Mood and modality in update semantics

Frank Veltman
University of Amsterdam
§ 1 Introduction

1.1 Topics and Themes

Topics

Monday:
Mainly on ‘might’

Tuesday and Wednesday:
Normally and Presumably

Thursday and Friday:
The counterfactual mood

Recurring Themes

o Semantics and pragmatics: is there any distinction left?

o How to distinguish between epistemic and alethic modalities?

o The role of expectations in interpretation and decision making
1.2 The standard view in 1980

**Semantics:**
- Logical validity defined in terms of truth
- Meaning equated with truth conditions
- Framework: possible worlds semantics

**Pragmatics:**
- Theory of conversation
- Pragmatic consequence (implicature)
- Framework: Grice’s theory of conversation

**Empirical criterion:**
Pragmatic consequences can be canceled, semantic consequences cannot.

In this picture information processing is part of pragmatics.

\[
S + \{\square\} = S\square
\]

*S*: input state (= set of possible worlds)

+: intersection

\( S\square \): output state
Example 1

Logical consequence: Most A are B. Therefore: Some A are B.
Pragmatic consequence: Most A are B. Therefore: Some A are not B.

Here the maxim of quantity is involved

The implicature is cancelable:

Most men are mortal. In fact, all are.

Example 2

Consider:

It’s raining but it might not be raining

Truth condition:

\textit{might} \textit{\box} \text{is true in} \textit{w} \text{iiff there is some epistemic alternative} \textit{w}' \text{to} \textit{w} \text{such that} \textit{\box} \text{is true in} \textit{w}'

This yields: \textit{p} \textit{\box} \textit{might} \textit{\box p} \text{is semantically consistent, but pragmatically absurd.}

Explanation involves Grice’s maxim of quality
Problem: this apparent inconsistency is not cancelable.
§2 The framework of update semantics

2.2 Basic definitions

*Slogan*

- To know the meaning of a sentence is to know the \textit{change} it brings about in the cognitive state of anyone who accepts the news conveyed by it.

Well, it’s a slogan, so…

- It’s much too narrow (think of questions, orders, etc.)
- It suggests we are interested in perlocutionary effect of an assertion.
Reminder: Speech acts

Recall the distinction between:

- **Locutionary act**
- **Illocutionary act**: the act I perform in performing the locutionary act
- **Perlocutionary act**: the act I may succeed in performing by performing the illocutionary act.

Example: in uttering the sentence “The door is open” (locutionary act), I state that the door is open (illocutionary act), and by performing this illocutionary act I may succeed in performing the perlocutionary act of getting you to shut it.

Note: even it may very well be the speaker’s intention to make the addressee shut the door, we would not want to say that “The door is open” has the same meaning as “Shut the door!”.

Examples of illocutionary acts: to make a statement, to make a promise, to warn, to order, to apologize.

Examples of perlocutionary acts: to convince, to persuade, to frighten
BOOH!

Webster’s New Collegiate Dictionary: ‘used to express contempt and disapproval, or to startle or frighten’

Consensus: semantics should restrict itself to the study of illocutionary acts. That’s where the conventional meaning of a sentence is to be found.
Anyway:
The meaning of a sentence is a function from information states to information states.

Let $S$ be any information state and $f$ a sentence with meaning $[f]$. I will write $S[f]$ for the information state that results when $S$ is updated with $f$.

This set up enables us to deal with sequences of sentences, texts:
$S[f_1] [f_2]$ is the result of updating $S$ with $f_1, f_2$ in that order.

*The dynamic circle:*
  
  Interpretation depends on context
  Interpretation builds context

The dynamic approach enables us to study all kinds of phenomena that have to do with text coherence.
Examples

• I lost ten marbles and found all but one. It’s under the sofa.
• I lost ten marbles and found nine of them. Presumably it is under the sofa.

• Two men came in. One was wearing blue suede shoes, the other wasn’t.

• John didn’t drink too much wine. He would have got sick.
• If John had drunk too much wine, he would have got sick.
• If John had not drunk too much wine, he would not have got sick.
• John drank too much wine. He would not have got sick (??).
Definition

An update system for a language L is given by

- A set \( S \) of relevant information states
- A function \( [\ ] \) that assigns a state \( S[\ ] \) to each sentence \( \square \) of L and each state \( S \) in \( S \). In general this function will be partial.

