NATURAL LOGIC: A VIEW FROM THE 1980S

Johan van Benthem, ILLC Amsterdam & Stanford, http://staff.science.uva.nl

March 2008

Abstract

This paper describes the main lines of the 'natural logic' research program on direct reasoning with natural language in the 1980s done by the author and colleagues in the formal semantics community. It also identifies and discusses some main challenges to the program given what we know today.

Summary

'Natural Logic' is a somewhat loose, but popular and suggestive term for recurrent attempts over the last decades at describing basic patterns of human reasoning directly in natural language without the intermediate of some formal system. One main type of inference (though not the only one) where this idea works well is so-called 'monotonicity reasoning', involving valid replacement of predicates by predicates with smaller or larger extensions. Essentially, this pattern of inference goes back to the distribution doctrine of traditional logic. With varying motivations, ideas for a simple 'surface-syntax' based calculus of reasoning have been rediscovered many times, in such diverse areas as logic, philosophy, linguistics, computer science, and nowadays also cognitive science. In particular, the natural logic program had a certain flowering, going beyond the ubiquitous example of plain monotonicity reasoning, in the formal semantics community in The Netherlands in the 1980s. This paper is the text of a survey lecture given in May 2007 at the request of some colleagues at Stanford University and PARC Research Center, who are working on practical information retrieval from texts by combinations of parsing and simple logical inference. Before re-inventing the wheel (a very common phenomenon in this area), it seems useful to see what rolling stock existed in that period. With a slightly more historical slant, the following text was also presented at the LORI Workshop in Beijing 2007. Its purpose is mainly to record the main lines of the work on natural logic done in the 1980s in my own Dutch logic and language environment, and also, to present the broader issues that seemed relevant to us then, and are still so today, even in a broader interdisciplinary environment.

1 Introduction: from classical to modern logic

For beginning logic students in Amsterdam arriving in the late 1960s like the present author, the following 'standard example' in our introductory course was supposed to show once and for all how the modern 19^{th} century logic of Boole and Frege came to supersede traditional logic. De Morgan's famous inference runs as follows:

> *"All horses are animals.* So, *all horse tails are animal tails."*

This is supposed to show the inadequacy of the traditional logic of 'monadic predicates', because binary relations are essential to understand the validity of the inference. And the latter are brought out in the standard first-order logical forms that students are trained in:

$$\forall x\, (Hx \rightarrow Ax) \;\;|= \;\; \forall x\, ((Tx \;\&\; \exists y\, (\boldsymbol{H}y \;\&\; Rxy)) \rightarrow (Tx \;\&\; \exists y\, (\boldsymbol{A}y \;\&\; Rxy))).$$

What is more, we can understand the general phenomenon behind this valid first-order inference as follows. Syntactically, we are replacing a predicate *'horse'* by one with a larger extension *'animal'* at some position in an assertion: *'having a – tail'*. And what licenses this 'upward replacement' in this context is the following semantic property:

Definition A formula $\varphi(X)$ is *upward monotone* with respect to the predicate X if for all models $\boldsymbol{M}$, if $\boldsymbol{M}, \boldsymbol{P}, \boldsymbol{s} \;|= \varphi(X)$ [here we interpret the syntactic predicate X as the set $\boldsymbol{P}$, while $\boldsymbol{s}$ is a tuple of objects whose length equals the arity of X], and $\boldsymbol{P} \subseteq \boldsymbol{Q}$, then $\boldsymbol{M}, \boldsymbol{Q}, s \;|= \varphi(X)$.

This property is ubiquitous in first-order logic, but also beyond it in generalized quantifier theory, and the model theory of inductive definitions in logics with fixed-point operators. [1]

Actually, there is another way of looking at the same notion which is closer to the original inferences. [2] One can also think of a formula $\varphi(P)$ with P already interpreted, and then view the inference as replacing this *'P'* by some concrete predicate *'Q'* with a larger extension. While perhaps a bit less rigorous, this is how most people understand upward monotonicity in practice, and we will talk this way henceforth. Incidentally, it will be clear that there is also a *downward* version of the preceding notions, allowing for replacement by stronger predicates having a smaller extension. If you own no animals, you own no horses.

This semantic behaviour has a syntactic counterpart, as may be seen in the occurrence of the predicate *'H'* in the first-order formula $((Tx \;\&\; \exists y\, (\boldsymbol{H}y \;\&\; Rxy))$. Let us call an occurrence of X in $\varphi(X)$ *positive* if it lies in the scope of an even number of negations, or stated differently, if the formula $\varphi(X)$ is created using only the following inductive syntax rules:

$$H\text{-free formulas} \;|\; \& \;|\; \vee \;|\; \forall \;|\; \exists.$$

The occurrence of *'H'* in $((Tx \;\&\; \exists y\, (\boldsymbol{H}y \;\&\; Rxy))$ is positive in this sense, and in a natural extended sense, so is that of the predicate 'horse' in the expression 'tail of a horse'.

---

[1] Often, monotonicity is defined looking at all occurrences of X in $\varphi(X)$. But our definition also works for just one specific occurrence. The single-occurrence based version seems more natural for seeing the fine-structure of inference. Monotonicity inferences can be executed occurrence by occurrence, with no need for more complex 'simultaneous replacements'.

[2] The observations made in this section may be found with many authors since the 1960s, witness Prior 1967, Barwise & Cooper 1981, and the sources in Sanchez Valencia 2004.

Now, it is easy to see that syntactic positive occurrence implies semantic monotonicity (a sort of 'soundness', if you wish), by a straightforward induction on the construction of formulas. [3]But it is much less trivial that the converse 'completeness' direction holds. But it does, witness a well-known model-theoretic result from the 1950s:

Lyndon's Theorem A first-order formula $\varphi(X)$ is semantically monotone in X iff $\varphi(X)$ is equivalent to a formula whose only occurrences of X are positive.

