotated in order to provide a gold standard which the performance of Alpino can be compared to. We employed our method to predict lexical types for each unknown word. The unknown words were then added to the lexicon of the grammar and the test sentences were parsed with this 'extended' version of Alpino. In order to have a baseline for comparison, we have also parsed the test sentences with the default Alpino version where the types of the unknown words are guessed by the built-in unknown word guesser, mentioned in Section 2.

The experiments showed that our method has a slightly better performance. The extended Alpino version achieves 87.92% accuracy on the test sentences against 87.5% for the default one. The improvement of 0.42% might look insignifficant but it is a serious step in the right direction. Alpino is already very robust and it is hard to outperform the 87.5% baseline it provides because the remaining 12% represent complicated linguistic phenomena which are challenge for any state-of-the-art deep parser.

4. Abstract Lexical Types

A possible solution for the remaining issues with the prediction process is a cascaded classifier architecture where we develop specific approaches to each problem. For the input to such a classification tool to be more accurate, we first try to predict the POS and the most important linguistic features of the unknown word correctly. What features are considered to be 'most important' depends strictly on the given lexical type. To this extent, we have analysed each of the 641 target types and transformed them into new ones by removing non-crucial features. Subcat frames are always removed.

(8) a. noun(de,count,pl), noun(het,count,pl,measure)⟶noun(pl)
b. verb(hebben,inf,refl), verb(hebben,pl,np_np), verb(zijn,pl,intransitive)⟶verb(inf)

In (8-a) we ignore the fact that the second type designates measure because it is not crucial in this initial stage. Furthermore, both types designate plural nouns, so the nouns are clearly countable and this feature is removed as redundant. Finally, since plural nouns in Dutch always go with the *de* article, the *de/het* distinction is not important either and we preserve only the main fact that both types designate plural nouns.

In (8-b) we ignore the subcat features. Since all 3 types do not designate past participles, it is not important on this stage to know about the auxiliary verb that is used together with the participle to form perfect tense and thus, we ignore the *hebben/zijn* distinction. Finally, since the infinitive verb form is also the plural form in Dutch, we also ignore this distinction and all 3 types are transformed to *verb(inf)*.

By doing such transformations, we end up with a reduced target type inventory of 42 lexical types which we call *abstract lexical types*. Using these types, we perform the lexical acquisition experiment described above. The results are shown in Table 5.

Model	Precision(%)	Recall(%)	F-measure(%)
Naive baseline	19.60	18.77	19.77
POS tagger baseline	74.40	71.12	72.72
Alpino baseline	64.20	61.75	62.95
ME-based model	91.85	91.55	91.70

Table 5: Experiment results with the abstract types

The naive baseline has not changed because *noun(de,count,sg)* is still the majority type in the lexicon. On the other hand, the performance of the other two baseline models has improved due to the reduced number of possible outcome types and the much smaller amount of morphosyntactic information they represent. It should be noted though that even under such 'ideal' conditions, the performance of the POS tagger is still far from state-of-the-art contrary to what would have been expected. Since similar results have been previously shown in (Zhang & Kordoni 2006) and (Cholakov et al. 2008), this clearly disproves a common opinion that a POS tagger could be sufficient for the purposes of sophisticated and linguistically interesting and useful LA.

The ME-based model still shows the best performance. We can rely on the accurate results it provides and use them as an input to task-specific classifiers. In this way we could be pretty sure that we do not send a noun to a classifier designed for guessing the right verb subcategorization frame, for example.

5. Conclusion and Future Research

In this paper, we presented accurate machine learning methods for the problem of LA for large-scale hand-crafted grammars. We applied them to the Dutch Alpino grammar and achieved about 84.5% F-measure which proved their applicability and accuracy. We also underlined the importance of morphology for the prediction process by using word paradigms to incorporate more morphological information into it which has significantly increased its performance.

The fact that the morphological features are specific to Dutch, does not harm the general language and formalism independence of the proposed LA approach. We have illustrated once again that morphology plays a crucial role in the prediction process for languages which exhibit a sufficiently large variety of morphological phenomena. The conclusion we draw is that more morphological information, carefully encoded in the form of suitable features, is very much likely to enhance significantly the LA for such languages. This sort of obvious fact has been often missed in the research.

Though morphology has improved the LA results, it will not help much for the prediction of certain type features. For example, subcategorization frames and other purely syntactic features will remain an obstacle. It is our opinion that the method of interaction between the type predictor and a task specific component (in our case, the finite state morphology) which we presented here is promising and in our future research we would try to develop other task specific components and integrate the type predictor and them into a unified system.

References

P. Barg & M. Walther (1998). 'Processing unknown words in HPSG'. In *Proceedings of the 36th Conference of the ACL*, Montreal, Quebec, Canada.

K. Cholakov, et al. (2008). 'Towards Domain-Independent Deep Linguistic Processing: Ensuring Portability and Re-Usability of Lexicalised Grammars'. In *Proceedings of COLING 2008 Workshop on Grammar Engineering Across Frameworks (GEAF08)*, Manchester, UK.

K. Cholakov & G. van Noord (2009). 'Combining Finite State and Corpus-based Techniques for Unknown Word Prediction'. In *Proceedings of the 7th Recent Advances in Natural Language Processing (RANLP) conference*, Borovets, Bulgaria.

J. Cussens & S. Pulman (2000). 'Incorporating linguistic constraints into inductive logic programming'. In *Proceedings of the Fourth Conference on Computational Natural Language Learning*.

G. Erbach (1990). 'Syntactic processing of unknown words. IWBS report 131'. Tech. rep., IBM, Stuttgart.

F. Fouvry (2003). 'Lexicon acquisition with a large-coverage unification-based grammar'. In *Companion to the 10th Conference of EACL*, pp. 87–90, Budapest, Hungary.

R. Prins & G. van Noord (2001). 'Unsupervised POS-tagging Improves Parcing Accuracy and Parsing Efficiency'. In *Proceedings of IWPT*, Beijing, China.

T. van de Cruys (2006). 'Automatically Extending the Lexicon for Parsing'. In J. Huitnik & S. Katrenko (eds.), *Proceedings of the Eleventh ESSLLI Student Session*, pp. 180–189.

G. van Noord (2006). 'At Last Parsing Is Now Operational'. In *Proceedings of TALN*, Leuven, Belgium.

Y. Zhang & V. Kordoni (2006). 'Automated Deep Lexical Acquisition for Robust Open Text Processing'. In *Proceedings of the Fifth International Conference on Language Resourses and Evaluation (LREC 2006)*, Genoa, Italy.

UNSUPERVISED SYNTAX LEARNING WITH CATEGORIAL GRAMMARS USING INFERENCE RULES

Xuchen Yao, Jianqiang Ma, Sergio Duarte and Çağrı Çöltekin
Center for Language and Cognition, University of Groningen

Abstract

We propose a learning method with categorial grammars using inference rules. The proposed learning method has been tested on an artificial language fragment that contains both ambiguity and recursion. We demonstrate that our learner has successfully converged to the target grammar using a relatively small set of initial assumptions. We also show that our method is successful at one of the celebrated problems of language acquisition literature: learning the English auxiliary order.

1. Introduction

Unsupervised learning of natural language grammar is a challenging task. One of the challenges is learning a finite description, a grammar, of an infinite language using a finite amount of input. Besides, human languages are full of ambiguity, which contributes to the challenge of the learning experience. In this paper we present a computational language learner that successfully learns an artificial grammar exhibiting both challenges. The method is based on learning a categorial grammar in an unsupervised fashion.[1]

Categorial Grammar (CG) is a lexicalized grammar formalism with a high level of transparency between syntax and semantics. These features make CG an attractive formalism for computational studies of language acquisition. The lexicalized nature of the CG reduces learning syntax to learning a lexicon, while the close connection between syntax and semantics helps learning one using the other.

One of the earliest studies of CG learners was proposed by Buszkowski & Penn (1989). Their system used unification of type-schemes to determine categorial grammars from functor-argument structures. Kanazawa (1998) extended this algorithm to learn from strings of words. A number of applied studies (e.g. Waldron 1999, Villavicencio 2002, Buttery 2006) followed similar approaches to learn CG based grammars. Waldron (1999) used a rule-based method to infer a CG from input labeled with basic syntactic types. Villavicencio (2002) proposed a method that improves the performance of Waldron's system by describing an unconventional universal grammar based on CG, and using semantically annotated input. Watkinson & Manandhar (2000) presented an unsupervised stochastic learner which aims to learn a compact lexicon. They assumed that the set of possible categories are known, which maps the problem of grammar induction to categorization. The system achieved perfect accuracy in an artificial corpus. However, its performance dropped to 73.2% in lexicon accuracy and 28.5% in parsing accuracy when tested on the more realistic LLL corpus (Kazakov, et al. 1998).

This paper proposes an unsupervised method to learn categorial grammars. The learner is provided with a set of positive sentences generated by a target grammar. Unknown categories are learned by applying a set of inference rules incrementally. When

[1] We use the term 'unsupervised' in the sense that the learner is not provided with the information about the structure of the input sentences. Although non-standard, this use of the term is common in computational linguistics literature (e.g. Klein 2005, Watkinson & Manandhar 2000)

there are multiple choices, a *simple category preference* (SCP) principle that is inspired by the MDL principle (Rissanen 1989) is used to minimize the size of the grammar. We intend to develop this algorithm further to learn from real language corpora. However, in this paper we show that the learner is able to infer a recursive and ambiguous artificial grammar and learn the English auxiliary word order from a set of input sentences that are considered insufficient for the task.

The next section gives a short introduction to CG. Section 3 describes our learning architecture. Section 4 presents two experiments and discussion of the results together with limitations of our approach. In the last section we provide brief conclusions and address future directions.

2. Categorial Grammar

Categorial grammar (Ajdukiewicz 1935, Bar-Hillel 1953) is a lexicalized grammar formalism. CG describes all the language specific syntactic information inside the lexicon, leaving only a small number of universal rules outside the lexicon. We present a very brief introduction to CG here, more comprehensive description can be found in Wood (1993).

Every word in a CG lexicon is assigned to a syntactic category. A limited set of categories constitutes the *basic categories* of the grammar. For example, *S* (sentence), *NP* (noun phrase), *N* (noun) are commonly assumed to be the basic categories for English. *Complex categories*, such as *NP/N*, $S\backslash NP$, $(S\backslash NP)\backslash(S\backslash NP)$, are formed by combining any two CG categories with a forward (/) or backward (\) slash. Given the lexicon with categories of this form, the only rules of the CG are given in (1).

(1) *Function application rules*

Forward application	A/B	B	$\rightarrow$	A	$(>)$
Backward application	B	$A\backslash B$	$\rightarrow$	A	$(<)$

CG as described above is weakly equivalent to Context Free Grammars, and cannot model the complexity of natural languages adequately. However, there are extensions such as *Combinatory Categorial Grammar* (CCG, Steedman 2000) that provide necessary descriptive and theoretical adequacy by introducing additional operations. In this work, we learn classical Categorial Grammars, while making use of some of the CCG operations in (2), namely *composition* and *type raising*, during the learning process.

(2) a. *Function composition rules:*

Forward	A/B	B/C	$\rightarrow$	A/C	$(>\mathbf{B})$
Backward	$B\backslash C$	$A\backslash B$	$\rightarrow$	$A\backslash C$	$(<\mathbf{B})$

b. *Type raising rules:*

Forward	A	$\rightarrow$	$T/(T\backslash A)$	$(>\mathbf{T})$
Backward	A	$\rightarrow$	$T\backslash(T/A)$	$(<\mathbf{T})$

3. Learning by Inference Rules

In this section we first introduce a series of inference rules used to perform grammar induction. Then we will present the complete learning architecture along with an example demonstrating the learning process.

3.1. Grammar Induction by Inference Rules

Our inference rules work when there is only one unknown category in the input. In the rule descriptions below, the letters *A, B, C* and *D* represent known categories, **X** represents the unknown category.

(3) *Level 0 inference rules:*

$$B/A \quad \mathbf{X} \quad \rightarrow \quad B \quad \Rightarrow \quad \mathbf{X} = A \quad if A \neq S$$
$$\mathbf{X} \quad B\backslash A \quad \rightarrow \quad B \quad \Rightarrow \quad \mathbf{X} = A \quad if A \neq S$$

(4) *Level 1 inference rules:*

$$A \quad \mathbf{X} \quad \rightarrow \quad B \quad \Rightarrow \quad \mathbf{X} = B\backslash A \quad if A \neq S$$
$$\mathbf{X} \quad A \quad \rightarrow \quad B \quad \Rightarrow \quad \mathbf{X} = B/A \quad if A \neq S$$

We define *level* as the number of *functioning* slash operators in a category. Functioning slash operators are functors that take an argument of one type and result in another during the derivation. Consequently, the basic categories are of level 0. The category $S\backslash NP$ belongs to level 1. Note that the category of adverbs $(S\backslash_f NP)\backslash_f(S\backslash NP)$ belongs to level 2. Although it has three slashes, only the slashes marked with subscript $_f$ are functioning, i.e. can be used in a derivation.

Level 0 and level 1 inference rules can be successfully used to learn the category of intransitive verbs, such as *slept* in *Peter slept*. The condition $if \quad A \neq S$ in (3) and (4), prevents learning a large number of incorrect categories.[2] For example, $S\backslash S$ for the word *well* from *Peter slept well*. As stated before, the category of adverbs belongs to level 2, so we need a level 2 inference rule to learn this category.

(5) a. *Level 2 side inference rules:*

$$\mathbf{X} \quad A \quad B \quad \rightarrow \quad C \quad \Rightarrow \quad \mathbf{X} = (C/B)/A$$
$$A \quad B \quad \mathbf{X} \quad \rightarrow \quad C \quad \Rightarrow \quad \mathbf{X} = (C\backslash A)\backslash B$$

b. *Level 2 middle inference rule:*

$$A \quad \mathbf{X} \quad B \quad \rightarrow \quad C \quad \Rightarrow \quad \mathbf{X} = (C\backslash A)/B$$

Level 2 inference rules are divided into two parts: the *side rule* and the *middle rule*, depending on whether an unknown category is at the beginning/end of a sentence or in the middle.

Notice that in (5b) the category $(C/B)\backslash A$ is as viable as the inferred category $(C\backslash A)/B$. This can be shown by the following example of *left-combining* rule and *right-combining* rule.

(6) a. left-combining *rule:*

$$A \quad \mathbf{X} \quad B \rightarrow C \quad \xrightarrow[divide\ A\mathbf{X}B]{left-combining} \quad (A\mathbf{X}) \quad B \rightarrow C \quad \xrightarrow[rule(4)]{level\,1} \quad A\mathbf{X} = C/B \quad \xrightarrow{divide\ A\mathbf{X}} \quad A \quad \mathbf{X} \rightarrow C/B \quad \xrightarrow[rule(4)]{level\,1} \quad \mathbf{X} = (C/B)\backslash A$$

b. right-combining *rule:*

$$A \quad \mathbf{X} \quad B \rightarrow C \quad \xrightarrow[divide\ A\mathbf{X}B]{right-combining} \quad A \quad (\mathbf{X}B) \rightarrow C \quad \xrightarrow[rule(4)]{level\,1} \quad \mathbf{X}B = C\backslash A \quad \xrightarrow{divide\,\mathbf{X}B} \quad \mathbf{X} \quad B \rightarrow C\backslash A \quad \xrightarrow[rule(4)]{level\,1} \quad \mathbf{X} = (C\backslash A)/B$$

---

[2]This is against the real world situation since the category $S\backslash S$ is a valid CG category. The problem can be possibly solved by putting S in a less favorite position in the simple category preference principle in later work.