Key notion

Sometimes the information conveyed by \( \square \) will already be subsumed by \( S \). In this case we write

\[
S \models \square.
\]

and we say that \( \square \) is accepted in \( S \), or that \( S \) supports \( \square \).

In simple cases the notion of acceptance can be defined as follows:

\[
S \models \square \iff S[S[\square]] = S
\]
Special case:

An update system \([\square, [\ ]\)] is additive iff there exists a (minimal) state \(\mathbf{0}\) in \(\square\), and a binary operation + on \(\square\) such that for any \(S, S', S''\)

(i) \(\mathbf{0} + S = S\)

\[S + S = S\]

\[S + S' = S + S\]

\[S + (S + S') = (S + S) + S''\]

(ii) for every sentence \(\square\) and state \(S, S [\square] = S + \mathbf{0} [\square]\)
2.2 Constraints that do not always hold

Generally speaking, the update operation will be partial rather than total. Take the case of a pronoun desperately looking for a referent:

*He is just joking.*

If it is not clear to whom the speaker is referring, the addressee will not know what to do with this statement.

Or take the case of presupposition. Indeed, the framework of update semantics offers a natural explanation of this notion:

\[ \square \text{presupposes} \square \text{iff for every } S, S [\square] \text{ is defined only if } S \models \square \]

Clearly, this definition can only be instrumental in systems in which \( S [\square] \) is sometimes undefined.
The principle of *Idempotence*:

For any state $s$ and sentence $\Box$, $S[\Box] \models \Box$

At first sight this principle goes without saying. But there are sentences for which it doesn't hold. Here paradoxical sentences like

*This sentence is false*

are a point in case: there is no state $S$ such that

$S[This \ sentence \ is \ false] = S$

There is no such thing as a successful update with the Liar, and that is where its paradoxicality resides.

For more details, see:

A system will not be additive unless the principle of permutation holds.

\[ S[f][y] = S[y][f] \]

Compare:

- Somebody is knocking at the door…. It might be John…. It’s Mary
- Somebody is knocking at the door…It’s Mary…It might be John

Counterexamples to the principle of permutation are easy to construct with sentences in which modal qualifications like ‘may’, ‘must’, ‘presumably’, ‘probably’, etc. occur.
2.3 Logical validity

There are various ways to define logical validity:

(i) Updating any state $S$ with the premises $\Box_1, \Box, \Box_n$ in that order yields a state that supports the conclusion $\Box$

$$\Box_1, \Box_n \models_1 \Box \iff \text{for any } S, S [\Box_1] \Box_n \models \Box$$

(ii) Updating the minimal state $0$ with the premises $\Box_1, \Box_n$ in that order yields a state that supports the conclusion $\Box$

$$\Box_1, \Box_n \models_2 \Box \iff 0 [\Box_1] \Box_n \models \Box$$

(iii) One cannot accept the premises without having to accept the conclusion as well:

$$\Box_1, \Box_n \models_3 \Box \iff S \models \Box \text{ for any } S \text{ such that } S \models \Box_1, \Box_n, S \models \Box_n$$

**Fact**

- Validity$_1$ implies validity$_2$
- Validity$_1$ implies validity$_3$
- In every additive system these three notions coincide.

Proof: left as exercise 1
Let us now concentrate on the second validity notion:
\[ 0_1,\ldots,0_n \vdash 0 \iff 0_1[0_1]\ldots[0_n] \vdash 0. \]

This notion of validity will in general be nonmonotonic, but it follows immediately from the definition that it satisfies the next two principles:

**Consequential Monotonicity:**

If \( 0_1,\ldots,0_n \vdash 0 \) and \( 0_1,\ldots,0_n, 0_1,\ldots,0_k \vdash 0 \)

then \( 0_1,\ldots,0_n, 0_1,\ldots,0_k \vdash 0 \)

**Consequential Cut:**

If \( 0_1,\ldots,0_n \vdash 0 \) and \( 0_1,\ldots,0_n, 0_1,\ldots,0_k \vdash 0 \)

then \( 0_1,\ldots,0_n, 0_1,\ldots,0_k \vdash 0 \)

Given the Principle of Idempotence, we also find

**Reflexivity:**

\[ 0_1,\ldots,0_n, 0 \vdash 0 \]

**Proposition**

Let \( \vdash \) be any consequence relation for which the principles mentioned above hold. Then there is an update system in which the principle of Idempotence holds such that

\[ 0_1,\ldots,0_n \vdash 0 \iff 0_1,\ldots,0_n \vdash 0 \]

The proof is left as Exercise 2
§3 Might

3.1 Updates and tests.

We add one new operator *might* to the language of propositional logic.