Lyndon's Theorem does not hold for arbitrary extensions of first-order logic, but the above soundness property is quite general: positive occurrence does imply monotonicity in many higher-order logics – and as we shall see, also in natural language. Thus, modern logic provides the right forms for the above family of inferences [4], and it backs these up with important meta-theorems probing the broader extent of the phenomenon.

2 Distribution in traditional logic

But the above De Morgan story is misleading and historically false. Inferences like the one with the horse tail were well within the scope of traditional logic, which was much subtler than many modern critics acknowledge. They blame it for defects it never had – all the way to the invective of Geach 1972, who even saw demoniacal political consequences, in his phrase 'The Kingdom of Darkness'. Indeed, monotonicity inference is closely related to the *Aristotelean Syllogistic*, the main tool of traditional logic. Now the latter is often viewed as a trivial theory of single-quantifier inferences with patterns *Q AB* on unary predicates *A, B*: at best a 'fragment of monadic first-order logic'. But this modern view in terms of formal systems does no justice to how the Syllogistic really functioned: as a method for *one-step analysis* of statements of any kind into one layer of quantification. In particular, *A, B* could be predicates with further structure, of whatever expressive complexity (first-order, higher-order, etc.): they are not constrained at all to one fixed formal language. This point may be worth emphasizing. The modern *bottom-up* view of inference as involving formulas constructed explicitly out of atoms is far removed from logical history and from the way we actually reason. We work *'top-down'*, unpack some surface-level quantificational patterns, the fewer the better, and then reason on the basis of those. [5]

Points like the preceding have been made before, and I do not pretend to give a complete history here. The reader is referred to Curry 1936, Prior 1967, for some early sources, Suppes 1982 for an original computational proposal linking up with automated learning, and

---

[3] Again, results for 'negative occurrence' and downward monotonicity are immediate.

[4] It also deals in a straightforward dual fashion with 'downward monotonic' counterparts.

[5] Incidentally, the top-down view, viewing formal systems as tools for analysis rather than synthesis from atomic components, might be more effective for *teaching* logic as well.

Sanchez Valencia 2004 for a more general history. This section is just a brief summary, including my private views about the border line between traditional and modern logic.

Later on, in addition to the syllogistic base system, the Medieval Scholastics developed the so-called *Doctrine of Distribution*, a general account of contexts $\varphi(P)$ where the statement was either about 'all of *P*' (the 'Dictum de Omni'), or about 'none of *P*' (the 'Dictum de Nullo'). Again, these contexts could be of any sort of linguistic complexity, where the expression φ might include iterated quantifiers such as "Someone loves every human", or even high-order constructions. Indeed, I would claim that, for the purpose of analyzing ordinary human inference, the modern 'first-order/higher-order' boundary is mainly a mathematical 'systems concern' without any clear matching jump in natural reasoning. However this may be, authors like Van Eijck 1982, van Benthem 1986, Sanchez Valencia 1991, and Hodges 1998 have pointed out in more detail how the Distribution principle called "Dictum de Omni et Nullo", corresponded to admissible inferences of two kinds: *downward monotonic* (substituting stronger predicates for weaker ones), and *upward monotonic* (substituting weaker predicates for stronger ones). Traditional logic investigated these phenomena for a wide range of expressions, without any boundary between unary and binary predication – another artefact of viewing history through predicate-logical glasses.

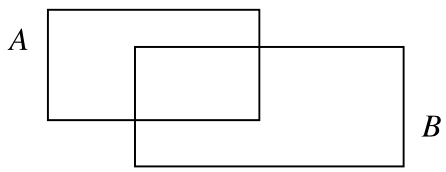
To be sure, this does not mean that all was clear. To the contrary, traditional logic had a major difficulty: providing a systematic account of complex linguistic constructions from which to infer, and in particular, despite lots of valid insights, it wrestled with a good general account of *iterations of quantifiers*. Dummett 1973 makes a lot of this, by saying that Frege's compositional treatment in terms of merely explaining single quantifiers and then letting compositionality do all the rest "solved the problem which had baffled traditional logicians for millennia: just by ignoring it". Again, while there is a kernel of truth to this, there is also a good deal of falsehood. Indeed, as the extensive historical study Sanchez 2004 remarks, it seems more fair to say that De Morgan represents a low point in logical history as far as understanding the scope of monotonicity reasoning is concerned. Things got better after him – but as the author points out tongue-in-cheek, they also got better and better moving back in time to Leibniz and then on to the Middle Ages...

Perhaps not surprisingly then, traditional logicians felt some unfairness when modern logic arrived on the scene, since it attacked a caricature of traditional logic. And indeed, until late in the 20th century, attempts have been made to further develop the Syllogistic into a full-fledged calculus of monotonicity reasoning, witness Sommers 1982 (going back to versions from the 1960s), and Englebretsen 1981. The claim of these authors was that this enterprise provided a viable alternative to first-order logic for bringing out key structures in actual human reasoning in a more congenial way. Still they did not propose turning back the clock altogether. E.g., Sommers' book is up to modern standards in its style of development, providing a systematic account of syntactic forms, an arithmetical calculus for computing positive and negative syntactic occurrence, as well as further inferential schemata generalizing traditional inference patterns like Conversion and Contraposition. While this has not led to a Counter-Revolution in logic, these points have resonated in other areas, such as linguistics, computer science, and recently also, cognitive science. We now move there.