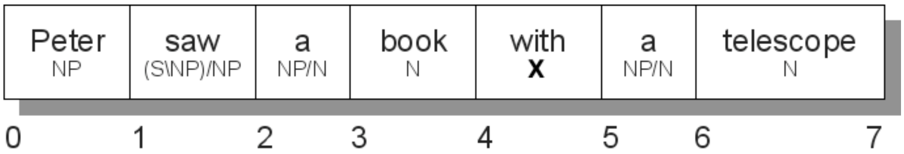


Figure 1: Index in the input string

As shown in the above example, left-combining and right-combining rules produce different but equivalent outputs. Our algorithm uses the right-combining rule when recursively dividing all the entities into two parts and whenever there are possibilities to combine an unknown category with either the left one or the right one, we always combine with the right one.[3]

It might seem that using (5b) we can learn the category of $(S\backslash S)/NP$ for the preposition *with* from the sentence *Peter slept with Mary*. But this will not happen: the level 2 inference rule is implemented by recursively calling level 0 and level 1 inference rules, all of which have the condition $if \quad A \neq S$ to prevent generating the category $S\backslash S$. As a matter of fact, none of the level 0-2 rules could help learning the category of *with* from the sentence *Peter slept with Mary*. So we need to use a level 3 inference rule.

(7) a. *Level 3 side inference rules:*

$$\mathbf{X} \ A \ B \ C \ \rightarrow \ D \ \Rightarrow \ \mathbf{X} = ((D/C)/B)/A$$
$$A \ B \ C \ \mathbf{X} \ \rightarrow \ D \ \Rightarrow \ \mathbf{X} = ((D\backslash A)\backslash B)\backslash C$$

b. *Level 3 middle inference rules:*

$$A \ \mathbf{X} \ B \ C \ \rightarrow \ D \ \Rightarrow \ \mathbf{X} = ((D\backslash A)/C)/B$$
$$A \ B \ \mathbf{X} \ C \ \rightarrow \ D \ \Rightarrow \ \mathbf{X} = ((D\backslash A)\backslash B)/C$$

3.2. The Learning Architecture

The learning framework consists of three parts: *the edge generator*, *the recursive learner* and *the output selector*. A schematic description of the learning process is provided in Figure 2. Below we provide a detailed description of the three parts, along with demonstration of learning the ambiguous and recursive categories of *with* in Figure 1.

The Edge Generator implements a variation of the CYK algorithm, which employs bottom-up chart parsing. Every known word in a sentence is an edge in the chart. The edge generator then tries to merge any consecutive edges into a single edge recursively. In order to produce as many edges as possible, besides function application rules (>,<), we have also used the composition ($> B$,$< B$) and the type raising ($> T$,$< T$) rules. Table 1 shows all possible edges generated for the example in Figure 1.

The Recursive Learner performs grammar induction by the rules given in Subsection 3.1. The learning process first tries to learn from level 0 or level 1 inference rules. If the unknown word cannot be learned by level 0 or level 1 inference rules, higher level rules are tried by recursively dividing all the edges in a sentence into two parts and then calling level 0 or level 1 inference rules to learn (This process is also shown in (6)). Following the *simple category preference* (SCP) principle, if a category can be inferred with a lower level rule, we do not attempt to use higher level rules.

[3]We realize that this rule may lead to wrong choices for other languages, and plan to relax it in future work.

	span	rule used	category		span	rule used	category
1	(0, 1)	>T	S/(S\NP)	6	(0, 3)	>B	S/N
2	(0, 2)	>B	S/NP	7	(1, 4)	>	S\NP
3	(1, 3)	>B	(S\NP)/N	8	(2, 4)	<T	S/(S\NP)
4	(2, 4)	>	NP	9	(0, 4)	<	S
5	(5, 7)	>	NP	10	(0, 4)	>	S

Table 1: Generated edges in a chart

	A		B		**X**		C	
	cat	span	cat	span	cat	span	cat	span
1	NP	(0, 1)	S\NP	(1, 4)	**((S\NP)\(S\NP))/NP**	(4,5)	NP	(5, 7)
2	S/(S\NP)	(0, 1)	S\NP	(1, 4)	((S\NP)\(S\NP))/NP	(4,5)	NP	(5, 7)
3	S/NP	(0, 2)	NP	(2, 4)	**(NP\NP)/NP**	(4,5)	NP	(5, 7)
4	S/NP	(0, 2)	S/(S\NP)	(2, 4)	(NP\(S/(S\NP)))/NP	(4,5)	NP	(5, 7)
5	S/N	(0, 3)	N	(3, 4)	**(N\N)/NP**	(4,5)	NP	(5, 7)

Table 2: Categories learned from the rule $A \quad B \quad \mathbf{X} \quad C \rightarrow S$
for the sentence in Figure 1.

For the input in Figure 1, the level 0 and level 1 inference rules are not enough. Only the level 3 middle inference rules (7b) can be applied. Table 2 gives a list of all the possible categories using this inference rule.

The Output Selector tests the learned categories produced by the recursive learner and selects the ones that can be parsed using only function application rules. The categories that do not produce a valid parse with function application rules are discarded.

In Table 2, the sentence cannot be parsed using the category in row 4, so this category is discarded. Rows 1 (or equal category in row 2), 3 and 5 provide the learned categories.

4. Experiments and Results

We conducted two experiments with our learning system. In the first experiment, we tested the system's capabilities on an artificial language exhibiting a certain level of ambiguity and recursion. In the second experiment, we tried to learn the English auxiliary order, a well known problem in language acquisition literature.

4.1. Experiment 1: Learning an Artificial Grammar

For this experiment, we have created a small English-like artificial grammar. The lexicalized grammar that is used as the target grammar for this experiment is listed in Table 3. The input to the learner consists of 160 sentences (2 to 7 words in length) generated by the target grammar. Only correct sentences are used. The input sentences are unlabeled, except for nouns (N) and noun phrases (NP). Thus the learner first searches sentences with only one unknown word and tries to learn this word. Then it takes into account the learned category and searches for other sentences with unknown words. Using this

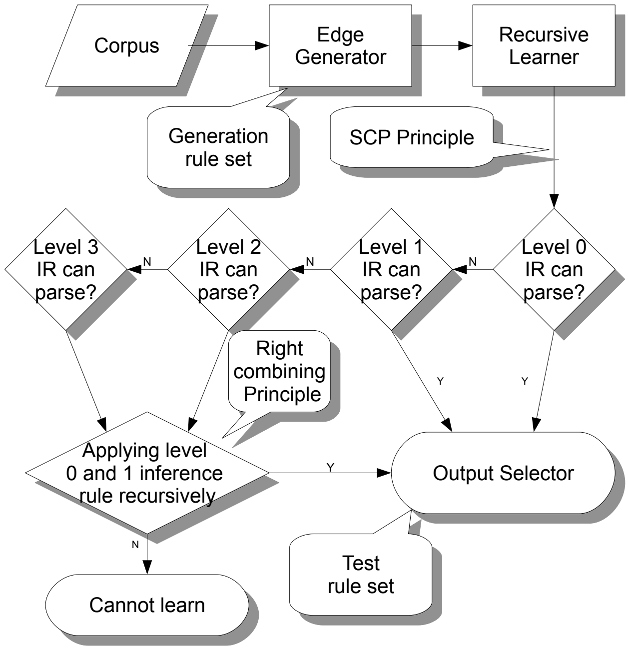


Figure 2: Learning process using inference rules

Peter	:= NP	*Mary*	:= NP	*with*	:= (N\N)/NP
book	:= N	*green*	:= N/N	*with*	:= ((S\NP)\(S\NP))/NP
colorless	:= N/N	*sleep*	:= S\NP	*furiously*	:= (S\NP)\(S\NP)
a	:= NP/N	*telescope*	:= N	*give*	:= ((S\NP)/NP)/NP
the	:= NP/N	*saw*	:= (S\NP)/NP	*read*	:= (S\NP)/NP
run	:= S\NP				

Table 3: Target grammar rules

"bootstrap"-like method the learner is expected to converge to the target grammar.

After only a single pass through input sentences, all categories in our target grammar presented in Table 3 are learned correctly. The learned grammar includes only one lexical item (*with* := $(NP\backslash NP)/NP$) that is not in the target grammar. This, however, is a useful generalization which allows deriving structures like *[Peter [saw [Mary [with [a telescope]]]]]*, while our original grammar does not.

4.2. Experiment 2: Learning Correct Word Order

The difficulty of learning English auxiliary order has also been used as a support for the poverty of the stimulus (POTS) argument, and hence for linguistic nativism. Introduced first by Kimball (1973), the problem can be summarized as follows: the English auxiliary verbs *should*, *have* and *be* occur exactly in this order and all of them are optional. The claim is that while sentences containing a single auxiliary (8a–8c) or two auxiliaries (8d–8f) are present in the input, sequences of three auxiliaries (8g) are not frequent enough. Hence, it is not possible to learn the correct three-auxiliary sequence from the input alone.

$$
\begin{array}{ll}
\textit{should} & := (S_s\backslash NP)/(S\backslash NP) \\
\textit{should} & := (S_s\backslash NP)/(S_h\backslash NP) \\
\textit{should} & := (S_s\backslash NP)/(S_b\backslash NP) \\
\textit{have} & := (S_h\backslash NP)/(S\backslash NP) \\
\textit{have} & := (S_h\backslash NP)/(S_b\backslash NP) \\
\textit{be} & := (S_b\backslash NP)/(S\backslash NP)
\end{array}
$$

Table 4: Categories of some auxiliary verbs.

(8) a. *I should go.*
b. *I have gone.*
c. *I am going.*
d. *I have been going.*
e. *I should have gone.*
f. *I should be going.*
g. *I should have been going.*
h. **I have should been going.*

The argument is controversial, and Pullum & Scholz (2002) have shown that there are more three-auxiliary sequences in children's input than claimed. In this study, we choose another approach: we present our learner with sentences containing only one or two auxiliaries (as in (8a-8f)), and we test if it can correctly recognize and generate sentences with three auxiliaries. The experiment setting is the same as in experiment 1. The only additional information provided is the type of sentences, i.e. every given input is marked with the "mood of the sentence". As well as simple declarative sentences (S), we used S_b, S_h and S_s for sentences with modal verbs *be*, *have* and *should* respectively.

Table 4 presents a fragment of the learned grammar. The derivation of the sentence (8g) using the learned grammar is given in Figure 3. As can be verified easily, the lexicalized grammar presented in Table 4 would not allow sequences as in (8h). The categories assigned to auxiliary verbs by the learner completely, and correctly derive the English auxiliary order.[4]

Success of the learner is again due to its assignment of words to syntactic categories. The categories induced from one- and two-auxiliary sequences in a logical way extend naturally to three-auxiliary sequences.

4.3. Discussion

We have presented a learner that learns syntax using CG. One of the characteristics of our method is that it learns from input without any structure, semantic annotation or negative evidence. Although there are theoretical results about learnability on only strings (Kanazawa 1998), and more applicable research about learning from sentences annotated with structures (Villavicencio 2002, Buttery 2006), applied work on learning from strings

[4]An alternative approach would be assuming that the sequences like '*should have been*' are learned as single units, at least at the beginning of the learning process. Lexical items spanning multiple input units are considered by some of the related learners (e.g. Zettlemoyer & Collins 2005, Çöltekin & Bozsahin 2007). However, to be compatible with the original claim, the learner presented in this paper assigns categories to single input units.

$$
\begin{array}{ccccc}
\text{I} & \text{should} & \text{have} & \text{been} & \text{going.} \\
\overline{NP} & \overline{(S_s\backslash NP)/(S_h\backslash NP)} & \overline{(S_h\backslash NP)/(S_b\backslash NP)} & \overline{(S_b\backslash NP)/(S\backslash NP)} & \overline{S\backslash NP} \\
 & & & \multicolumn{2}{c}{\underline{\hspace{6em}}{>}} \\
 & & & \multicolumn{2}{c}{S_b\backslash NP} \\
 & & \multicolumn{3}{c}{\underline{\hspace{12em}}{>}} \\
 & & \multicolumn{3}{c}{S_h\backslash NP} \\
 & \multicolumn{4}{c}{\underline{\hspace{18em}}{>}} \\
 & \multicolumn{4}{c}{S_s\backslash NP} \\
\multicolumn{5}{c}{\underline{\hspace{22em}}{<}} \\
\multicolumn{5}{c}{S_s}
\end{array}
$$

Figure 3: Derivation of the correct word order.

is rather limited. This study is our first attempt to fill this gap. Although the experiments are based on artificial data, our aim is to further develop the method and apply it on real-world linguistic input.

The simple inference rules used here are admittedly ad-hoc and we have not yet attempted to provide the guarantee of convergence. Our main goal with this method is to experiment with the possibilities of exploiting the information in the linguistic input, rather than to find a learning algorithm that is guaranteed to learn in a wide range of input distributions. For the fragment of the English grammar captured by our artificial language learning experiments, the results are promising.

The method is in essence similar to unification based learner of Buszkowski & Penn (1989), which learns from structurally annotated input. Unfortunately, Kanazawa's extension of the algorithm to learn from strings is computationally intractable. The use of ad-hoc rules and constraints instead of standard learning framework is motivated by the aim of using of a reasonable amount of input and computational resources.

The input to our learner is partially annotated. This approach carries an affinity to the *partial learning* system described by Moreau (2004). However, crucially, the annotation provided to our learner does not contain any structure. Moreau (2004) especially makes use of high-frequency closed-class words with categories that give hints about the structure of the input sentences. This is useful in practical computational linguistic applications, as it helps inducing a grammar with relatively small amount of annotation. However, our approach is more plausible for language acquisition, as children are known to learn nouns earlier than other word classes (Gentner 1982).

Another apparent limitation of our learner is that it only learns from the input that contains only one unknown word. This avoids the combinatorial expansion of hypothesized categories and keeps the required computational resources low, and this worked fine for our artificial grammar learning experiments.[5] For language acquisition, this is not a wildly wrong assumption. We assume that the children do not understand and, hence, make use of complicated input at first sight. However, the category of the word can still be inferred, when the same word later appears in an understandable context.

The output selector only selects the categories that can be parsed by the AB grammars. This could lead to inconsistency with the target grammar: if the target grammar can only be parsed by more complicated CCG rules while the output selector only uses AB rules, or if the target grammar can be parsed by AB rules while the output selector uses CCG

[5] It should also be noted that, the algorithm can be adapted to use the input with '*k* unknown words' with the expense of additional computational resources.

rules. Here we do not think there will be big problems: the initial setting is to use AB rules, when there are no candidates under AB rules, the output selector adjusts to more rules to produce an output. The simple category preference guarantees that no matter how complex rules are used, the output category will stay as simple as possible.

The inference rules are simple and even intuitive. Although there is no solid psychological evidence that children learn a grammar in an inference-rule-based way, the 100% correct results in parsing and generation by our model suggests that it is sufficient to assume that children use a small set of rules together with a plausible inference procedure for learning the categories of unknown words. The only other additional piece in our learning algorithm is a preference towards simpler categories.

This method performs well on learning a typical language phenomenon such as learning English auxiliary order. We have shown that only being exposed to one- and two-auxiliary sequences, our simple algorithm generalized correctly to sentences containing three auxiliary verbs. Even if the POTS claims are correct, children can still learn correct forms with simple inference mechanisms.

5. Conclusion

We described a method to learn categorial grammars using inference rules. Our method has learned all the categories of the target grammar. We use simple logical and intuitive inference rules to solve the problem of unknown categories in the input. The only additional aid provided to our learner is the simple category preference. Using only this set of initial assumptions, our system is also able to learn a phenomenon that has been considered difficult. The learner is able to infer English auxiliary order correctly without being presented with all possible sequences.

However, it is necessary to note that our system has a number of limitations. First, these results were obtained using data that was generated artificially. Second, since we do not use any statistical inference mechanism, our system is not robust against noise. Using statistical patterns in the input language, it may also be possible to relax some of the assumptions presented here. This is a limitation when the amount of data is not large enough to "bootstrap" the learner.

Future work includes developing the algorithm further and evaluating it on real data, such as child-directed speech from CHILDES database (MacWhinney 2000).