3 Monotonicity in natural language and generalized quantifier theory

Very similar issues concerning the analysis of reasoning came up in the 1970s and 1980s when linguists and logicians started looking at natural language together (there are several famous joint papers, including Barwise & Cooper 1981) with fresh eyes, following Richard Montague's pioneering work (Montague 1974). [6] Suddenly, natural language was no longer 'misleading' as to the correct logical forms, but rather a gold-mine of intriguing insights with remarkable staying power. After all, no one, not even pure mathematicians, has ever seriously switched to predicate logic as a tool of reasoning (but see below). In particular, Montague provided a categorial/type-theoretic analysis of *quantifier expressions* Q as taking linguistic noun phrases A to noun phrases QA that denote properties of properties B. Deconstructing this forbidding phrasing somewhat, quantifiers may then be viewed semantically as denoting binary relations between predicates, on the pattern

$Q\,AB$



E.g., in this Venn-style diagram format, "All A are B" says that the area $A–B$ is empty, "Some A are B" that the intersection $A \cap B$ has at least one object in it, while the more complex "Most A are B" says that the number of objects in $A \cap B$ exceeds that in $A–B$. Moreover, in addition to judgments of (non-)grammaticality, judgments about valid and invalid inference were now considered relevant to our understanding of natural language, and even linguists admitted that semantic theories had to account for these.

Generalized quantifiers and monotonicity More semantic depth was provided beyond the mechanics of Montague Grammar by *Generalized Quantifier Theory*: a research program which charted the variety of the quantifier repertoire of natural human languages (cf. Keenan & Westerståhl 1997), and tried to formulate general laws about expressive power. Again, Monotonicity turned out to play a crucial role across this repertoire. One influential example here was the observation by Bill Ladusaw that so-called 'negative polarity items' like "at all", or "ever" flag 'negative contexts' in linguistic expressions which allow for downward monotonic entailments from predicates to sub–predicates:

---

[6] For the state of the art, see various chapters in van Benthem & ter Meulen, eds., 1997.

"If you ever *feel an ache*, I will cure it" implies
"If you ever *feel a headache*, I will cure it".

As a rule, negative polarity items do not occur in upward entailing positive contexts.

Here are some general facts about monotonicity for basic quantifiers: they can be either upward or downward, in both arguments. E.g., the quantifier "All" is downward monotonic in its left-hand argument, and upward in its right-hand argument, exemplifying the patterns

$\downarrow MON$ *if* $Q\ AB$ *and* $A' \subseteq A$, *then* $Q\ A'B$
$MON\uparrow$ *if* $Q\ AB$ *and* $B \subseteq B'$, *then* $Q\ AB'$

It is easy to exemplify the other three possible combinations: e.g., "Some" is $\uparrow MON\uparrow$. By contrast, a quantifier like "Most" is only $MON\uparrow$, being neither 'down' nor 'up' on the left. Quantifiers with monotonicity ($\uparrow$ or $\downarrow$) in both arguments are called 'doubly-monotonic'.

Conservativity But Monotonicity is not the only key property found with natural language quantifiers (and *NP*-forming determiner expressions in general, such as "Mary's"). Here is a further principle which shows how the first argument sets the scene for the second:

Conservativity $Q\ AB$ iff $Q\ A(B \cap A)$

Conservativity seems to hold in all human languages. One can think of this as a sort of domain or role restriction imposed by the initial predicate *A* on the predicate *B*. More generally, the nouns in sentences give us relevant domains of objects in the total universe of discourse, and quantifiers impose a sort of coherence on the total predication expressed.

Taken together, the preceding semantic properties explain what makes particular (logical) notions so special. Generalized Quantifier Theory then charted what sort of expressions pass these tests, getting a grip on what natural languages can say. The following result from van Benthem 1986 is a typical sample, and it shows that the traditional quantifiers may be viewed as the simplest level of *conservative, inference-rich* linguistic expressions, provided we add one more technical condition saying that the quantifier makes maximal distinctions:

Variety   If $A \neq \emptyset$, then $Q\ AB$ for some *B*, and $\neg Q\ AC$ for some *C*.

Theorem The quantifiers "All", "Some", "No", "Not All" in the Square of Opposition are the only ones satisfying Conservativity, Double Monotonicity, and Variety.

But natural language quantifiers can also be classified in other ways, for instance, by means of more algebraic types of inferential property for specific lexical items. A typical example is the rule of 'Conversion' already found in traditional logic:

Symmetry $Q\ AB$ iff $Q\ BA$.

This holds typically for expressions like “Some”, “At least *n*”, “No”, “All but at most *n*”. Characterizations exist of all quantifiers satisfying Symmetry and other basic properties.

With this bare minimum, we conclude our survey here. For much more detailed information on Generalized Quantifier Theory and its richer agenda, cf. earlier sources like van Benthem 1986, Keenan & Westerståhl 1997, or the recent monograph Peters & Westerståhl 2006.

4 The ‘natural logic’ program of the 1980s

In the 1980s, the idea arose that the preceding observations had a more general thrust. Natural language is not just a medium for saying and communicating things, but it also has a ‘natural logic’, viz. a system of simple modules capturing ubiquitous forms of reasoning that can operate directly on natural language surface form without the usual logical formulas. This idea was developed in some detail in van Benthem 1986, 1987, whose main proposals we outline here. [7]The main ingredients were to be three modules:

(a) Monotonicity reasoning, or *Predicate Replacement*,
(b) Conservativity, or *Predicate Restriction*, and also
(c) Algebraic laws for inferential features of specific lexical items.

But of course, there are many further natural subsystems in natural language, including reasoning about collective predication, prepositions, anaphora, tense and temporal perspective. The systematic challenge is then to see how much of all this inference can be done directly on natural language surface form, and we will look at some details below. Another challenge might be how these subsystems manage to *work together* harmoniously in one human mind, and we will return to this somewhat neglected issue below.