References

K. Ajdukiewicz (1935). 'Die syntaktische Konnexität'. *Studia Philosophica* **1**:1–27.

Y. Bar-Hillel (1953). 'A Quasi-Arithmetical Notation for Syntactic Description'. *Language* **1**:47–58.

W. Buszkowski & G. Penn (1989). 'Categorial grammars determined from linguistic data by unification'. Tech. rep., Chicago, IL, USA.

P. J. Buttery (2006). 'Computational models for first language acquisition'. Tech. rep., University of Cambridge, Churchill College.

Ç. Çöltekin & C. Bozsahin (2007). 'Syllables, Morphemes and Bayesian Computational Models of Acquiring a Word Grammar'. In *Proceedings of 29th Annual Meeting of Cognitive Science Society*, Nashville.

D. Gentner (1982). 'Why nouns are learned before verbs: Linguistic relativity versus natural partitioning.'. In S. Kuczaj (ed.), *Language development*, vol. 2. Erlbaum, Hillsdale, NJ.

M. Kanazawa (1998). *Learnable classes of categorial grammars*. Cambridge University Press, New York, NY, USA.

D. Kazakov, et al. (1998). 'The FraCas dataset and the LLL challenge'. Tech. rep., SRI International.

J. P. Kimball (1973). *The Formal Theory of Grammar*. Englewood Cliffs, NJ: Prentice-Hall.

D. Klein (2005). *The Unsupervised Learning of Natural Language Structure*. Ph.D. thesis, Stanford University.

B. MacWhinney (2000). *The CHILDES project: tools for analyzing talk*. Lawrence Erlbaum, Mahwah, NJ u.a. EN.

E. Moreau (2004). 'Partial Learning Using Link Grammars Data'. In G. Paliouras & Y. Sakakibara (eds.), *Grammatical Inference: Algorithms and applications. 7th International Colloquium: ICGI 2004*, vol. 3264 of *Lectures Notes in Artificial Intelligence*, pp. 211–222, Athens, Greece. Springer.

G. K. Pullum & B. C. Scholz (2002). 'Empirical assessment of stimulus poverty arguments'. *The Linguistic Review* **19**:9–50.

J. Rissanen (1989). *Stochastic Complexity in Statistical Inquiry Theory*. World Scientific Publishing Co., Inc., River Edge, NJ, USA.

M. Steedman (2000). *The syntactic process*. MIT Press, Cambridge, MA, USA.

A. Villavicencio (2002). *The acquisition of a unification-based generalised categorial grammar*. Ph.D. thesis, University of Cambridge.

B. Waldron (1999). 'Learning grammar from corpora'. Master's thesis, Cambridge University.

S. Watkinson & S. Manandhar (2000). 'Unsupervised lexical learning with Categorial Grammars using the LLL corpus'. In *Learning language in logic*, pp. 218–233. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

M. M. Wood (1993). *Categorial Grammars*. Routledge, London.

L. S. Zettlemoyer & M. Collins (2005). 'Learning to Map Sentences to Logical Form: Structured Classification with Probabilistic Categorial Grammars'. In *Proceedings of the Twenty First Conference on Uncertainty in Artificial Intelligence (UAI-05)*.

EXTENDED QUALIA-BASED LEXICAL KNOWLEDGE FOR DISAMBIGUATION OF JAPANESE POSTPOSITION *No*

Sumiyo Nishiguchi
Department of Language and Culture, Osaka University

Abstract. This paper proposes the elaboration of the qualia structure of the Generative Lexicon in Pustejovsky (1995) and the Extended Generative Lexicon theory (Lenci et al., 2000). My proposal is based on the Japanese genitive postposition *no*. The Japanese "NP_1-no NP_2" construction expresses a wider range of relations between two entities than does the English possessive "NP_1*'s* NP_2," such that the Pustejovskian qualia roles encoded in NP_2 do not supply the necessary relations between two entities, which Vikner and Jensen (2002) succeeded to certain degree. The disambiguation of possessive relations requires that lexical entries be augmented by incorporating a referential module comprising subcategories such as LOCATION, TIME, and MANNER.

1. Different Types of Relations and Argument Reversal

The Japanese genitive marker is semantically very ambiguous. "NP1-GEN NP2" not only expresses possession as in *Naomi's bag* and inalienable relations as in *Naomi's face*, but also aspects such as location, accompaniment, property, and quantity, as presented in Table 1.

Relation	**Japanese Possessive**	**English Possessive**	**English Compound**	**English Prepositional Phrase**
I possession	Naomi-no kaban	Naomi's bag	*Naomi bag	a bag of Naomi
II part-whole	Naomi-no kao	Naomi's face	*Naomi face	the face of Naomi
III location	Tokyo-no shinseki	*Tokyo's relative	*Tokyo relative	a relative (living) in Tokyo
IV time	yugata-no koen natsu-no kyuka 7-ji-no nyusu	*evening's park *summer's vacation *7 o'clock's news	evening park summer vacation 7 o'clock news	a park in the evening vacation in summer the news at 7 o'clock
V accompaniment	kaban-no hito boshi-no fujin	*bag's man *hat's lady	the bag man hat lady	the man with a bag the lady with a hat
VI trade	Kaban-no Kochi	*Bags' Coach	Bags Coach	Coach for bags
VII activity	maaruboro-no kuni biiru-no machi	*Marlboro's country *the beer's city	Marlboro country *the beer city	the country of Marlboro the city of beer
VIII property	chisee-no hito osu-no tora aoi-me-no ningyo tsutsuji-no koen	*intelligence's man *male's tiger *blue eyes' doll *azaleas' park	*intelligence man a male tiger blue eyes doll azalea park	a man of intelligence a tiger of the male kind the doll with blue eyes a park with azaleas
IX weight	1-kiro-no pasokon	*1 kg's computer	1 kg computer	*the computer of 1kg
X quantity	3-bon-no pen	*three's pen	three pens	
XI intensional property	nise-no fukahire nise-no keisatsukan	*fake's shark fin *fake's police officer	fake shark fin fake police officer	

Table 1: **Ambiguity of Japanese Postposition *No***

Note the reversal of the possessor argument between (I) and (V–VI). The possessor argument is NP_1 in (I), as in English *Naomi's bag* whose possessor argument is *Naomi*. On the contrary in (V), the possessor of the bag is NP_2 *hito* "man" and there is no English equivalent *big bag's person*. In (VI) *Kaban-no Kochi* "Bags Coach," *Coach* is a store, and therefore the possessor of a bag. The controller-controllee relation is also reversed, for example, in *Naomi-no kuruma* "Naomi's car" (type I), *Naomi* is the controller of the car, i.e., NP_2 the car is at *Naomi*'s disposal as in English *the girl's car* (Vikner and Jensen, 2002). On the contrary, in *boshi-no fujin* "the lady with a hat," NP_1 *boshi* is at the person's disposal. *Aoi-me-no ningyo* "the doll with blue eyes," literally, "blue eyes' doll" in (VIII) even expresses the part-whole relation in the reverse direction, compared with *ningyo-no me* "the doll's eyes."

Such non-canonical relations, i.e., other than those expressing possession or a part-whole relation, are more likely expressed in noun compounds such as *magic land* or prepositional phrases using *of*, *in*, or *with* in English (Teramura, 1980,234)(Makishita, 1984,193).

Table 2 presents examples taken from *Balanced Corpus of Contemporary Written Japanese*; here, the Japanese postposition *no* is translated into *in* in English since $NP_1 - no$ expresses location, time, and property.

Relation	**Japanese Possessive**	**English Possessive**	**English Compound**	**English Prepositional Phrase**
III location	Osuro kogai-no mura Hachioji-shi-no borantia guruupu	*Oslo suburb's village Hachioji city's volunteer group	*Oslo suburb village Hachioji city volunteer group	a village in the suburb of Oslo a volunteer group in Hachioji city
IV time	katsute-no ikioi manatsu-no hyozan natsu-no kaidan-jiki	*past's force summer peak's iceberg *summer's horror season	past force ?summer peak iceberg ?summer horror season	force in the past/ former influence iceberg in the peak of summer ?horror season in summer
VIII property	jutai-no Shakuruton	*serious condition's Shackleton	*serious condition Shackleton	Shackleton in serious condition
X quantity	9-nin-no esukimo	*nine's Eskimos	nine Eskimos	*Eskimos in nine

Table 2: **Data Translated from *Balanced Corpus of Contemporary Written Japanese, 2008 edition* by The National Institute of Japanese Language**

2. Problems with Deriving Different Types of Possessive Relations from NP_2

Possessive relations are ambiguous in both English and Japanese. For example, there is more than one interpretation for *John-no hon* "John's book." *John's book* may refer to the book that *John* owns or the book that *John* wrote (Barker, 1995,87).

In view of such ambiguity, Partee (1997) assumes two syntactic types for *John's* depending on whether or not the following noun is inherently relational. If the following noun is a non-relational common noun (CN) such as *car*, *John's* composes with *car* which is regular (et) type, and the relation between *John* and *car* is contextually supplied (1a). On the contrary, when *John* is followed by inherently relational nouns such as *brother*, *employee* and *enemy*, which are (e,et) type with an extra argument slot, the relation between John and his brother in *John's brother* inherits kinship from the two-place predicate *brother* (1b). (2) exemplifies the computation related to another relational noun, *friend*.

(1) a. Free R type:

Syntax: [John's]$_{NP/CN}$

Semantics: $\lambda Q\lambda P[John'(\lambda z[\exists x[\forall y[[Q(y) \wedge R(y)(z)] \leftrightarrow y = x] \wedge P(x)]])]$

b. Inherent relation type:

Syntax: [John's]$_{NP/TCN}$ (TCN: transitive common noun)

Semantics: $\lambda R\lambda P[John'(\lambda z[\exists x[\forall y[R(z)(y) \leftrightarrow y = x] \wedge P(x)]])]$

(2) Syntax: [[John's]$_{NP/TCN}$[friend]$_{TCN}$]$_{NP}$

Semantics: $\lambda R\lambda P[John'(\lambda z.\exists x[\forall y[R(z)(y) \leftrightarrow y = x] \wedge P(x)]](friend - of')$ = $\lambda P[John'(\lambda z.\exists x[\forall y[friend - of'(z)(y) \leftrightarrow y = x] \wedge P(x)]]$

If we apply Partee's theory to Japanese examples, most of the possessive relations are unpredictable, and the contextually supplied relation R remains largely ambiguous. Possession relation (I) is prototypical, and part-whole relation (II) can be derived lexically from a possessive *kao* "face" (Barker, 1995). However, other possessee nominals are not necessarily relational.

In order to reduce the cost of pragmatics, Vikner and Jensen (2002) apply the Qualia Structure (Pustejovsky, 1995) of the possessee noun and type-shift even non-inherently relational NP$_2$ into a relational noun. For example, even though *poem* is not a relational noun, *John's poem* can be interpreted as the *poem* that John composed because the internal semantic structure of *poem* contains an author-of relation as AGENTIVE role. The meaning shifting operator Q$_A$ raises a one-place holder *poem* in (3a) into a two-place holder in (3b). The type-shifted NP$_2$ can now combine with the possessive NP, which has a uniformly type ((e,et),(et,t)) so that the authorship relation is inherited from NP$_2$ *poem*, and R is no longer a free variable.

(3) Q$_A$(poem) = $\lambda x\lambda y[poem'(x) \wedge compose'(x)(y)]$

However, even Vikner and Jensen (2002)'s method is not sufficient to systematically compute the meaning of the Japanese 'NP$_1$-no NP$_2$' construction. For example, in terms of location (III), *shinseki* "relative" in *Tokyo-no shinseki* "a relative living in Tokyo" is a relational noun, i.e., *relative-of x*, but the relation between NP$_1$ and NP$_2$ is not *relative of* but of location, namely, NP$_2$ is in NP$_1$. We also encounter a problem with *boshi-no fujin* "the lady with a hat." Since wearing a hat is not part of the qualia roles, that are AGENTIVE (origin), TELIC (purpose), CONSTITUTIVE (part-whole) and FORMAL (isa) roles, of the non-inherently relational noun *fujin*"lady," even Vikner and Jensen's system is unable to supply the binder for R.

3. Selective Binding

Pustejovsky (1995) proposes *selective binding* when computing the meaning of the noun phrases modified by non-intersective adjectives. For example, *fast* in *a fast typist* does not denote a typist who is also generally fast in activities apart from typing, but specifically a typist who is fast at typing. In other words, *fast* does not modify the physical presence of the typist but it does modify the way the typist types; in other words, *fast* modifies the event argument of the TELIC quale of the noun *typist*—to type.

(4) a. a fast typist

b. $[\![fast\_typist]\!]$ = λx[typist'(x) ∧... [TELIC = λe[type'(e) ∧ agent(e) = x ∧ fast'(e)]]...]

This section examines some of the aforementioned cases in which $NP_1 - no$ phrases modify the event argument of the qualia of NP_2. Specifically, the postpositional phrases denoting time, trade, activity, location and outstanding property selectively bind events contained in the TELIC and AGENTIVE qualia. However, there remain many examples in which selective binding does not apply as discussed in section 4.

3.1. TELIC Quale Modification

3.1.1. Time

Some of the $NP_1 - no$ phrases are temporal modifiers, such as *7-ji-no nyusu*, or "7 o'clock news." The purpose, or the TELIC role, of *news* is to describe current events or information; therefore, *7-ji-no* or "7 o'clock's" modifies the TELIC role of *nyusu* or "news" so that the TELIC role of the *7-ji-no nyusu* or "7 o'clock news" is to describe the events taking place at 7 o'clock.

(5) 7-ji no nyusu "7 o'clock news'

$$\begin{bmatrix} \textbf{NYUSU “NEWS”} & \\ \text{TYPESTR} = & \left[\text{ARG1} = \boxed{x}\text{MEDIA\_INFORMATION}\right] \\ \text{EVENTSTR} = & \left[\text{D-E1} = \boxed{e1}\text{PROCESS}\right] \\ \text{ARGSTR} = & \left[\text{D-ARG1} = \boxed{y}\text{OBJ}\right] \\ \text{QUALIA} = & \left[\text{TELIC} = \text{DESCRIBE}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right)\right] \end{bmatrix}$$

(6) $[\![7\_o'clock\_news]\!]$ = λx[news'(x) ∧ [...TELIC = λe[describe'(e, x, p) ∧ at-seven'(e)]...]

$$\begin{bmatrix} \textbf{7-JI-NO NYUSU “7 O'CLOCK NEWS”} & \\ \text{TYPESTR} = & \left[\text{ARG1} = \boxed{x}\text{MEDIA\_INFORMATION}\right] \\ \text{EVENTSTR} = & \left[\text{D-E1} = \boxed{e1}\text{PROCESS}\right] \\ \text{ARGSTR} = & \left[\text{D-ARG1} = \boxed{y}\text{INFO}\right] \\ \text{QUALIA} = & \left[\text{TELIC} = \text{DESCRIBE}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right) \wedge \text{AT-SEVEN}\left(\boxed{e1}\right)\right] \end{bmatrix}$$

3.1.2. Trade and Activity

Genitive phrases that represent trade and activity of the referent of NP_2 in Table 1 are considered to be modifiers of the event contained in the TELIC role of the NP_2. For example, in *biiru-no machi* "a city of beer" and *Kaban-no Kochi* "Bags Coach," beer and bags comprise the theme of the event in the TELIC role, i.e., the making act and the selling act, respectively.

(7) biiru-no machi "a city of beer"

$$\begin{bmatrix}\textbf{MACHI “TOWN”} & \\ \text{TYPESTR} = & \begin{bmatrix}\text{ARG1} = \boxed{x}\text{LOCATION}\end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix}\text{D-ARG1} = \boxed{y}\text{HUMAN} \\ \text{D-ARG3} = \boxed{z}\text{PHYS\_OBJ} \\ \text{D-E1} = \boxed{e1}\text{STATE} \\ \text{D-E2} = \boxed{e2}\text{PROCESS}\end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix}\text{FORMAL} = \text{LIVE}\left(\boxed{e1}, \boxed{y}, \boxed{x}\right) \\ \text{TELIC} = \text{MAKE\_ACT}\left(\boxed{e2}, \boxed{y}, \boxed{z}\right)\end{bmatrix}\end{bmatrix}$$

(8) $[\![biiru - no\_machi]\!] = \lambda x[\text{town}'(x) \wedge \lambda e_2[\text{TELIC} = \text{make\_act}'(e_2, y, z) \wedge \text{theme}(e_2) = \text{beer}']\ldots]$

(9) kaban-no Kochi “Bags Coach”

$$\begin{bmatrix}\textbf{KOCHI “COACH”} & \\ \text{TYPESTR} = & \begin{bmatrix}\text{ARG1} = \boxed{x}\text{STORE}\end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix}\text{D-ARG1} = \boxed{y}\text{HUMAN} \\ \text{D-ARG2} = \boxed{z}\text{PHYS\_OBJ} \\ \text{D-E1} = \boxed{e1}\text{PROCESS} \\ \text{D-E2} = \boxed{e2}\text{PROCESS}\end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix}\text{FORMAL} = \boxed{x} \\ \text{TELIC} = \text{SELL\_ACT}\left(\boxed{e1}, \boxed{y}, \boxed{z}\right)\end{bmatrix}\end{bmatrix}$$

(10) $[\![kaban - no\_kochi]\!] = \lambda x[\text{Store}'(x) \wedge [\text{TELIC} = \lambda e_1[\text{sell\_act}'(e_1, y, z)] \wedge \text{theme}(e_1) = \text{bag}']\ldots]$

3.2. AGENTIVE Role Modification

3.2.1. Location

A locative genitive phrase *chikaku-no* or “nearby” in *chikaku-no koen* which means “a nearby park” modifies the event contained in the AGENTIVE role of the *park*; in other words, the park was created in a nearby location.

(11) chikaku-no koen “a nearby park”

$$\begin{bmatrix}\textbf{KOEN “PARK”} & \\ \text{TYPESTR} = & \text{ARG1} = \boxed{x}\ \text{outdoor's\_location} \\ \text{ARGSTR} = & \begin{bmatrix}\text{D-ARG1} = \boxed{w}\text{HUMAN} \\ \text{D-ARG2} = \boxed{y}\text{HUMAN} \\ \text{D-ARG3} = \boxed{l}\text{LOCATION} \\ \text{D-ARG4} = \boxed{t}\text{TIME} \\ \text{D-E1} = \boxed{e1}\text{PROCESS} \\ \text{D-E2} = \boxed{e2}\text{PROCESS}\end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix}\text{FORMAL} = \boxed{x} \\ \text{CONSTITUTIVE} = \left\{\text{LAWN, BENCH, FOUNTAIN,}\ldots\right\} \\ \text{TELIC} = \text{RECREATIONAL\_ACTIVITY}\left(\boxed{e1}, \boxed{w}, \boxed{x}\right) \\ \text{AGENTIVE} = \text{MAKE\_ACT}\left(\boxed{e2}, \boxed{y}, \boxed{x}\right)\end{bmatrix}\end{bmatrix}$$

(12) $[\![a\_nearby\_park]\!]$ = λx[...AGENTIVE = λ e_1[make_act(e_1, z, x)$\wedge$ in-neighborhood'(e_1)]...]

Osuro kogai-no mura or "a village in a suburb of Oslo" can be analyzed in a similar manner. Here, a village in a suburb of Oslo exists in the location in a suburb of Oslo since the construction time, therefore, *Osuro kogai-no* or "in a suburb of Oslo" predicates the location where the village was formed.

$$\begin{bmatrix} \textbf{MURA “VILLAGE”} & \\ \text{TYPESTR} = & \begin{bmatrix}\text{ARG1} = \boxed{x}\text{LOCATION}\end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix} \text{D-ARG1} = \boxed{y}\text{HUMAN} \\ \text{D-ARG2} = \boxed{z}\text{PHYS\_OBJ} \\ \text{D-ARG3} = \boxed{a}\text{HUMAN} \\ \text{D-E1} = \boxed{e1}\text{STATE} \\ \text{D-E2} = \boxed{e2}\text{PROCESS} \\ \text{D-E3} = \boxed{e3}\text{PROCESS} \end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix} \text{FORMAL} = \text{LIVE}\left(\boxed{e1}, \boxed{y}, \boxed{x}\right) \\ \text{TELIC} = \text{MAKE\_ACT}\left(\boxed{e2}, \boxed{y}, \boxed{z}\right) \\ \text{AGENTIVE} = \text{MAKE\_ACT}\left(\boxed{e3}, \boxed{a}, \boxed{x}\right) \end{bmatrix} \end{bmatrix}$$

(13) $[\![a\_village\_in\_a\_suburb\_of\_Oslo]\!]$ = λx[village'(x) $\wedge$ [AGENTIVE = λe_3[make_act(e_3, a, x) $\wedge$ in-Oslo-suburb(e_3)]...]]

3.2.2. Outstanding Property

NP_1 can be an outstanding property of NP_2. Since azaleas are the outstanding feature of the park, *tsutsuji-no* "with azaleas" modifies the event in the AGENTIVE role of the park.

(14) tsutsuji-no koen "a park with azaleas"

(15) $[\![a\_park\_with\_azaleas]\!]$ = λx.park'(x) $\wedge$ [AGENTIVE = λe_1[make_act(e_1, z, x)$\wedge$ with-azaleas'(e_1)]...]

Azaleas were planted at the time of building of the park, and hence, the feature of azaleas modifies the AGENTIVE role.

4. Extensional Module Modification

4.1. Augmenting Qualia Structure with a Referential Module

Even though many Japanese postpositional phrases predicate events of one of the qualia of the possessee nominals, we need to account for other cases that cannot be explained by the existing qualia modification.

When possessive nominals represent temporary or changeable features of the possessee nominals, there exists no bindee within the four qualia. For example, the following patterns cannot be accounted for within the existing framework:

(16) a. TIME: *yugata-no koen* "an evening park"

b. LOCATION: *Tokyo-no shinseki* "a relative (living) in Tokyo"

c. ACCOMPANIMENT: *boshi-no fujin* “the lady with a hat lady”

d. PROPERTY: *jutai-no shakuruton*[1] “Shackleton in serious condition”

In order to accommodate noun modification by postpositional phrases that denote temporary location, time, accompaniment, and property, I propose that additional information be encoded into the lexicon. Specifically, I suggest adding a referential module into Generative Lexicon (GL) (Pustejovsky, 1995):

Referential Module:

1. TIME = AT
2. LOCATION = IN
3. MANNER = WITH

Musan (1999) assumes that all noun phrases have the time argument. For example, in (17) below, the person referred to as the intern could have been a hard-working intern in the past or at present. It is possible that he was not an intern when he worked hard as far as he is an intern at present. In other words, the time argument of *the intern* can refer to the past or the utterance time.

(17) The intern worked hard.

Because all physical objects usually occupy some space and time (Sowa, 1999), LOCATION and TIME are subcategories of the REFERENTIAL module.

The following sections demonstrate how the extended GL renders the genitive modification that cannot predicate the events in the present qualia structure.

4.2. Locative Modification

The lexical input for *shinseki* or “relative” in GL does not allow modification by a locative genitive phrase *Tokyo-no* or “in Tokyo,” since *Tokyo-no* “in Tokyo” does not modify the FORMAL role of *shinseki* “relative” because Tokyo is not the location for the kinship relation but the location of the referent of a relative. This word does not modify the AGENTIVE role of *shinseki* “relative” either, since the kinship relation between this relative and another relative could have been formed in another location. Therefore, selective binding is not available for such locative modification.

(18) Tokyo-no shinseki “a relative (living) in Tokyo”

$$\begin{bmatrix} \textbf{SHINSEKI “RELATIVE”} & \\ \text{TYPESTR} = & \begin{bmatrix} \text{ARG1} = \boxed{x}\text{RELATIVE} \\ \text{ARG2} = \boxed{y}\text{HUMAN} \end{bmatrix} \\ \text{EVENTSTR} = & \begin{bmatrix} \text{E1} = \boxed{e1}\text{STATE} \\ \text{E2} = \boxed{e2}\text{PROCESS} \end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix} \text{D-ARG1} = \boxed{z}\text{HUMAN} \end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix} \text{FORMAL} = \text{KINSHIP\_RELATION}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right) \\ \text{AGENTIVE} = \text{KINSHIP\_RELATION}\left(\boxed{e2}, \boxed{z}, \boxed{x}\right) \end{bmatrix} \end{bmatrix}$$

[1]Taken from *Balanced Corpus of Contemporary Written Japanese, 2008 edition* by The National Institute of Japanese Language

Therefore, we incorporate location as part of the referential module such that the location of a relative can be modified by the locative postpositional phrase.

$$\begin{bmatrix} \textbf{SHINSEKI ``RELATIVE''} & \\ \text{TYPESTR} = & \begin{bmatrix}\text{ARG1} = \boxed{x}\text{RELATIVE}\\ \text{ARG2} = \boxed{y}\text{HUMAN}\end{bmatrix} \\ \text{EVENTSTR} = & \begin{bmatrix}\text{E1} = \boxed{e1}\text{STATE}\\ \text{D-E1} = \boxed{e2}\text{TRANSITION}\\ \text{D-E2} = \boxed{e3}\text{STATE}\end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix}\text{D-ARG1} = \boxed{z}\text{HUMAN}\\ \text{D-ARG2} = \boxed{l}\text{LOCATION}\end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix}\text{FORMAL} = \text{KINSHIP\_RELATION}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right)\\ \text{AGENTIVE} = \text{KINSHIP\_RELATION}\left(\boxed{e2}, \boxed{z}, \boxed{x}\right)\end{bmatrix} \\ \text{EXT} = & \begin{bmatrix}\text{LOC} = \text{IN}\left(\boxed{e3}, \boxed{x}, \boxed{l}\right)\end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textbf{TOKYO-NO SHINSEKI ``TOKYO RELATIVE''} & \\ \text{TYPESTR} = & \begin{bmatrix}\text{ARG1} = \boxed{x}\text{RELATIVE}\\ \text{ARG2} = \boxed{y}\text{HUMAN}\end{bmatrix} \\ \text{EVENTSTR} = & \begin{bmatrix}\text{E1} = \boxed{e1}\text{STATE}\\ \text{D-E1} = \boxed{e2}\text{TRANSITION}\\ \text{D-E2} = \boxed{e3}\text{STATE}\end{bmatrix} \\ \text{ARGSTR} = & \begin{bmatrix}\text{D-ARG1} = \boxed{z}\text{HUMAN}\\ \text{D-ARG2} = \boxed{l}\text{LOCATION}\end{bmatrix} \\ \text{QUALIA} = & \begin{bmatrix}\text{FORMAL} = \text{KINSHIP\_RELATION}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right)\\ \text{AGENTIVE} = \text{KINSHIP\_RELATION}\left(\boxed{e2}, \boxed{z}, \boxed{x}\right)\end{bmatrix} \\ \text{EXT} = & \begin{bmatrix}\text{LOC} = \text{IN}\left(\boxed{e3}, \boxed{x}, \boxed{l}\right) \wedge \text{IN-TOKYO'}\left(\boxed{e3}\right)\end{bmatrix} \end{bmatrix}$$

(19) $[\![Tokyo - no\_shinseki]\!]$ = λx[...[LOC = λ e_3[in-Tokyo'(e_3)]...]

4.3. Temporal Modification

The temporal genitive phrase such as *yugata-no* or "evening's" does not modify any of the AGENTIVE or TELIC role because *yugata-no koen* "a park in the evening" means neither the park build in the evening nor that build solely for visiting in the evenings. Rather it refers to the appearance of a park during an evening visit. Thus, walking in the evening park means walking in the park in the evening.

(20) Yugata-no koen-o sanposhi-ta.
evening-GEN park-ACC walk-PAST
"(I/we) walked in a park in the evening."

Hence, *yugata-no* or "evening's" modifies the referential time of the park.

$$\begin{bmatrix}
\textbf{YUGATA-NO KOEN} & \\
\textbf{“A PARK IN THE EVENING TIME”} & \\
\text{TYPESTR} = & \text{ARG1} = \boxed{x}\ \text{outdoor's_location} \\
\text{ARGSTR} = & \begin{bmatrix} \text{D-ARG1} = \boxed{w}\text{HUMAN} \\ \text{D-ARG2} = \boxed{y}\text{HUMAN} \\ \text{D-ARG3} = \boxed{l}\text{LOCATION} \\ \text{D-ARG4} = \boxed{t}\text{TIME} \\ \text{D-E1} = \boxed{e1}\text{TRANSITION} \\ \text{D-E2} = \boxed{e2}\text{STATE} \\ \text{D-E3} = \boxed{e3}\text{PROCESS} \end{bmatrix} \\
\text{QUALIA} = & \begin{bmatrix} \text{FORMAL} = \boxed{x} \\ \text{CONSTITUTIVE} = \left\{\text{LAWN, BENCH, FOUNTAIN,...}\right\} \\ \text{TELIC} = \text{RECREATIONAL_ACTIVITY}\left(\boxed{e3}, \boxed{w}, \boxed{x}\right) \\ \text{AGENTIVE} = \text{MAKE_ACT}\left(\boxed{e1}, \boxed{y}, \boxed{x}\right) \end{bmatrix} \\
\text{EXT} = & \begin{bmatrix} \text{LOC} = \text{IN}\left(\boxed{e2}, \boxed{x}, \boxed{l}\right) \\ \text{TIME} = \text{AT}\left(\boxed{e2}, \boxed{x}, \boxed{t}\right) \wedge \text{IN-EVENING'}\left(\boxed{e2}\right) \end{bmatrix}
\end{bmatrix}$$

4.4. Accompaniment by Manner

Carrying a hat or bag is a temporary feature, that does not modify any inherent qualia roles. It does, however, modify the MANNER role in the EXTENSIONAL structure as shown below.

(21) boshi-no fujin “a hat lady”

$$\begin{bmatrix}
\textbf{FUJIN “LADY”} & \\
\text{TYPESTR} = & \text{ARG1} = \boxed{x}\ \text{human} \\
\text{ARGSTR} = & \begin{bmatrix} \text{D-ARG1} = \boxed{l}\text{LOCATION} \\ \text{D-ARG2} = \boxed{y}\text{PHYS_OBJ} \\ \text{D-E1} = \boxed{e1}\text{STATE} \\ \text{D-E2} = \boxed{e2}\text{STATE} \end{bmatrix} \\
\text{QUALIA} = & \begin{bmatrix} \text{FORMAL} = \boxed{x} \end{bmatrix} \\
\text{EXT} = & \begin{bmatrix} \text{LOC} = \text{IN}\left(\boxed{e1}, \boxed{x}, \boxed{l}\right) \\ \text{MANNER} = \text{WITH}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right) \end{bmatrix}
\end{bmatrix}$$

$$\begin{bmatrix}
\textbf{BOSHI-NO FUJIN}\ \text{“THE LADY WITH A HAT”} & \\
\text{TYPESTR} = & \text{ARG1} = \boxed{x}\text{human} \\
\text{ARGSTR} = & \begin{bmatrix} \text{D-ARG1} = \boxed{l}\text{LOCATION} \\ \text{D-ARG2} = \boxed{y}\text{PHYS_OBJ} \\ \text{D-E1} = \boxed{e1}\text{STATE} \\ \text{D-E2} = \boxed{e2}\text{STATE} \end{bmatrix} \\
\text{QUALIA} = & \begin{bmatrix} \text{FORMAL} = \boxed{x} \end{bmatrix} \\
\text{EXT} = & \begin{bmatrix} \text{LOC} = \text{IN}\left(\boxed{e1}, \boxed{x}, \boxed{l}\right) \\ \text{MANNER} = \text{WITH}\left(\boxed{e1}, \boxed{x}, \boxed{y}\right) \wedge \text{WITH-HAT'}\left(\boxed{e1}\right) \end{bmatrix}
\end{bmatrix}$$

(22) $[\![boshi - no\_fujin]\!]$ = λ x[lady'(x) $\wedge$ λe[manner(e)=with-hat']]

4.5. Computation

As far as compositional calculation of meaning is concerned, I assume that the ϵ operator and the ι operator lower the types of CN into (e). The use of the ϵ operator follows its use for Japanese nouns in Cann et al. (2005).