Notice how this way of thinking cuts the cake of reasoning differently from the syntax of first-order logic – redrawing the border-line between traditional and modern logic. E.g., monotonicity inference is both richer and weaker than first-order predicate logic. It is weaker in that it only describes part of all valid quantifier inferences, but it is richer in that it is not tied to any particular logical system, as we observed above (it works for second-order just as well as first-order logic). One intriguing aspect then becomes *where the surplus of first-order logic is really needed*. We will give two possible answers to this below.

5 Compositional structure, parsing, and free rides for inferences

To get a natural logic going, it is not enough to display single-quantifier inferences of the sort we had before. One also needs to account for the way in which these inferences play in arbitrary *complex sentences*, and here is how this may be done in general, merging ideas

[7] Later on, Sanchez Valencia 1991 found an interesting ancestry in the work of C.S. Peirce.

from Generalized Quantifier Theory and Categorial Grammar. The earliest source for this calculus seems van Benthem 1986, while our exposition mainly follows van Benthem 1991.

Spreading conservativity A first example concerns the broader impact of Conservativity. Consider an iterated quantifier sentence of the sort "Every man loves a woman":

$Q_1A\ R\ Q_2B$

Clearly, *both* predicates A and B should have restriction effects. How doe they do this? It is not hard to see, and in fact it can be computed in many parsing formalisms, that we have

$Q_1A\ R\ Q_2B \quad \textit{iff} \quad Q_1A\ R \cap (AxB)\ Q_2B$

That is, the first predicate restricts the first argument of the binary relation R, while the second predicate restricts the second argument. Based on this and other cases, the semantic and inferential mechanism behind Conservativity might be called *Predicate Restriction*, a first major aspect of natural logic that seems to be at work all through natural language:

Nouns constrain correlated predicate roles.

Monotonicity in complex sentences Likewise, to get a full-fledged account of how monotonicity inference functions in complex sentences, we need a grammatical theory that provides an analysis of hierarchical syntactic structure. Just flat strings of words will not do. For instance, I cannot tell whether the occurrence of "women" in

"Johan admires some men and women"

is downward monotonic unless I resolve the scope of the determiner "some". To achieve disambiguation, we can use a logic-friendly grammar formalism, viz. *Categorial Grammar* to obtain a systematic *Monotonicity Calculus*. For a start, consider the sentence

"No mortal man can slay every dragon."

We would like compute all predicate markings in this sentence, but they cannot be taken at face value. E.g., whether "dragon" is negative depends on the scope of other expressions inside which it occurs. Maybe the reader wants to check that, on the narrow scope reading for "every dragon", [8] readings should come out (intuitively) as follows:

– +
"No mortal man can slay every dragon."

E.g., it follows that no mortal *Dutchman* can slay every dragon, or that no mortal man can slay every *animal*. For the wide scope reading of 'every dragon" (more artificial, but still acceptable to many people [9]) these markings should come out as follows:

---

[8] Think of the first-order formula $\neg\exists x\ (Mortal\text{-}Man(x)\ \&\ \forall y\ (Dragon(y) \rightarrow Can\text{-}Slay(x, y)))$.

[9] Now think of the formula $\forall y\ (Dragon(y) \rightarrow \neg\exists x\ (Mortal\text{-}Man(x)\ \&\ Can\text{-}Slay(x, y)))$.

– –
"No mortal man can slay every dragon."

But we can be more general. Perhaps surprisingly, other expressions than predicates can also be marked here, when we take a suitably general perspective on sentence construction. For instance, the sentence "*No* mortal man can slay every dragon" clearly implies that "*No or very few* mortal men can slay every dragon". Here the determiner "No or very few" is intuitively weaker than the determiner "No", just as the predicate "animal" is weaker than "dragon". Thus, like predicates, determiners themselves allow for monotonic replacement by suitable items in their linguistic category. This phenomenon is totally general, and the requisite notion of inclusion in arbitrary categories was made precise in the monotonicity calculi of the 1980s, which could handle inferences in arbitrary kinds of expression. Of course, this is best checked against some intuitions. We invite the reader to check that, on the narrow scope reading of "every dragon", markings should come out as follows:

\+ – – – – – +
"No mortal man can slay every dragon."

Categorial monotonicity calculus Speaking generally, we need a linguistic mechanism marking positive/negative occurrences in any category in tandem with the syntactic analysis of given expressions. As it turned out, this can be done quite elegantly in a *categorial grammar*, of the Ajdukiewicz function application type, or the more sophisticated Lambek type which is more like a simple system of function application plus a limited additional operation of ('single-bind') lambda abstraction. For details, we refer to van Benthem 1991, Sanchez Valencia 1991. Here we only state the major rules of the procedure:

Rules come in two kinds:

(a) *General rules of composition*:
occurrences in a function head A in applications $A(B)$ retain their polarity,
occurrences in a body A of a lambda abstract $\lambda x \bullet A$ retain their polarity,

where function heads can block the monotonicity marking in their arguments. Notice, e.g., how "Most *AB*" made its left-hand argument 'opaque': it is neither upward or downward. But "Most *AB*" does pass on monotonicity information in its right-hand argument *B*, and this demonstrates a second crucial source of information for our calculus:

(b) *Specific information about lexical items*:
e.g. "All" has functional type $e^{-} \rightarrow (e^{+} \rightarrow t)$.

Here is how the two kinds of information combine. First, in general, a function application $A(B)$ may block the polarity marking of positions in the argument B. E.g., "best (friend)" has no marking left for “friends”, as there is no inference to either “best girlfriend” or “best acquaintance”. The adjective “best” is highly context-dependent and hence steals the show.

But sometimes monotonicity marking does percolate upwards, when the meaning of the function head *A* ‘helps’. E.g., “blonde friend’ does imply “blonde acquaintance’, because the adjective “blonde’ has a simple ‘intersective’ meaning forming a Boolean conjunction “blonde$\cap B$”. More generally, if a function head *A* has type $a \rightarrow b$ where the argument type *a* is marked, the argument position *B* in applications *A(B)* will assume that same polarity. This explains how negations switch polarity, how conjunctions just pass them up, and so on. Such markings will normally come from lexical information, but there is one nice twist. They can also be introduced via lambda abstractions $\lambda x_a \bullet M_b$ of type $a \rightarrow b$, where the type *a* gets positive marking. Readers may want to check the semantics for an explanation.

The final ingredient to make this work is the following self-evident rule of calculation. Markings can be computed as long as there is an unbroken string in the parse tree:

(c) *The arithmetic of polarity combination:*

$+\ +\ =\ +\qquad -\ -\ =\ +\qquad +\ -\ =\ -\qquad -\ +\ =\ -$

This is just one mechanism for making natural logic precise. But its basic categorial structure has been rediscovered by many people: it just *is* rather natural! Thus we find one more key aspect of natural logic. Monotonicity marking works in tandem with one’s preferred syntactic analysis for natural language – providing fast inferences ‘on the fly’.

Digression: Boolean lambda calculus For technically inclined readers, here is a summary in terms of *Boolean typed lambda calculus*, the system behind much of the above (van Benthem 1991). Expressions now have ‘marked types’, and we define inductively what it means for an occurrence of a sub-expression to be positive or negative in an expression:

The occurrence x_a is positive in the term x_a,
The head *M* occurs positively in applications $M_{a\rightarrow b}\ (N_a)$,
If *M* has type $a^+ \rightarrow b$, then *N* occurs positively in $M_{a\rightarrow b}\ (N_a)$,
If *M* has type $a^- \rightarrow b$,, then *N* occurs negatively in $M_{a\rightarrow b}\ (N_a)$,
The body *M* occurs positively in $\lambda x_a \bullet M_b$, and the resulting type is $a^+ \rightarrow b$.

The rest is the earlier computation rule (c), or alternatively, we could have built this feature into the inductive definition. Clearly, this definition can be extended to deal with, not just functional types, but also *product types* $a \bullet b$ allowing for pair formation of objects.

Meta-theory: Lyndon theorems This logical perspective raises further issues of its own. One is the ‘completeness’ of the above syntactic marking procedure. Can we be sure that every semantically monotone inferential position will be found in this way?

> Is there a Lyndon Theorem stating that every semantically monotone occurrence must be positive in the above categorical sense?

This would extend the first-order result. The answer is 'Yes' in the categorial single-bind Lambek Calculus (this is a model-theoretic result proved by brute force in van Benthem 1991) – but the problem is still open for type theory with Boolean operators in general.

A lambda calculus for natural language has no internal first-/higher-order boundary. The only special thing to first-order quantifiers here is that they have more monotonicity markings than others, in line with our earlier observation about their inferential richness.

Depth rather than surface after all? We conclude this section with a worry. When all is said and done, our 'natural logic' turned out to be a rather modern system, requiring a fully-fledged parse of a sentence – just as logics require construction of formulas from the atomic level up. Is not this a 'kiss of death' for the natural logic program, that was supposed to work on surface syntax? Now, the above analysis can be twisted to work in a top-down manner (van Eijck 2005). The only thing we need to know for the monotonicity marking of a constituent is the hierarchical structure of the sentence *above it*, leaving all other 'side parts' unanalyzed. Even so, I must confess that I am not entirely happy. The categorical monotonicity calculus is definitely not carefree surface analysis, and like much of current linguistics and logic, it analyzes more than the bare minimum which seems involved in natural reasoning. How can we 'hit and run' as reasoners, or is that idea just a chimera?

But even with this worry, the main insight from the 1980s remains intriguing. Our natural inferences based on predicate restriction and monotonicity do not need special logical apparatus: they get a free ride on syntactic analysis, a task we have to perform anyway.

6 Descriptive challenges to natural logic

There are many further questions at this stage about the range of our 'natural logic' so far.

Polyadic quantifiers There is much more to quantifier patterns in natural language than the above single and iterated cases. Through the 1980s, further combinations have come to light which do not reduce to simple iterations, such as *cumulative forms* "Ten firms own 100 executive jets" or *branching patterns* "Most boys and most girls knew each other". These require new forms of monotonicity marking, depending on how one takes their meanings. Also, quantifiers also lead to *collective predication* ("The boys lifted the piano"), as well as *mass quantification* ("the teachers drank most of the wine"), whose inferential behaviour is far from being generally understood – either in linguistic semantics or in modern logic.

Other fast subsystems Maybe more interesting is the issue whether there are *other fast surface inference systems* in natural language. I already mentioned the general functioning of Conservativity as a mechanism of general *Role Restriction* for predicates in sentences.

And I can think of several other examples, such as 'individual positions' X in expressions allowing for arbitrary *distribution over disjunctions*, as in $\varphi(X_1 \vee X_2) \leftrightarrow \varphi(X_1) \vee \varphi(X_2)$. [10]

Interactions So much for separate inferential systems. Another issue is of course how these interact with other major features of natural language. E.g., *anaphora* with pronouns can wreak havoc with monotonicity inferences, as in the following example (van Benthem 1986; but the observation really goes back to Geach 1974 and earlier):

> "Everyone with a child owns a garden.
> Every owner of a garden waters it. *So:*
> Everyone who has a child sprinkles it?"

Here, the pronoun "it" has picked up the wrong antecedent. Again, information about the total sentence composition is crucial to block these inferences, and keep the correct ones.