(23) *boshi* "hat": ϵ x.hat'(x): some x satisfying hat'(x), if there is one

fujin "lady": ιy[lady'(y)]: the unique x satisfying person'(x), if there is such a thing

no: $\lambda X \lambda Y \iota y[Y(y) \wedge R(\epsilon x.X)(y)]$

boshi-no fujin "the lady with a hat": $\iota y[\text{lady'}(y) \wedge \lambda e[\text{manner}(e) = \text{with'}(\epsilon x.\text{hat'})(y)]$

5. Conclusion

Japanese genitive postpositions cannot be disambiguated in terms of the existing qualia of the possessee nominals. We need to augment the lexical input by adding another module, REFERENTIAL or EXTENSIONAL structure. As Vikner and Jensen (2002) did not propose any method for restricting the quale to be used for type-shifting, the present analysis does not provide any suggestions for identifying the quale to be used for the interpretation of the possessive noun phrases. However, it provides the enriched lexical entry which enables access to the sense of NP_2 and determines the semantic relation expressed by Japanese genitive postpositions.

Acknowledgements

I would like to thank the anonymous reviewers for their valuable comments, as well as Yoshihiko Hayashi, Norihiro Ogata, Harry Bunt, and Chris Barker for their suggestions at different stages of this paper's development.

References

Barker, C. (1995). *Possessive Descriptions*. Stanford: CSLI Publications.

Cann, R., R. Kempson, and L. Marten (2005). *The Dynamics of Language: an Introduction*, Volume 35 of *Syntax and Semantics*. Amsterdam, San Diego: Academic Press.

Lenci, A., N. Bel, F. Busa, N. Calzolari, E. Gola, M. Monachini, A. Ogonowski, I. Peters, W. Peters, N. Ruimy, M. Villegas, and A. Zampollo (2000). Simple: A general framework for the development of multilingual lexicons. *International Journal of Lexicography 13:4*, 249–263.

Makishita, Y. (1984). *Nihongo-kara Mita Eigohyogen: Eigo Jutsubu-no Imiteki Kosatsu-o Chushin-to Shite (English Expression from the perspective of Japanese: Centering on Semantic Investigation of English Predicates*. Tokyo: Kenkyusha.

Musan, R. (1999). Temporal interpretation and information-status of noun phrases. *Linguistics and Philosphy 22*, 621–661.

Partee, B. H. (1983, 1997). Genitives - a case study. appendix to theo m.v.janssen, ”compositionality”. In *Handbook of Logic and Language*, pp. 464–470. Amsterdam: Elsevier Science Publishers.

Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge: MIT Press.

Sowa, J. F. (1999). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Course Technology.

Teramura, H. (1980). Meishi shushokubo-no hikaku (comparison of noun modifiers).

Vikner, C. and P. A. Jensen (2002). A semantic analysis of the english genitive. interaction of lexical and formal semantics. *Studia Linguistica 56*, 191–226.

CAN DP BE A SCOPE ISLAND?

Simon Charlow
Department of Linguistics, New York University

1 Introduction

Sauerland (2005) uses data from inverse linking—cf. May (1977)—to motivate QR out of DP and proposes to derive Larson's (1987) generalization regarding the scopal integrity of DP via an economy-based constraint on QR (cf. Bruening 2001).

This squib is in four parts. I first lay out Sauerland's (2005) three arguments for QR out of DP. I present (a slightly modified version of) his mechanism for constraining QR. I show that it both over- and under- generates. I conclude with some alternative explanations for the readings Sauerland takes to motivate QR out of DP.

2 Sauerland's data

2.1 Modal intervention

Sauerland points out that (1) can be true if Mary doesn't have any specific individuals in mind and doesn't want to get married twice (say she's placed a classified ad indicating she wants to meet and marry a man from *either* Finland or Norway):

(1) Mary wants to marry someone from these two countries. (Sauerland's ex. 8a)

Sauerland concludes that (a) the non-specificity of Mary's desire suggests that the indefinite remains within the scope of the bouletic operator $\mathscr{O}$, and (b) the fact that Mary needn't desire two marriages requires that *these two countries* be outside the scope of $\mathscr{O}$. In sum, the scope ordering is $2 > \mathscr{O} > \exists$ and requires QR out of DP:

(2) [these two countries]$_x$ [Mary wants [$\lambda w'$. PRO marry [someone from x] in w']]

The first of these points seems correct. If the semantics of *want* involves quantification over want-worlds w', this scope ordering entails that the individual Mary marries can vary with each w'_{Mary}. This derives a non-specific desire.

Point (b) is more subtle. Depending on the semantics assigned to *want*, scoping *two* over it may be insufficient to derive the disjunctive reading. In brief: the existence of two individuals x such that Mary marries x in each of her want-worlds w' still requires, on a naïve semantics for *want*, that Mary marry twice in each w'. More sophisticated semantics for *want*—e.g. Heim (1992)—may obviate this worry.

Grant that scoping *two* over *want* can derive the disjunctive reading. Sauerland's (2005) point then requires that leaving *two* within the scope of $\mathscr{O}$ be incompatible with a disjunctive desire. We return to this below.

2.2 Antecedent-contained deletion

Sauerland corroborates this observation by noting the grammaticality of "wide" non-specific readings of ACD constructions like (3):

(3) Mary wants to marry someone from every country Barry does.

(3) can be true if (a) neither Mary nor Barry has anyone specific in mind and (b) the ACD is resolved "wide"—viz. anaphoric to the larger VP *want to...*

As before, the first of these points suggests that the indefinite remains within the scope of *want*. Additionally, standard assumptions require that the DP containing the ellipsis site QR past the verb heading the antecedent VP in order to resolve the antecedent-containment paradox—cf. May (1985). The scope ordering $\forall > \mathcal{O} > \exists$ again requires QR out of DP.

2.3 Negation intervention

Finally, Sauerland echoes Huang's (1998) observation regarding (4):

(4) John didn't see pictures of several baseball players.

(4) has a reading judged true if there are several baseball players x such that John didn't see any pictures of x—*several* $> \neg > \exists$. Sauerland assumes that the scope of the existential quantifier marks the LF position of the bare plural (a proposition I dispute below) and safely establishes that the cardinal indefinite occupies an LF position above negation. The by-now-familiar conclusion is that this reading requires QR out of DP.

3 Larson's generalization and constraining QR

Larson (1987) observes that a QP external to a DP X must scope either below or above all scopal elements in X (i.e. no interleaved scope):

(5) Three men danced with a woman from every city. (*$\forall > 3 > \exists$)

(6) Several students ate a piece of every pie. (*$\forall >$ *several* $> \exists$)

The conclusion usually drawn from this datum is that QR out of DP is illicit. Inverse linking instead results from QR of the embedded QP to a DP-adjunction position:

(7) $[_{\text{DP}}$ [every city]$_x$ $[_{\text{DP}}$ someone from x]] left

This approach is adopted in e.g. May (1985); Rooth (1985); Büring (2004). If QR into DP is likewise illicit, Larson's generalization is derived.

3.1 Superiority

Sauerland rejects DP's scope island-hood, arguing that subjecting QR to Superiority in the sense of Richards (1997); Bruening (2001) accounts for generalizations (8) and (9).

(8) $\text{QP}_1\ [_{\text{QP}_2}\ [\text{QP}_3]] \rightsquigarrow {}^*\text{QP}_3 > \text{QP}_1 > \text{QP}_2$ (Larson 1987)

(9) $\mathcal{O}\ [_{\text{DP}}\ [\text{QP}]] \rightsquigarrow \text{QP} > \mathcal{O} > \text{DP}$ (Sauerland 2005)

We don't dwell on the syntactic minutiae of Sauerland's account here. It will be sufficient to note that the relative scope of two QPs can be reversed iff the reversal is required for interpretation. This is effected by ordering QR of higher QPs before QR of lower QPs and requiring that QR be to the nearest node of type t (thus the lower QP in general lands below the higher one).[1] "Canonical" inverse scope is derived by total reconstruction of the subject QP (which Sauerland conceptualizes in terms of subject movement at PF).

Sauerland assumes that absent DP-internal clausal syntax, DP-embedded QPs are uninterpretable in situ. Accordingly, they QR to the nearest node of type t. If the embedding DP is quantificational, this entails a scope inversion (note that surface scope readings of inversely linked constructions are predicted impossible, something we revisit in §4.1). If the QP-containing DP is itself uninterpretable—e.g. in object position—a proper characterization of Superiority (cf. Bruening 2001; Charlow 2009) requires that it QR before the embedded QP.

This is all that's needed to derive Larson's generalization in the extensional case.

(10) $[_{v\text{P}}$ $[_{\text{QP}_1}$ three men$]$ danced with $[_{\text{QP}_2}$ a woman from $[_{\text{QP}_3}$ every city$]]]$

Two scenarios are possible. Either (a) QP_1 moves to [Spec,TP] at LF (it QRs), or (b) it moves there at PF (it doesn't). In the first case each QP QRs. The only inversion required for interpretation is between QP_2 and QP_3, and so the scope ordering $1 > 3 > 2$ is derived. In scenario (b) QR applies twice. One inversion comes for free (QP_1 and QP_2; since QP_1 doesn't raise, QP_2 goes to the nearest node of type t—viz. above QP_1), and one is required for interpretation (QP_2 and QP_3). Superiority also requires that QP_2 raise over QP_1 before QP_3 raises out of QP_2. Thus the scope ordering $3 > 2 > 1$ is derived. In both scenarios QP_2 and QP_3 scope *together* relative to QP_1.

3.1.1 Non-QP operators

The following constructions replace the subject QP with an intensional operator/negation:

(11) Mary wants $[_{\text{TP}}$ PRO to marry $[_{\text{QP}_1}$ someone from $[_{\text{QP}_2}$ these two countries$]]]$

(12) $[_{\text{NegP}}$ not $[_{v\text{P}}$ John see $[_{\text{QP}_1}$ pictures of $[_{\text{QP}_2}$ several baseball players$]]]]$

Both structures require QR of QP_1 and QP_2 to a TP/vP-adjunction position for interpretation. QP_2 may subsequently continue climbing the tree. It's free to raise over *want*/*not*; Superiority doesn't come into play. Thus the scope ordering $2 > \mathcal{O} > 1$ is derived (similarly, the ACD example is predicted grammatical).

4 Problems with the account

4.1 Surface scope

As noted in §3.1, Sauerland's account predicts that a DP-embedded QP can never scope inside its embedding DP.

(13) John bought a picture of every player on the team. (Sauerland 2000's ex. 40a)

(14) John bought a picture of each player on the team. (Sauerland 2000's ex. 40b)

[1] This oversimplifies the mechanism Bruening and Sauerland propose. I don't think this affects any of my points. See Sauerland (2000); Charlow (2009) for further discussion.

(15) Everyone/no one from a foreign country eats sushi. (after Heim & Kratzer 1998 221, ex. 1)

As Sauerland (2000) notes, example (13) has a reading on which it's true if John bought a single picture with everyone in it and false if he bought many individual pictures but no single picture with everyone. Though this reading seems to require surface scope (viz. $\exists > \forall$) Sauerland suggests it may stem from a "group interpretation" of wide-scoping *every player on the team*—i.e. roughly equivalent to *all the players on the team*. If a group interpretation is unavailable for e.g. *each player on the team* in (14), Sauerland argues, we have an explanation for why surface scope (viz. $\exists > \forall$) is "unavailable" here.

A few comments are in order. First, the ungrammaticality of $\exists > \forall$ in (14) is actually not clear. Though the surface-scope reading may be marked, this follows from *each*'s oft-noted strong preference for wide scope. Second, the grammaticality of (15) on its surface-scoping reading—viz. every/no x such that x is from a foreign country is such that x eats sushi ($\forall/\neg\exists > \exists$)—cannot be answered by appeal to group interpretations of the embedded QP. A theory of inverse linking must, it seems, account for "surface" linking. Absent an ad hoc appeal to abstract clausal syntax inside DP, Sauerland (2000, 2005) cannot.

4.2 Reliance on covert clausal syntax

Rooth (1985); Larson (1987) observe that QPs embedded in nominal intensional complements can be read *de dicto*:

(16) Max needs a lock of mane from every unicorn in an enchanted forest.

(16) (Larson's ex. 4a) has a reading on which it's true if Max is trying to perform a spell which requires him to pick an enchanted forest and then procure a lock of mane from every unicorn in it. Max's need in this scenario is nonspecific with respect to both unicorns and locks of mane, suggesting that both remain within the scope of the intensional verb *need*.

The DP-as-scope-island approach to inverse linking predicts this state of affairs. QR of the embedded QP targets the DP-internal adjunction site rather than the nearest node of type t. The embedded QP can—indeed must—remain within the scope of *need*.

Something more needs to be said on Sauerland's account. Following Larson et al. (1997) he proposes that intensional transitives take abstractly clausal complements. Informally, the syntax of (16) is something like *Max needs PRO to have...* The infinitive clause offers a type-t landing site for the embedded QP below *need*. Abstract clausal syntax in complements of intensional transitives is thus an essential feature of Sauerland's account.

4.3 Double-object behavior in intensional cases—an over-generation issue

Consider now the following two cases:[2]

(17) Two students want to read a book by every author. (*$\forall > 2 > \exists$)

(18) Two boys gave every girl a flower. ($\forall > 2 > \exists$)

Example (17) lacks the starred reading—unsurprisingly given Larson's generalization. Example (18)—discussed by Bruening in unpublished work, and given as Sauerland's

[2] I thank an anonymous reviewer for comments which helped me sharpen this point.

(2000) ex. 49—permits an intervening scope reading (i.e. on which boys vary with girls and flowers with boys).

(19) $[_{QP_1}$ two students] want [$[_{QP_3}$ every author$]_x$ [$[_{QP_2}$ a book by $x]_y$ [PRO to read y]]]

(20) $[_{QP_1}$ two boys] gave $[_{QP_3}$ every girl] $[_{QP_2}$ a flower]

(19) and (20) represent intermediate steps in the derivations of (17) and (18), respectively. In (19) QP_2 has raised from object position, and QP_3 has raised out of QP_2. The difficulty for Sauerland here is that (19) and (20) are predicted to license the same subsequent movements—the numbering is intended to highlight this. If in both cases QP_1 moves only at PF we may derive the following structures:

(21) $[_{QP_3}$ every pie$]_x$ [$[_{QP_1}$ two students] want [x [$[_{QP_2}$ a piece of $x]_y$ [PRO to eat y]]]]]

(22) $[_{QP_3}$ every girl$]_x$ [$[_{QP_1}$ two boys] gave x $[_{QP_2}$ a flower]]

This is a good result for (18) but a bad one for (17). In sum, Sauerland predicts that though inversely linked DPs in extensional contexts obey Larson's generalization, those in intensional contexts do not. The contained and containing QPs in (17) are incorrectly predicted to behave like the indirect-object (IO) and direct-object (DO) QPs in (18) (viz. permitting $3 > 1 > 2$).

Note, moreover, that appealing to the obligatoriness of the QR in (22) as compared to the non-obligatoriness of the QR in (21) won't help:

(23) A (different) child needed every toy. ($\forall > \exists$)

(24) Two boys want to give every girl a flower. ($\forall > 2 > \exists$)

(23) possesses an inverse-scope reading (on which children vary with toys), and (24) possesses the interleaved scope reading that (17) lacks. As per Sauerland's assumption, the syntax of (23) is actually as in (25):

(25) [a (different) child] needed [PRO to have [every toy]]

(26) [two boys] want [PRO to give [every girl] [a flower]]

QR of *every toy/girl* above the subject isn't obligatory in either case. In both instances obligatory QR targets a position below the intensional verb (and thus below the subject QP). In short, Sauerland *needs* to allow non-obligatory QR to reorder subject and object QPs. Ruling this mechanism out in order to save (17) dooms (23) and (24).