Inference without scope resolution? Finally, here is a more daunting challenge for surface reasoning. If we are to do inference as close to the linguistic surface string as possible, it would be nice to not have to resolve all quantifier scope ambiguities – and infer as much as possible from ambiguous expressions. [11] This is the way language often comes to us. For instance, in the above two dragon examples with monotonicity markings, note that 5 out of the 7 positions in the string retain the same marking in either scope reading. Can we find a still more surfacy natural logic for inferences unaffected by ambiguity? A stream of recent work on inference with 'underspecified' syntax seems relevant here: cf. van Deemter 1996, van Eijck & Jaspars 1996, Fernando 1997, and subsequent work.

7 Digression: traditional logic and small inference languages

Here is another direction in current research on natural logic. We saw that the medieval logicians started classifying multi-quantifier inferences. Thus, they were well-aware that

> "Some P R all Q" implies "All Q are R-ed by some P"

---

[10] Another candidate came up at the Stanford RTE seminar in 2007: *disjointness of predicates*. Which expressions preserve this? Monotonicity was about inferences of the form: '$P \leq Q$ implies $\varphi(P) \leq \varphi(Q)$'. But now, we are now given, not an inclusion premise, but an *exclusion premise* (note that $P \cap Q = \emptyset$ iff $P \subseteq \neg Q$), and we want to know what follows then in the right contexts: '$P \leq \neg Q$ implies $\varphi(P) \leq \neg \varphi(Q)$'. In first-order logic, this amounts to stating a monotonicity-like inference between the formula and its *dual* obtained by working the prefix negation inward switching operators: '$P \leq Q$ implies $\varphi(P) \leq \varphi^{dual}(Q)$'. I think one can find first-order syntax which guarantees this. More generally, many classical model-theoretic preservation results in first-order logic can be re-interpreted as giving simple special-purpose syntax for specialized inferences.

[11] This is the view of natural language high-lighted by Lexical-Functional Grammar; cf. van Benthem, Fenstad, Halvorsen & Langholm 1987.

and that the converse fails. Now the iteration depth in ordinary discourse seems limited (at best 3 levels, as in “You can fool some people all of the time” seem to occur in practice). Thus it makes sense, pace Dummett, to define small special-purpose notations for such combinations, and try to axiomatize them. Moss 2005 is a first attempt in this direction of rehabilitating Scholastic efforts, with very appealing sets of valid principles. [12]

In addition to this focus on deductive completeness for small languages, Pratt 2006 has performed computational complexity analysis on several small decidable fragments of natural language. Here outcomes are a bit more gloomy, in that high complexity can arise – but as so often with bad news of this sort, it remains to be seen what this means in practice.

8 From language to computation

The story of natural logic has also crossed to computer science.

Efficient information processing First, it has often been observed that simple reasoning with relational databases only uses small parts of predicate logic, and that monotonicity accounts for most of what is needed. Accordingly, Purdy 1991 and follow-up publications have come close to the material covered in the above. Likewise, modern research on natural language processing, and in particular, intelligent text analysis appears to be arriving at similar aims and results. Results from the linguistic phase are promoted in van Eijck 2005, using various programming techniques for optimizing the monotonicity calculus and related forms of inference. For an extensive empirical investigation of actual data, see Manning & MacCartney 2007. Likewise, polarity calculi for hierarchical marking to extend the scope of natural lexical inferences with factive verbs have been proposed in Nairn, Condoravdi & Karttunen 2006. Whether all this is really less complex than first-order alternatives is a matter of detailed complexity-theoretic analysis, as said earlier – and the jury is still out.

Aside: fixed-point logics The computational setting also suggests natural constructions outside of first-order logic, such as transitive closure (Kleene iteration) and recursive definitions involving *fixed-point operators*. But monotonicity still makes sense here, and indeed, it is crucial. A recursive definition $Px \leftrightarrow \varphi(P)(x)$ of a new predicate P does not make sense in general – but it does when $\varphi(P)$ is semantically monotone with respect to P (Ebbinghaus & Flum 1996). Is this just a technical coincidence, or does this mean something from the perspective of natural logic? Maybe it supports circular definitions?

AI and default logic: ‘non-monotonicity calculus’? But there are other, less standard aspects to the computer science connection. In particular, ‘common sense reasoning’ has been analyzed in Artificial Intelligence by John McCarthy and his school. Now this involves not just monotonic inferences, but also *non-monotonic* ones, where inclusion of predicates

[12] It has to be admitted that higher-order quantifiers like “Most” have proved recalcitrant so far.

need not keep the conclusion true. This brings us to the area of default logics, which involve both classical reasoning with rock-solid conclusions, and *defeasible inferences* which can be retracted when new information comes in. For instance, by default "birds fly", but there are exceptions such as penguins, who tend to march… A systematic extension of the above monotonicity calculus to deal also with default implications based on predicate inclusions would be a highly interesting project! Existing logical systems in this area, which already combine material and default conditionals might provide a guide.

Combination, architecture and complexity My final computational theme returns to the natural logic program as such, and in particular, to what I myself consider a major open problem not recognized in the 1980s and before. Analyzing natural inference as a large family of simple fast subsystems is not enough! In reality, we are not a bare set of isolated processors. All this information must be *combined*, and one module must be able to feed quickly into another. So, what sort of 'natural' reasoning system are we?

Here, an insidious challenge to the whole enterprise emerges. Much experience with logical systems over the past decade has shown that the analytical strategy of divide and conquer' does not always hold. The complexity of the total logic is not just a maximum of the complexities of the components. It can explode dramatically, because the additional source of complexity is *the nature of the combination*, i.e., the communication between the different inferential modules. Indeed, several natural examples are known where apparently innocuous combinations of decidable logics create *undecidable* over-all inference systems. Note, this does not have to happen in every case, and the technique of 'fibered logics' (Gabbay 1996) provides a way out in many cases. But the danger is there.

Our conclusion here is this. Unless we have an additional idea about the *global architecture* of natural logic, claims about its performance may be premature and ill-founded.

9 Cognitive science

A final arena where natural logic is coming up these days is experimental cognitive science. We definitely know that inference in the human brain is not one unified phenomenon, but a joint venture between many modules, some related to our language abilities, some more to immediate visual processing or to schematic representation, and yet others to brain areas dedicated to planning and executive function (Knauff 2007). Current neuroscience experiments, guided by hypotheses about linguistic data (Geurts and van der Slik, 2005) are now beginning to map out how, for instance, monotonicity inference may be located in different parts of the brain than heavy-duty first-order logic (if that is available at all).

10 Conclusion: modern versus traditional logic once more

Natural logic is a proposal for taking a new look at natural inferential systems in human reasoning, not through the lenses of modern logic which see 'formal systems' everywhere. I think this is well-worth doing for all the reasons that I mentioned. But let me also be clear that I do not see this as a resumption of warfare between traditional logic and modern logic. If only by this 'cheap shot': the only respectable systems of natural logic that I know use thoroughly modern logical techniques and standards of exposition. There is no way back.

Architecture and transitions To me, by now, the more interesting question is one of *architecture* and *transitions*. Clearly, modern logic provides more subtle tools for analyzing reasoning than traditional logic, and it is of great interest to see where these must come into play. The analysis of mathematical reasoning is of course one clear case in point where traditional logic failed eventually in its accounts of density, continuity, limits, etcetera (cf. Friedman 1985 on Kant's views of mathematics). But the transitions do not lie where De Morgan and many modern teachers claim they lie. Traditional logic was much richer than many people think, and it still deals in attractive ways with large areas of natural language and common sense reasoning. First-order logic probably takes over when explicit variable management and complex object constellations become inevitable. Thus, it seems to me, there is no inevitable conflict between 'natural logic' versus modern logic.

Redefining the issue: mixtures and merges But again, one cannot just see this peaceful co–existence in sweeping terms like 'mathematics is modern logic', 'natural language is traditional logic'. The more interesting perspective for research may rather be

mixtures of natural and formal language and reasoning!

For instance, it is a telling fact that mathematicians have never abandoned natural language in favour of logical formalisms. Just read any mathematics paper, or go to any mathematics colloquium. The real situation is that mathematicians use mixtures of both, with the logical notation coming in *dynamically* when natural language needs to be made *more precise*. This mixture suggests that 'natural logic' and 'modern logic' can coexist harmoniously, because both have their place. And modern logic might even learn something from natural logic in combating its 'system imprisonment', trying to look for more general systems-free formulations of its basic insights, the same way, say, Monotonicity seems a general insight about human reasoning, which does not really seem to depend on any specific formal language and semantics. But I take this also as a shift in the whole issue. What we really need to understand is how out 'natural logic' can be 'naturally extended' with formal notations, and other technical inventions which somehow seem to fit our cognitive abilities.

The price of inferential holism Here is a final speculation. There might be another role then for modern first-order logic. Maybe modern logic is the only system which really *integrates* *all separate natural reasoning modules*. And then, as in my story of architecture and module combination, there may be a price to pay. This might be the true reason for the *undecidability* of first-order logic: not because its subsystems are so hard by themselves, but *because their combination is*. This may be seen by linking undecidability to Tiling Problems and interaction axioms (van Benthem 1996), but I will leave the matter here.

In summary, natural logic seems an inspiring, if not always well-defined, research program into human language and human reasoning which raises many new questions of its own, while helping us rethink the achievements of modern logic in new and unexpected ways.

11 References

J. Barwise & R. Cooper, 1981, 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* 4, 159 – 219.

J. van Benthem, 1986, *Essays in Logical Semantics*, Reidel, Dordrecht.

J. van Benthem 1987, 'Meaning: Interpretation and Inference', *Synthese* 73:3, 451-470.

J. van Benthem, 1991, *Language in Action. Categories, Lambdas and Dynamic Logic*, North-Holland Amsterdam & MIT Press, Cambridge (Mass.).

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford.

J. van Benthem, J-E Fenstad, K. Halvorsen & T. Langholm, 1987, *Situations, Language and Logic*, Reidel, Dordrecht, Studies in Linguistics and Philosophy, Vol. 34.

J. van Benthem & A. ter Meulen, eds., 1997, *Handbook of Logic and Language*, Elsevier, Amsterdam.

I. Bochenski, 1961, *A History of Formal Logic*, University of Notre Dame Press, Notre Dame.

H. Curry, 1936, 'A Mathematical Treatment of the Rules of the Syllogism', *Mind* 45, 209 – 216.

K. van Deemter, 1996, 'Towards a Logic of Ambiguous Expressions', in K. van Deemter & S. Peters, eds., *Semantic Ambiguity and Underspecification*, Cambridge University Press, Cambridge, 203 – 237.

M. Dummett, 1973, *Frege. The Philosophy of Language*, Duckworth, London.

H-D Ebbinghaus & J. Flum 1996, *Finite Model Theory*, Springer, Berlin.

G. Englebretsen, 1981, *Three Logicians: Aristotle, Leibniz, Sommers and the Syllogistic*, Van Gorcum, Assen.

J. van Eijck, 1985, *Aspects of Quantification in Natural Language*, Dissertation, Philosophical Institute, University of Groningen.

J. van Eijck, 2005, 'A Natural Logic for Natural Language', in B. ten Cate & H. Zeevat, eds., *The Batumi Proceedings*, FoLLI LNAI Series, Vol. 4363, Springer, Berlin, 216 – 230. Updated version, CWI Amsterdam & UIL-OTS Utrecht 2007.