4.4 ECM QPs—an under-generation issue

The following constructions are grammatical when the ECM indefinite scopes below the matrix-clause intensional operator $\mathscr{O}$ (evidenced by NPI grammaticality in 27 and a nonspecifically construed indefinite in 28) and the bracketed QP scopes above $\mathscr{O}$:[3]

(27) Frege refused to let any students search for proofs of [at least 597 provable theorems]

(28) Frege wanted many students to desire clear proofs of [every theorem Russell did]

In (27) the bracketed QP can be (indeed, on its most salient reading *is*) construed *de re*. Frege need never have wanted anything pertaining *de dicto* to $\geq$597 provable theorems.

[3] For Sauerland, anyway. I discuss below why I don't think *de re* diagnoses wide scope.

Say he made a habit of dissuading students from searching for proofs of theorems he considered unprovable, but by our reckoning he engaged in no fewer than 597 erroneous dissuasions. For Sauerland this requires 597 > *refused*. In (28) wide ACD resolution is permitted. As Sauerland observes, this suggests that the bracketed QP scopes at least as high as *wanted*.

Both of these "wide" readings are compatible with $\mathscr{O} > \exists_{\mathrm{ECM}}$, a situation Superiority predicts impossible. Obligatory QR of the bracketed QPs in both cases targets a node below the ECM indefinite (N.B. the verbs in the infinitival complements are intensional transitives; on the Larson et al. analysis of these constructions their complements are clausal; obligatory QR of the bracketed QPs thus targets a position *below* the infinitival intensional transitive). If the ECM indefinite stays within the scope of $\mathscr{O}$, Superiority predicts—barring total reconstruction of the indefinite[4]—that the bracketed QP will be unable to take scope over $\mathscr{O}$, contrary to fact. I return to both of these constructions in §5.

5 Re-evaluating Sauerland's data

5.1 Modal intervention?

Does the non-specific disjunctive-desire reading of (1)—repeated here as (29)—require QRing *these two countries* over the intensional operator? Here's some evidence it doesn't:

(29) Mary wants to marry someone from these two countries.

(30) (Pointing to "Toccata and Fugue in D minor" and "O Fortuna" on a list of classical music) When these two songs play in a movie, someone's about to die.

(31) a. The paranoid wizard refuses to show anyone these two amulets.
b. The paranoid wizard refuses to show more than two people these two amulets.

(32) a. You may show a reporter (lacking a security clearance) these two memos.
b. [Ms. Goard] declined to show a reporter those applications.
c. At least some states consider it to be attempted murder to give someone these drugs.
d. When you give someone these viruses, you expect to see a spike as gene expression changes.

(33) #Mary wants to marry someone from every Scandinavian country.

(34) #When every Stravinsky song plays in a movie, someone's about to die.

To the extent that (29) can express something felicitous, so can (30), despite the fact that QR of *those two songs* over the modal operator *when* is blocked by a tensed clause boundary. Specifically, (30) needn't quantify over situations in which two songs play. The availability of a disjunctive reading in this case (viz. ≈*when either of those two songs plays in a movie, someone's about to die*) suggests that QR out of DP may not be required for a felicitous reading of (29).

Example (31a), whose infinitival complement hosts a double object configuration, corroborates this assessment. Double object constructions are known to disallow QR of the DO over the IO—cf. Bruening (2001). Here the NPI/nonspecific IO remains within the scope of the downward-entailing intensional operator *refuse*. Accordingly, *these two amulets* cannot QR above *refuse*. Nevertheless, the most salient reading of (31a) involves a wizard who doesn't show anyone *either* of the two amulets.[5] Similarly (31b) permits a

[4]The aforementioned anonymous reviewer notes that total reconstruction as generally understood only applies to A-chains, not QR chains. True enough.

[5]Superiority theorists may counter that NPIs aren't subject to QR and thus that the DO is free to QR over *anyone* in (31a). This leaves (31b) and (32) mysterious. Additionally, Merchant (2000) shows that NPIs can host ACD gaps, suggesting they QR, after all—cf. *that boy won't show anyone he should his report card.*

reading such that the paranoid wizard won't show *either* of the two amulets to any group of three or more people.

Similarly, (32a) allows a disjunctive construal of *these two memos*. On this reading, you are conferred permission to show any reporter lacking a security clearance *either* of the two memos (and possibly both). So you're being compliant if you show such a reporter memo #1 but not memo #2. This is again despite a nonspecific IO, which should prohibit QR of *these two memos* to a position over the deontic modal. Examples (32c)–(32d) likewise permit nonspecific IOs alongside disjunctively construed DOs, despite double object configurations (and a tensed clause boundary in 32d).[6]

Finally, (33) and (34) lack felicitous readings (given certain norms surrounding marriage and film scores). They are incompatible with scenarios in which Mary wants to marry once, and every Stravinsky song playing in a given situation isn't a clue about anything. This suggests that plural demonstratives may be necessary for disjunctive readings of (29)–(32).[7]

In sum, (30)–(32) show that QR over an intensional operator cannot be necessary for a disjunctive construal of a plural demonstrative. Examples (33) and (34) show that in certain cases the plural demonstrative is a necessary component of the disjunctive reading. These facts follow if we assume that disjunctive readings in these cases aren't (necessarily) due to QR over an intensional operator but may instead arise when plural demonstratives occur in the scope of modal (or downward-entailing, cf. §5.2) operators.[8]

5.2 Negation intervention?

Recall Huang's (1998) negation-intervention cases—e.g. (4), repeated here as (35):

(35) John didn't see pictures of several baseball players (at the auction).

As Huang (1998) observes and Sauerland confirms, constructions like (35) allow a reading with *several* $> \neg > \exists$. Several baseball players x, in other words, are such that John didn't see any pictures of x.

Independently motivated semantic apparatus for bare plurals helps explain these data. If DP *is* a scope island, scoping *several* over *not* requires QRing the bare plural over *not*:

(36) [[several baseball players]$_x$ pictures of x]$_y$ [John didn't see y]

We assume following Chierchia (1998) that bare plurals sometimes denote kinds and that combining a kind-level argument with a predicate of objects creates a type-mismatch resolved by an operation called 'D(erived) K(ind) P(redication).' Following Magri (2004), the semantics of DKP is as follows:

(37) For any P denoting a predicate of objects:
DKP$(P) = \lambda x.[\exists y : y \leq x][Py]$, where $y \leq x$ iff y instantiates the kind x.

DKP generalizes to n-ary relations in the usual way (cf. Chierchia 1998 fn. 16), introducing an existential quantifier within the scope of the shifted verb.

[6](32b)–(32d) were obtained via Google search. They can be accessed at the following links: 1, 2, and 3—each of which displays the nonspecific-IO/disjunctive-DO reading.

[7]Though plural indefinites seem to work similarly in certain cases. See §5.2.

[8]Disjunctive readings might involve something like a free-choice effect or exceptional scope (i.e. scope out of islands which doesn't require QR out of islands).

That DPs of the form *pictures of several baseball players* denote kinds on their inversely linked readings is confirmed by (a) the felicity of (38) and (b) the *absence* of a *several* $>$ $\exists > \neg$ reading for (35) (repeated as 39):

(38) Pictures of several baseball players are rare.

(39) John didn't see pictures of several baseball players (at the auction).

Returning to (36), note that the trace y left by QR of the bare plural will (presumably) be kind-level.[9] This creates a mismatch between *see* and the bare plural's trace y. DKP applies to $[\![\text{see } y]\!]$, introducing an $\exists$ within the scope of a $\neg$:

(40) $\lambda z \,.\, \textbf{see}\, y\, z \;\mapsto_{\text{DKP}}\; \lambda z \,.\, [\exists x : x \leq y][\textbf{see}\, x\, z]$

This derives *several* $> \neg > \exists$, despite the prohibition on QR out of DP.

5.2.1 *Plural indefinites and demonstratives under negation*

Other factors may be at work in these cases. Recall (31a), repeated here as (41):

(41) The paranoid wizard refuses to show anyone these two amulets.

(42) The paranoid wizard refuses to show anyone several (of his) amulets.

As noted previously, (41) requires the NPI IO to remain under *refuse*, while permitting a (disjunctive) reading truth-conditionally equivalent to 2 $>$ *refuse*. Interestingly, the same goes for (42), which replaces the demonstrative with a plural indefinite and admits a (disjunctive) reading equivalent to *several* $>$ *refuse*. In both cases scope freezing should prohibit QR of the DO over the IO to a position above *refuse*. It is hypothesized that these readings instead result from disjunctively construed DOs.

Might QR of (35)'s inversely-linked bare plural over negation thus be unnecessary for a reading which gives *several* apparent scope over negation? Consider the following cases:

(43) John didn't read any books by these two authors. ($\approx 2 > \neg > \exists$)

(44) John didn't read any books by several authors. (??? $\approx$ *several* $> \neg > \exists$)

These examples replace (35)'s bare plural with full determiner phrases. (43) allows a (disjunctive) reading equivalent to $2 > \neg > \exists$, whereas the disjunctive reading of (44) is borderline ungrammatical. Why this might obtain is unfortunately beyond the scope of what I can consider here, but it shows that the QR+DKP story may be necessary for (35) but not for an example with a plural demonstrative in lieu of the plural indefinite.[10,11]

5.3 ACD and scope shift

Sauerland's ACD data remains to be discussed. (45) is grammatical with a nonspecific indefinite and wide ACD resolution, suggesting QR out of DP.

(45) Mary wants to marry someone from every country Barry does.

---

[9] *Pictures of several baseball players* will denote something like the set of predicates κ of kinds such that for several baseball players x, the $y = [\![\text{pictures of } \ulcorner x \urcorner]\!]$ (where $[\![\ulcorner x \urcorner]\!] = x$) is such that $\kappa y = 1$.

[10] The contrast between (42) (permits a disjunctive reading) and (44) (doesn't) is also unexplained.

[11] Note also that the contrast between (43) and (44) doesn't follow for Sauerland, who permits the embedded QP to QR over negation in both cases.

I wish to suggest that QR to resolve ACD is a more powerful scope-shift mechanism than QR which isn't required for interpretation. A similar claim is made in von Fintel & Iatridou (2003), who distinguish "ACD-QR" from "Scope-QR"—viz. QR which doesn't resolve antecedent containment. von Fintel & Iatridou note that ACD licenses QR across a tensed clause boundary and negation, both islands for Scope-QR:

(46) John hopes I marry everyone you do.

(47) John said that Mary will not pass every student that we predicted he would (~~say...~~).

In the following I consider some additional evidence in favor of an ACD-QR mechanism distinct from and more powerful than Scope-QR.

5.3.1 ACD and DE operators

Examples (48) and (49) differ in that (49) hosts an ACD gap, whereas (48) does not. The reading of (49) we're interested in involves wide ACD resolution:

(48) Mary denied kissing everyone. (??$\forall > deny$)
cf. Mary imagined kissing everyone. ($\forall > imagine$)

(49) Mary denied kissing everyone Barry did. ($\forall > deny$)

QPs headed by *every* do not readily QR over downward-entailing operators—cf. Beghelli & Stowell (1997); von Fintel & Iatridou (2003).[12] (48) doesn't permit $\forall > deny$ without focal stress on *everyone*. The wide ACD reading of (49), by contrast, permits (actually, requires) $\forall > deny$ (note that although *Barry* is focused, *everyone* is not).

5.3.2 Double object constructions

Imagine a scenario as follows: a bus full of Red Sox players pulls up. Mary and Barry both mistake them for the Yankees. Each of them wants to give the same presents to some player (or other?) on the bus.

(50) Mary wants to give a Yankee everything Barry does.

(50) is grammatical with *a Yankee* read *de dicto*—as required by the mistaken beliefs of Mary and Barry in our scenario—and wide ACD resolution, *pace* Bruening (2001).[13] Whether this reading permits gift recipients to vary with gifts is a more difficult matter.[14] Nevertheless, the grammaticality of a wide ACD site hosted in a DO ($\forall > \mathscr{O}$), combined with a *de dicto* IO ($\mathscr{O} > \exists$), requires subverting double-object scope freezing.

(51) The paranoid wizard's wife refuses to show anyone the same two amulets her husband does.

(51) is grammatical with wide ACD. Given the NPI IO, this requires $2 > \mathscr{O} > \exists$. Again, ACD QR subverts the prohibition on QR of DO over IO in double object configurations.

[12] N.B. these authors only consider QR of *every* over *not*.

[13] Larson (1990) discusses an example like (50) in his fn. 10 but doesn't consider whether it allows the IO to be read *de dicto*. Bruening (2001) considers two examples like (50)—his (27a) and (27b)—but concludes they don't admit a *de dicto* IO, contrary to the judgments of my informants and myself.

[14] As Bruening correctly notes, *Mary gave a child every present Barry did* is grammatical but doesn't allow girls to vary with presents (*$\forall > \exists$). He proposes that both IO and DO QR, the DO since it contains the ACD gap and the IO to get out of the DO's way (i.e. to preserve Superiority). Thus $IO > DO$ is preserved. Examples (50) and (51) suggest that this may not be the right approach.

5.3.3 *Larson's generalization*

Recall (28), repeated here (slightly modified) as (52):

(52) Frege wanted a student to construct a proof of [every theorem Russell did]

Previously we focused on how (28) represented a problem for Sauerland. Superiority predicts that a nonspecific reading of the ECM indefinite will be incompatible with wide ACD resolution, contrary to fact.

Note, however, that this reading represents a problem for just about anybody. Specifically, its grammaticality entails a violation of Larson's generalization:[15]

(53) [every theorem Russell ~~wanted a student to construct a proof of x~~]$_x$ [Frege wanted a student to construct a proof of x]

LF (53) entails that a QP intervenes between a DP-embedded QP and its remnant! In other words: the same strategy that Sauerland uses to argue that DP isn't a scope island allows us to construct examples wherein Larson's generalization doesn't. But of course we don't want to conclude that Larson's generalization doesn't *ever* hold.

In sum, since ACD-QR can cross islands, Sauerland's ACD examples aren't dispositive for the DP-as-scope-island hypothesis.

6 Conclusions

This squib has offered evidence that the conclusions reached in Sauerland (2000, 2005) favoring QR out of DP may not be warranted. Most seriously, the mechanism Sauerland proposes to derive Larson's generalization only really works for extensional cases, overgenerating when inversely-linked DPs occur in intensional contexts. Sauerland's account also struggles with "surface-linked" interpretations of inversely linked DPs, a reliance on covert clausal syntax in intensional transitive constructions, and ECM interveners which should block certain readings but appear not to.

On closer inspection, readings analogous to those which Sauerland takes to motivate QR out of DP occur in constructions where (we have independent reason to believe that) QR above a certain relevant operator isn't an option. Importantly, each of Sauerland's arguments for QR out of DP is given a double-object-construction rejoinder. I have speculated that plural demonstratives/indefinites in the scope of modal/negation operators can be construed disjunctively in the absence of QR. Additionally, if (following Chierchia 1998) the scope of an $\exists$ quantifier isn't diagnostic of the LF position of a kind-denoting bare plural, Huang's negation-intervention cases don't require a split DP.

Finally, in line with von Fintel & Iatridou (2003), I've provided new arguments that ACD-QR can do things Scope-QR can't: namely, scope an *every*-phrase over a downward-entailing operator, carry a DO over an IO in double object configurations, and precipitate violations of Larson's generalization.

Some of the criticisms I've mounted against Sauerland will also militate against Bruening's (2001) characterization of QR as Superiority-governed. Additionally, it remains

[15]Similar comments don't apply to (27). As many authors—e.g. Farkas (1997); Keshet (2008)—have shown, scoping a QP over an intensional operator may be sufficient for a *de re* reading of that QP, but it cannot be necessary.

to be determined to what extent plural demonstratives and indefinites behave as a piece with respect to disjunctive readings, why this might be the case, and what any of this has to do with modals/negation. I must leave consideration of these matters to future work.