J. van Eijck and J. Jaspars, 1996, 'Ambiguity and Reasoning', FRACAS Report, CWI Amsterdam, ISSN 0169-118X.

T. Fernando, 1997, 'Ambiguity under Changing Contexts', *Linguistics and Philosophy* 20, 575 – 606.

M. Friedman, 1985, 'Kant's Theory of Geometry', *The Philosophical Review* XCIV:4, 455 – 506.

D. Gabbay, 1996, *Labeled Deductive Systems*, Oxford University Press, Oxford.

P. Geach, 1972, *Logic Matters*, Blackwell, Oxford.

B. Geurts and F. van der Slik, 2005, 'Monotonicity and Processing Load', *Journal of Semantics* 22:1, 97 – 117

W. Hodges, 1998, 'The Laws of Distribution for Syllogisms', *Notre Dame Journal of Formal Logic* 39, 221 – 230.

E. Keenan & D. Westerståhl, 1997, 'Quantifiers', in J. van Benthem & A. ter Meulen, eds., *Handbook of Logic and Language*, Elsevier, Amsterdam.

M. Knauff, 2007, 'How our Brains Reason Logically', Center for Cognitive Science University of Freiburg. In H. Hodges, W. Hodges & J. van Benthem, eds., Special Issue of *Topoi* on "Logic and Psychology", May 2007.

F. Liu & J. Zhang, 2007, 'Some Thoughts on Mohist Logic', Chinese Academy of Social Science, Beijing, and ILLC, University of Amsterdam.

Ch. Manning & B. MacCartney, 2007, 'Natural Logic for Textual Inference', Departments of Linguistics/Computer Science, Stanford University. To appear at the Workshop on Textual Entailment and Paraphrasing, *ACL 2007.*

R. Montague, 1974, *Formal Philosophy*, Yale University Press, New Haven.

L. Moss, 2005, 'Completeness in Natural Logic: What and How', LSA Linguistics Summer Institute, Harvard University.

R. Nairn, C. Condoravdi & L. Karttunen, 2006, 'Computing Relative Polarity for Textual Inference', Palo Alto Research Center, in *Proceedings Inference in Computational Semantics*.

S. Peters & D. Westerståhl, 2006, *Quantifiers in Language and Logic*, Oxford University Press, Oxford.

I. Pratt, 2004, 'Fragments of Language', *Journal of Logic, Language and Information* 13:2, 207 – 223.

A. Prior, 1967, 'Traditional logic', in P. Edwards, ed., *The Encyclopedia of Philosophy*, Vol. 5, McMillan, London, 34 – 45.

W. Purdy, 1991, 'A Logic for Natural Language', *Notre Dame Journal of Formal Logic* 32:3, 409–425.

V. Sanchez Valencia, 1991, *Studies on Natural Logic and Categorial Grammar*, Dissertation, Institute for Logic, Language and Computation, University of Amsterdam.

V. Sanchez Valencia, 2004, 'The Algebra of Logic', in D. Gabbay & J. Woods, eds., *Handbook of the History of Logic*, Elsevier, Amsterdam 389 – 544.

F. Sommers, 1982, *The Logic of Natural Language*, Clarendon Press, Oxford.

P. Suppes, 1982, 'Variable-Free Semantics with Remarks on Procedural Extensions', in T. Simons & R. Scholes, eds., 1982, *Language, Mind and Brain*, Lawrence Erlbaum, Hillsdale (N.Y.), 21 – 31.

Appendix History once more: logic in China

When this material was presented at a lecture in Beijing, some interesting coincidences came to light. As is becoming known by now, logic started simultaneously in at least three geographical areas and cultures: Greece, India, and China. Here are a few telling examples from Liu & Zhang 2007 about Mohist logic (5th century B.C.), a school of thought clearly manned by logicians. The following inference is straight from the Mohist Canon:

> "A white horse is a horse. To ride a white horse is to ride a horse."

This is clearly the pattern of upward monotonicity.

But now, here are two further Mohist examples which seem to contradict this:

> "Robbers are people, but to abound in robbers is not to abound in people."
> "A cart is a wooden object. To ride a cart is not to ride a wooden object."

These examples are subtle, and both highlight a further phenomenon of logical interest. The first seems a failure of upward monotonicity due to *context dependence* of a quantifier. If "Many" just means 'more than a fixed threshold value *N*', it is upward monotonic in both arguments. But if we assume the norm is dependent on the predicate, as seems more likely, "Many" is not upward monotonic in either argument. [13] To manage correct and incorrect

[13] Another form of context dependence played in an example brought up in a colloquium

inferences here, one would need a dynamic mechanism of 'context management'. The second example seems one of *intensionality*. 'To ride a cart' may be read extensionally as just 'being transported', and then the conclusion does follow by upward monotonicity. But the Mohist colleagues surely meant something more subtle. Intensionally, one rides a cart *as a vehicle*, and read in that way, the stated inference comes out invalid, since one does not ride a wooden object qua wooden object. These refinements of the above monotonicity setting are as valid now as they were then. Intensional contexts have been a well-known challenge to simple monotonicity reasoning, ever since Richard Montague made intensional expressions a benchmark for the semantic analysis of natural language.

Mohist logic had many further subtle features than inferences like these. It also includes versions of the Paradox of the Cretans, and very nicely, the following pragmatic paradox of conversation, which was new at least to this author (though one referee of this paper saw an analogy with Buridan's Paradox that "Every proposition is negative"):

> Telling someone that "You can never learn anything" can never be successful.

The somewhat Whorfian question has sometimes been raised how one can ever recognize people in other cultures as colleagues. I would say, with the great historical work Bochenski 1961 that there is no such problem: '*only logicians* worry about crazy things like this'.

---

at PARC Palo Alto: Does "They verbally attacked the president" imply "They attacked the president"? The conclusion suggests (incorrectly) physical attacks.