References

F. Beghelli & T. Stowell (1997). 'Distributivity and negation: the syntax of *each* and *every*'. In A. Szabolcsi (ed.), *Ways of Scope Taking*, pp. 71–109. Kluwer, Dordrecht.

B. Bruening (2001). 'QR Obeys Superiority: Frozen Scope and ACD'. *Linguistic Inquiry* **32**(2):233–273.

D. Büring (2004). 'Crossover Situations'. *Natural Language Semantics* **12**(1):23–62.

S. Charlow (2009). *Inverse linking, Superiority, and QR*. Ms., New York University.

G. Chierchia (1998). 'Reference to Kinds across Languages'. *Natural Language Semantics* **6**:339–405.

D. Farkas (1997). 'Evaluation Indices and Scope'. In A. Szabolcsi (ed.), *Ways of Scope Taking*, pp. 183–215. Kluwer, Dordrecht.

K. von Fintel & S. Iatridou (2003). 'Epistemic Containment'. *Linguistic Inquiry* **34**(2):173–198.

I. Heim (1992). 'Presupposition Projection and the Semantics of Attitude Verbs'. *Journal of Semantics* **9**:183–221.

I. Heim & A. Kratzer (1998). *Semantics in Generative Grammar*. Blackwell, Oxford.

C.-T. J. Huang (1998). *Logical relations in Chinese and the theory of grammar*. Garland, New York. (Ph.d. thesis, MIT, 1982).

E. Keshet (2008). *Good Intensions: Paving Two Roads to a Theory of the De re/De dicto Distinction*. Ph.d. thesis, MIT.

R. K. Larson (1987). *Quantifying into NP*. Ms., MIT.

R. K. Larson (1990). 'Double Objects Revisited: Reply to Jackendoff'. *Linguistic Inquiry* **21**(4):589–632.

R. K. Larson, et al. (1997). *Intensional Transitive Verbs and Abstract Clausal Complementation*. Ms., SUNY at Stony Brook, Vrije Universiteit Amsterdam.

G. Magri (2004). 'Constraints on the readings of bare plural subjects of individual-level predicates: syntax or semantics?'. In L. Bateman & C. Ussery (eds.), *Proceedings from NELS35*, vol. I, pp. 391–402. GLSA Publications, Amherst, MA.

R. May (1977). *The Grammar of Quantification*. Ph.d. thesis, MIT.

R. May (1985). *Logical Form: Its Structure and Derivation*. MIT press, Cambridge.

J. Merchant (2000). 'Antecedent-contained deletion in negative polarity items'. *Syntax* **3**(2):144–150.

N. Richards (1997). *What moves where when in which language?* Ph.d. thesis, MIT.

M. Rooth (1985). *Association with Focus*. Ph.d. thesis, UMass, Amherst.

U. Sauerland (2000). *Syntactic Economy and Quantifier Raising*. Ms., Universität Tübingen.

U. Sauerland (2005). 'DP Is Not a Scope Island'. *Linguistic Inquiry* **36**(2):303–314.

Mildly non-planar proof nets for CCG

Armin Buch
Department of Linguistics, University of Tübingen

Abstract. I introduce a *mildly non-planar* subclass of proof nets (Girard 1987) as an abstraction over derivational ambiguity in Combinatory Categorial Grammar (Steedman 2000). Proof nets encode the linear order of a sentence, its scope relations and (long distance) dependencies, and its syntactic structure and semantics — via the Curry-Howard correspondence — simultaneously and unambiguously.

1. Introduction

Derivations in most variants of categorial grammars (CGs) are valid proofs in linear logic (Girard 1987), or a subtype thereof, e.g. Lambek Calculus (Lambek 1958). Parsing a sentence with a CG corresponds to proving a sequent. Different derivations can denote the same λ-term via the Curry-Howard correspondence. In linear logic, proof nets abstract over this spurious ambiguity. A proof net denotes exactly one λ-term; and different proof nets denote different λ-terms. Thus one can use proof nets to abstract over spurious ambiguity in CG parses.

Proof nets for derivations in the Lambek calculus enjoy the property of being *planar*. Lambek calculus is weakly equivalent to context-free grammars (Pentus 1993). So planarity corresponds to context-freeness. In this paper, I explore proof nets as a presentation of parses in a mildly context-sensitive categorial grammar: Combinatory Categorial Grammar (Steedman 2000, Steedman & Baldridge to appear, henceforth CCG). The process of parsing (with or without proof nets) is an entirely different matter, which I will not address here.

2. Combinatory Categorial Grammar

I repeat basic definitions for CCG, and address derivational ambiguity. In search for an unambiguous and transparent presentation I briefly discuss a normal form and a chart presentation, which both turn out unsatisfactory. A satisfactory solution is found with proof nets, in the section hereafter.

2.1. Definition of CCG

As its name suggests, CCG is a combinatory logic characterized by a set of recursively defined categories and rules how to combine them. Let $\textsc{cat}_0$ be the set of basic categories, e.g. $\{S, NP, N, PP\}$. The set of categories $\textsc{cat}$ is the smallest set such that

- $\textsc{cat}_0 \subset \textsc{cat}$
- $A\backslash B, A/B \in \textsc{cat}$ *iff* $A \in \textsc{cat} \wedge B \in \textsc{cat}$

A note of caution: Talking about CCG, it makes sense to use Steedman's *result left-most* notation, which I do throughout the paper. Read $A\backslash B$ (A/B) as "yielding A if given B on the left (right)".

CCG is a lexicalized formalism. Each word in a sentence introduces a category. There are no empty words.[1] For semantic interpretation via the Curry-Howard correspondence, a typed λ-term is assigned to every category. Types and categories match structurally; see Roorda (1991). The variant of CCG I will use in this paper is given by the following set of combination rules, along with their interpretation:

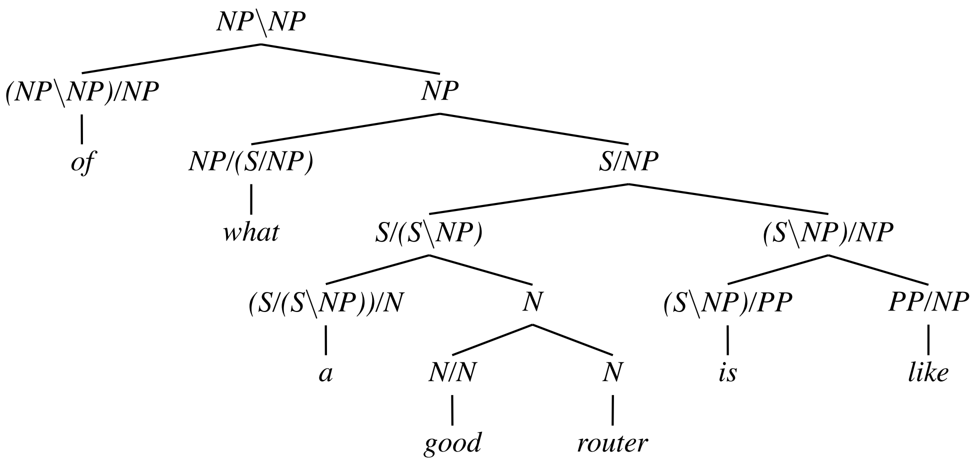
(>)	right functional application	$X/Y : f$	$Y : a$	$\rightarrow$	$X : f(a)$
(<)	left functional application	$Y : a$	$X\backslash Y : f$	$\rightarrow$	$X : f(a)$
(>B)	right functional composition	$X/Y : f$	$Y/Z : g$	$\rightarrow$	$X/Z : \lambda z.f(g(z))$
(<B)	left functional composition	$Y\backslash Z : g$	$X\backslash Y : f$	$\rightarrow$	$X\backslash Z : \lambda z.f(g(z))$
(>$\text{B}_\times$)	forward crossing composition	$X/Y : f$	$Y\backslash Z : g$	$\rightarrow$	$X\backslash Z : \lambda z.f(g(z))$
(<$\text{B}_\times$)	backward crossing composition	$Y/Z : g$	$X\backslash Y : f$	$\rightarrow$	$X/Z : \lambda z.f(g(z))$
(>T)	forward type-raising		$X : a$	$\rightarrow$	$T/(T\backslash A) : \lambda f.f(a)$
(<T)	backward type-raising		$X : a$	$\rightarrow$	$T\backslash(T/A) : \lambda f.f(a)$

Variants of CCG come with specialized rules, e.g. for conjunction, or generalized composition (Steedman & Baldridge to appear). We will only consider the rule set given above in this paper. Crossing rules essentially make CCG a mildly context-sensitive grammar formalism (ibid.). That is why I include them here. I will not use type raising productively here, but pre-compile the necessary type raises into the lexical categories.

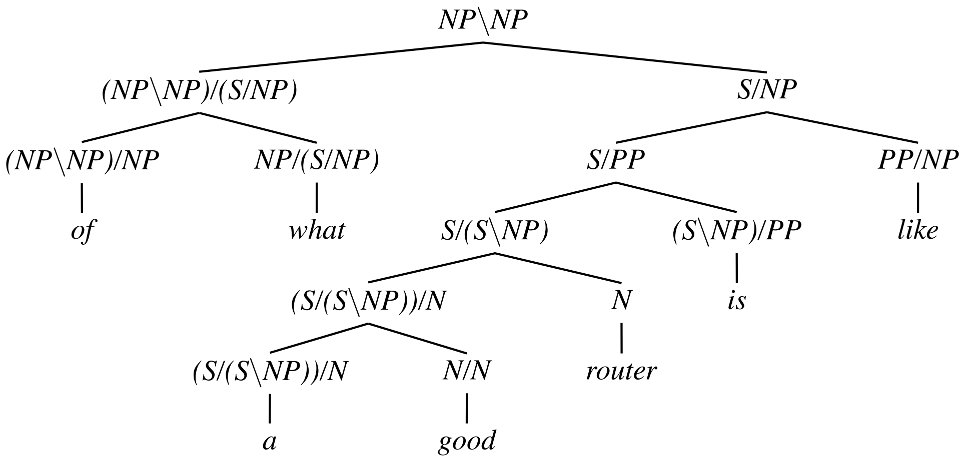
2.2. Derivations in CCG

A derivation in CCG is a sequence of rule applications, which turns an initial string of lexical categories (a sentence) into a single resulting category. For any given input string, there may be more than one successful derivation (if any), conceptually due to two different reasons. The first one is syntactic ambiguity; if the sentence has more than one reading, we want a different derivation for each. The second one is spurious ambiguity: Two derivations are *equivalent* if they correspond to the same λ-term, after β-conversion.

Example 1. *A* normal *and an equivalent* incremental *derivation tree in CCG. For λ-terms, see example 4. The original type of 'a', NP/N, is internally raised to $(S/(S\backslash NP))/N$. 'What is your opinion ...' (from a CCG annotation of the Switchboard corpus (Reitter, et al. 2006))*



[1]Hypothetical reasoning as in related calculi finds its counterpart in the compositional rules.

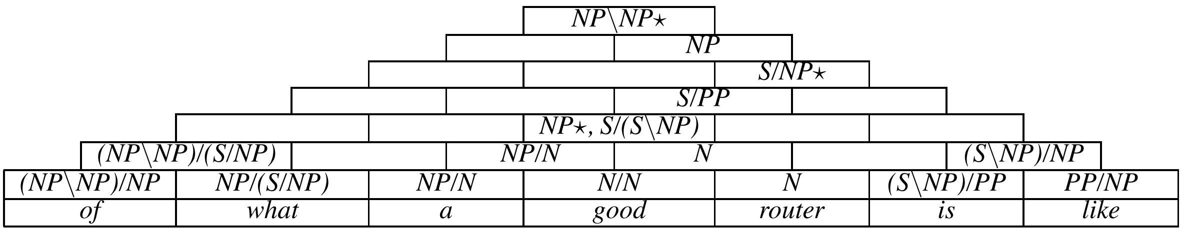


The two derivations shown are constructed to be as left- or right-branching as possible, corresponding to an incremental derivation or to phrase structure (in English), respectively. For example, ‘a good router’ is constructed via application only in the normal form, but with composition and application in the incremental form.[2] Between the two example derivations, there is a variety of equally valid derivations.[3] Note that CCG does not employ traces, and therefore the fact that ‘what’ also serves as an argument to ‘like’ escapes the eye.[4]

It is desirable to abstract over this spurious ambiguity, to encode the *result* without the *process* of the derivation. This is not accomplished by a normal form, which is unique, but does denote the application of certain rules where several choices are possible, as shown for ‘a good router’.

In CYK-parsing of context-free grammars, a *chart* abstracts over spurious ambiguity. Chart-parsing techniques are also used for CCG, and for our example they yield the following chart (2).

Example 2. *The chart for our example.*



I shall not prove soundness and completeness of the chart presentation here, as the following argument renders its use impractical. As it is, there is no information on how a cell was derived. If I include it, the ambiguity for the categories marked with ⋆ is retained. Neither is satisfactory: we want both unambiguous and complete information.

In linear logic, this is given in the form of *proof nets* (Girard 1987). In the next section I show how to convert a CCG derivation into a proof net, and how proof nets capture syntax and semantics transparently.

---

[2]A rule of generalized composition would allow total incrementality here. It would also fill more cells in the chart (example 2).

[3]Reitter et al. (2006) suggest the incremental analysis comes closer to human sentence processing, especially under time pressure.

[4]CCGbank (Hockenmaier & Steedman 2007) captures these directly in an additional dependency structure. This actually reveals more dependencies than seen in proof nets.

3. Proof nets

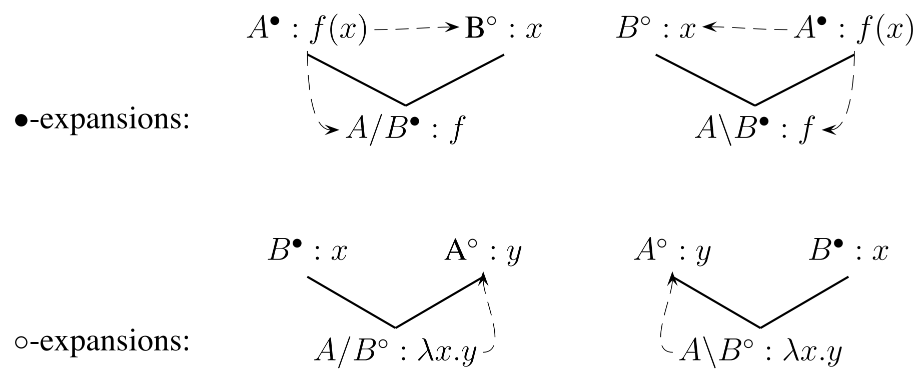
A proof net is a visualization of a proof in linear logic (Girard 1987, Roorda 1991). It consists of two parts: a list of tree expansions of the lexical categories, and a pairwise linking of their leaves. This linking has to satisfy several correctness criteria, depending on the logic.

Proof nets can be used for categorial grammars, if their derivations are valid proofs. Here I identify properties of proof nets for CCG, calling them mildly non-planar proof nets. Crossing composition leads to non-planar nets. The Curry-Howard correspondence carries over.

3.1. Definition and construction of proof nets

From Roorda (1991) and Morrill (2000), I repeat the following basic definitions, adapted to Steedman's notation. Further definitions which are needed later are also included.

Category trees are binary branching trees. Each node carries a category labeled with input (•) or output (∘) polarity, and a λ-term. They are recursively expanded from lexical categories (along with lexical semantics) by the following rules. The dashed lines denote steps (up, down, left, right) on the *semantic trip* (Carpenter & Morrill 2005, Lamarche 1994).



Every lexical category has input polarity. Following the path from its root only over •-nodes we reach the *head leaf* of the tree.

A *proof frame* is a sequence of category trees with roots A°, $A_1^\bullet$, …, $A_n^\bullet$ where A is the result category and A_1 to A_n are the lexical categories of words w_1 to w_n in a sentence. The result receives output polarity.

An *axiom linking* is an exhaustive, pairwise matching of leaves of category trees, such that each leaf A° is connected to a leaf $A^\bullet$. Each axiom link denotes substituting the semantic variable of the ∘-node with the λ-term of the •-node. The semantic trip step (left or right) is from A° to $A^\bullet$. A *proof structure* is a proof frame with an axiom linking on its leaves. A *proof net* is a proof structure satisfying two criteria:

(a) The axiom linking is planar (no crossing links).

(b) Every cycle[5] (over axiom links and edges of the rule expansions) crosses one $\circ$-expansion (Morrill 2000).

Planarity is required for Lambek calculus (Lambek 1958) and similar categorial grammars, but not for linear logic in general. Let a proof structure satisfying (b) but not (a) be a *non-planar proof net*.

Semantics Starting from the λ-term of the result node, and following the semantic trip, all substitutions denoted by the links are carried out. (b) guarantees that this results in a valid λ-term. A logical form (LF) tree depicts both the semantic trip and the unreduced λ-term (see example 4). We call links to head leaves *primary* and all others *secondary*. Primary links build the LF tree, secondary links are interpreted as abstractor (binder) and variable (trace); see the curved arrows in the example.

The tree retains all directions, resulting in a directed λ-calculus; cf. Wansing (1992). Planar proof nets translate to LF trees retaining the linear order of words. Non-planar proof nets either violate the linear order, or require crossing branches.

3.2. Turning a CCG derivation into a proof net

Derivations in a categorial grammar which correspond to proofs in a type calculus (Lambek (1958) and descendants) are straightforward to turn into proof nets. Constructing the proof net for a given CCG derivation is as follows. First we convert the derivation to a **raising**-free form by raising the necessary types already in the lexicon, as in example 1. We then construct the proof frame and find the partial linking.

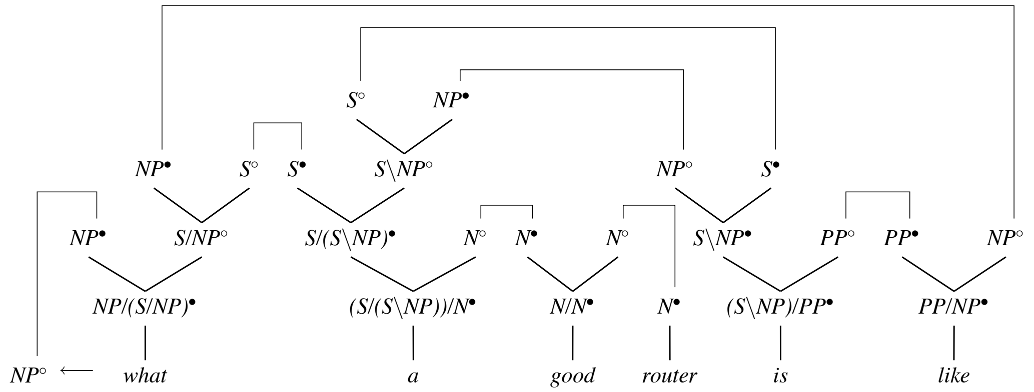
Each instance of an **application** or **composition** rule identifies a pair of (possibly complex) categories B° and $B^\bullet$. As an abbreviation, we might connect them directly (see especially ex. 5). In full: The sequence of terminal nodes from expanding B° is that from $B^\bullet$ with linear order and polarities reversed. Each terminal is linked to its counterpart, yielding a nested linking.

Crossing rules "bury" nodes under their links, because the two terminal sequences are not adjacent. To connect the buried nodes, links have to cross (also see ex. 5). Thus planarity (a) is not satisfiable for CCG.

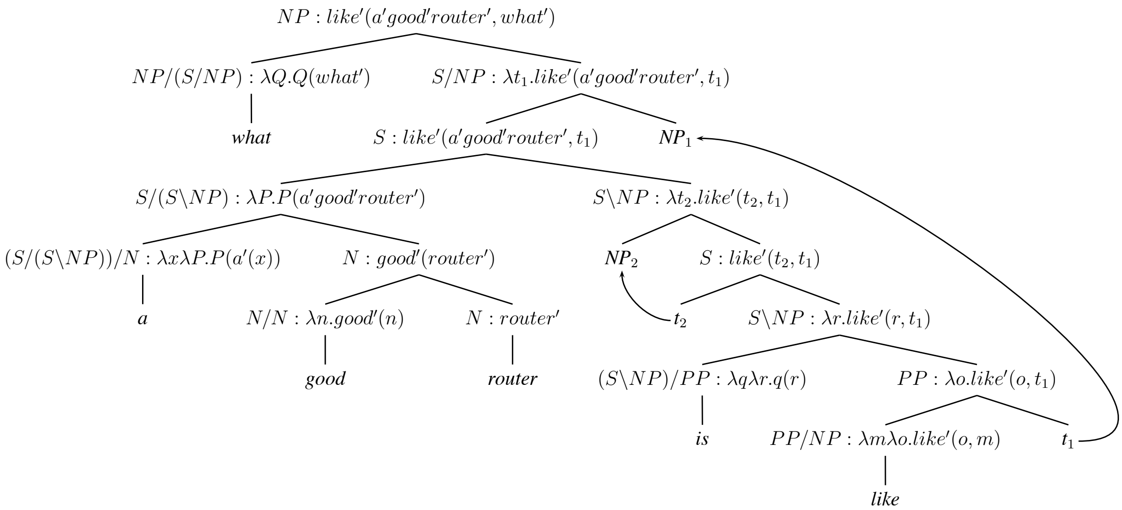
The remaining open category $X^\bullet$ is connected to the **result** node, which then is X°. The same holds for complex categories as said before. Then there are no more open terminals; the net is complete.

Example 3. *The sentence from example 1 as a proof net (minimally abridged). It abstracts over which CCG rules were used. Thus it depicts the normal form, the incremental form (see example 1), and all possible others simultaneously. The long distance dependency is manifested in a direct link (from the content-NP$^\bullet$-node of 'what' to the argument-NP$^\circ$-node of 'like').*

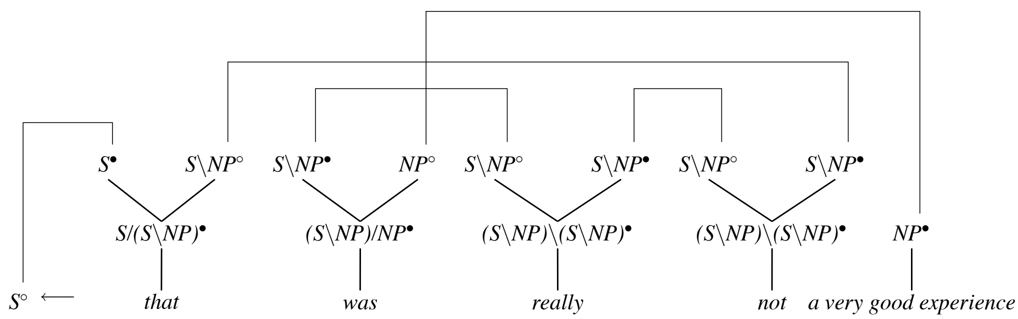
[5]Not to be confused with paths on the semantic trip, which does not allow cycles.



Example 4. *The following tree shows the semantic trip for the example sentence. Primary links, which are unary branchings, are directly substituted; secondary links appear as arrows. Category labels are inherited from the CCG parse / the proof net. Pre-terminals are annotated with (sketched) lexical λ-terms. The tree's bracketing structure over the lexical λ-terms reads as the unreduced λ-term for the sentence. I have added reduced λ-terms for non-terminal nodes.*



Example 5. *This net depicts a CCG derivation with $<B_\times$, for a sentence from the CCG-annotated Switchboard corpus (Reitter et al. 2006). Depending on the derivation, $<B_\times$ is either used to combine 'was' with 'really' and again to add 'not'; or to combine 'was' with 'really not' (constructed via composition). The proof net denotes both, and in any case $<B_\times$ leads to crossing links.*



3.3. Properties of CCG proof nets

Characteristic for combinatory rules is the identification of a ○-node with a ●-node. The ●-node itself might arise from functional composition, so it is not represented by a single node in the net. I capture this in criterion (c). Directionality is expressed by a bipartite criterion (d).

(c) Every ○-expansion is part of a cycle (Morrill 2000), and all those cycles are *balanced*: Their traversal involves going up an edge as many times as going down. See e.g. the S/NP°-node in the expansion of 'what' in ex. 3, and follow the circles it is part of, counting steps up and down. In the LF tree this amounts to well-nestedness of abstractions: the arrows in ex. 4 do not cross.

(d′) Primary links have the same direction as the preceding left or right step (ignoring any intervening steps up) on the semantic trip. That is to say, arguments wanted on the right (left) are on the right (left).

(d″) Secondary links go from a right daughter in an expansion to a left daughter, or vice versa. In the LF tree this makes binders and traces both left (or both right) daughters (cf. example 4), because order is reversed in ○-expansions.

I call proof structures satisfying (b)–(f) *mildly non-planar proof nets*. They are less restricting than planar proof nets, yet more restricting than proof nets in general. My choice of combinatory rules in section 2.1. restricts proof nets for CCG even more. Let *reachability* (of A from B) denote existence of a path on the semantic trip (from B to A). On LF, reachability is dominance. Let a *gap* be a sequence of root nodes spanned over by a (primary) link α. Let the *governing node* A for α be defined as the node reachable via the longest path from α such that every step up has to be cancelled out by a later step down. There are two restrictions on gaps.

(e) Coherence: All roots in the gap are reachable from the governing node. This might be via α, but:

(f) If a (primary) link β, reachable from α, leads to a word in the gap (that is, α and β cross), α and β must not have the same direction.

Criteria (b)–(f) define *CCG proof nets*, for the variant of CCG used here.

Proposition 1. *Every CCG derivation yields a CCG proof net, that is, a net satisfying (b)–(f).*

The proof follows the construction of the net. Furthermore all criteria are designed to be valid for (this variant of) CCG. We omit it here, for sake of more space for the proof for the following proposition.

Proposition 2. *Every CCG proof net denotes a valid CCG derivation.*

Proof. The proof proceeds in a constructive, bottom-up way. I first prove the existence of two adjacent words connected via a primary link; then prove that they in fact are combinable by a rule; and finally treat them as one word to continue the argument recursively, thus proving that there is a valid CCG derivation for any CCG proof net.

1. Primary links form a directed acyclic graph G over the words. Take the path H given by starting at any leaf of G and going back until encountering a branching (a word with more than one argument), or the root of G. By condition (e) on gaps the links in $G \backslash H$ never cross those in H, so they do not account for non-adjacency of linked words in H.

Assume H without adjacent connected words. Without loss of generality, assume the first link α goes right, leaving a gap x. If the next link β goes right again, call the new gap x and continue there. If β goes left, spanning over α, x is rendered unreachable: Because of (f), no succedent link may cross α or β. Thus β goes left and stays within x, leaving a new gap y. The next link γ renders y unreachable, if it leaves x, in either direction. Continue with γ as the new β, β as the new α and a new gap y. Each new link either leaves a containing link, rendering the remaining gaps unreachable, or opens a new gap. Therefore, no finite graph H without adjacent, connected words exists.

2. Now I prove the combinability of adjacent words a and b connected via a primary link α. By construction of H, the argument side of α ($A^\circ \rightarrow A^\bullet$) may have maximally one argument itself, so $A^\bullet$ is the head of $B_2^\bullet$, which is either the category of b or of the head-daughter thereof. By definition of primary links, A° is the head of the non-head daughter B_1° of a. By virtue of (c) not only the cycle over B_1° and $B_2^\bullet$ is balanced, but also all cycles over parts of these categories, if they are complex. This amounts to perfect symmetry: $B_1 = B_2$. Thus a and b combine via functional application or composition.

3. Now reduce those two words into a single category tree according to the combination rule, inheriting all remaining links. Treating two adjacent words as one does not add any nodes or links to the net, and it retains all relative positions. Hence the net still satisfies (b)–(f).

This argument applies recursively, until there is only one word left, to be combined with the resulting category tree. We have thus found a CCG derivation for an arbitrary CCG proof net. □

Proposition 3. *CCG proof nets stand in a bijection to equivalence classes of CCG derivations.*

Proof. All equivalent derivations are mapped to the same net. Assume the opposite. The proof frames are the same, so the nets differ in their linking. Different linkings denote different semantics by definition. Contradiction to the derivations' equivalence.

The net denotes *all* equivalent derivations because of prop. 1 □

4. Conlusions and future work

Proof nets abstract over derivational ambiguity in categorial grammars. A proof net is an unambiguous presentation of the linear order, the dependencies (by direct links), and scope relations (by precedence on the semantic trip) of a sentence. Syntactic structure and semantics stand in (Curry-Howard) correspondence.

I have identified the properties of proof nets for a certain mildly context-sensitive, yet almost arbitrary variant of CCG, given by the rule set in section 2.1.. Accordingly, I call them mildly non-planar. Not too surprisingly, this notion is not as natural a subclass of proof nets compared to the notion of planarity. Future research may reveal a landscape of subclasses of proof structures corresponding to the variants of categorial grammars. Special interest lies on multimodal CCG. Also, a better treatment of type raising is desirable.

The work reported here is part of joint work with Christian Pietsch (University of Bielefeld) and our supervisor Gerhard Jäger. We have compiled a preliminary proof net corpus from a CCG annotation of the Switchboard corpus (the one used by (Reitter et al. 2006)), containing about 65,000 sentences of spoken language. A conversion of CCGbank (Hockenmaier & Steedman 2007) is planned. In subsequent research we will utilize it in data-oriented methods (e.g. for parsing) by defining reasonable sub-structures of the nets, exploiting the direct links between grammatically dependent words. We will also investigate syntactic priming between these sub-structures.

Acknowledgements

I would like to thank my supervisor Gerhard Jäger, Christian Pietsch and two anonymous ESSLLI student session reviewers for helpful suggestions and corrections. All remaining errors are mine. This research has been supported by the EC 277 “Cognitive Interaction Technology” of the DFG at the University of Bielefeld.

References

B. Carpenter & G. Morrill (2005). ‘Switch Graphs for parsing Type Logical Grammars’. In *Proceedings of the International Workshop on Parsing Technology, IWPT05, Vancouver*.

J.-Y. Girard (1987). ‘Linear Logic’. *Theoretical Computer Science* **50**:1–102.

J. Hockenmaier & M. Steedman (2007). ‘CCGbank: A Corpus of CCG Derivations and Dependency Structures Extracted from the Penn Treebank’. *Computational Linguistics* **33**(3):355–396.

F. Lamarche (1994). ‘Proof nets for intuitionistic linear logic: Essential nets’. Tech. rep., Imperial College, London.

J. Lambek (1958). ‘The mathematics of sentence structure’. *American Mathematical Monthly* **65**(3):154–170.

G. Morrill (2000). ‘Incremental processing and acceptability’. *Computational Linguistics* **26**(3):319–338.

M. Pentus (1993). ‘Lambek grammars are context free’. In *Proceedings of the Eighth Annual IEEE Symposium on Logic in Computer Science, Montreal*.

D. Reitter, et al. (2006). ‘Priming effects in Combinatory Categorial Grammar’. In *Proc. EMNLP*, pp. 308–316.

D. Roorda (1991). *Resource Logics: Proof-theoretical investigations.* Ph.D. thesis, Universiteit van Amsterdam.

M. Steedman (2000). *The Syntactic Process*. MIT Press.

M. Steedman & J. Baldridge (to appear). 'Combinatory Categorial Grammar'. In R. Borsley & K. Borjars (eds.), *Constraint-based approaches to grammar: Alternatives to transformational syntax*. Oxford: Blackwell.

H. Wansing (1992). 'Formulas-as-types for a hierarchy of sublogics of intuitionistic propositional logic'. In D. Pearce & H. Wansing (eds.), *Nonclassical Logics and Information Processing*, Springer Lecture Notes. Springer, Berlin.