*Basic idea*

One has to agree to *might* • if • is consistent with ones knowledge — or rather with what one takes to be ones knowledge. Otherwise is to be rejected.
Definition

Let $W$ be the set of possible worlds*.

(i) $S$ is an information state iff $S \subseteq W$

(ii) The minimal state, $\mathbf{0}$, is the state given by $W$

The absurd state, $\mathbf{1}$, is the state given by $\emptyset$.

A state $S$ represents the agent's knowledge of the facts. It contains those worlds which — for all the agent in state $S$ knows — may yet turn out the real one. As the agent's knowledge increases, $S$ shrinks.

*) We think of a possible world as being determined by the atomic sentences that are true in it. Thus, we identify $W$ with the powerset of the set of the atomic sentences.
**Definition**

Updating $s$ with $\Box$

- atoms: $S[p] = S \{ w \mid W \mid p \Box w \}$
- $\neg$: $S[\neg \Box] = S \setminus S[\Box]$
- $\Box$: $S[\Box \Box \Box] = S[\Box] [\Box]$
- *might*: $S[might \Box] = S$ if $S[\Box] \neq \mathbf{1}$
  
  $S[might \Box] = \mathbf{1}$ if $S[\Box] = \mathbf{1}$

If $S[\Box] \neq \mathbf{1}$, we say that $\Box$ is acceptable in $S$.

If $S[\Box] = \mathbf{1}$, $\Box$ is not acceptable in $S$;

- these notions are meant to be normative.
- these notions have little to do with truth and falsity.
- We are not dealing with revision here (!)
3.2 Moore-like paradoxes

Sentences of the form *might □* provide an invitation to perform a test on $S$ rather than to incorporate some new information in it.

- $0 [\text{might } \neg p][p] \neq 1$;
- $0 [p][\text{might } \neg p] = 1$.

- A text $\Box_i, \Box, \Box_n$ is *consistent* iff there is some $S$ such that $S [\Box_i] [\Box, ]$ exists and $S [\Box_i] [\Box_i] \neq 1$
- A text $\Box_i, \Box, \Box_n$ is *coherent* iff there is some $S$ such that $S [\Box_i] [\Box, ] = S$ and $S \neq 1$

Consistency is hearer oriented
Coherence is speaker oriented

- $p \Box \text{might } \neg p$ is not consistent
- $\text{might } \neg p \Box p$ is consistent, but not coherent

**Claim:**

Semantics and pragmatics can be so attuned that the dividing line between logical validity and pragmatic correctness is drawn exactly as the criterion of cancellability prescribes.

More on Moore’s paradox:

3.3 Logical Matters

- $\text{might } \neg p \models_{1,2} \text{might } \neg p$, but $\text{might } \neg p, p \not\models_{1,2} \text{might } \neg p$.

  Hence, validity$_1$ and validity$_2$ are not monotonic.

**Facts:**

- $\models_2 \text{might } p$, but $\not\models_{1,3} \text{might } p$
- $\text{might } p, \neg p \models_3 q$, but $\text{might } p, \neg p \not\models_{1,2} q$
- $\text{might } p, \neg p \models_{2,3} q$, but $\text{might } p, \neg p \not\models_1 q$

If you are interested in proof systems for these various validity notions, look at

Proposition

Let \( \Box \) be a sentence in which \textit{might} does not occur. Set
\[
\Box = \text{df} \ 0 \ [\Box]
\]
Then for any information state \( S \),
\[
S [\Box] = S \Box \ [\Box].
\]

If \( \Box \) is a sentence of ordinary propositional logic, we can speak of ‘the \( \Box \)-worlds’, the worlds in which the \textit{proposition expressed by} \( \Box \) holds. It would be nonsense to speak of the ‘\textit{might} \( \Box \)-worlds’. Sentences of the form \textit{might} \( \Box \) do not express a proposition.

Proposition

Let \( \Box_1, \ldots, \Box_n, \Box \) be sentences in which \textit{might} does not occur.
\[
\Box_1, \ldots, \Box_n \models_{1,2,3} \Box \text{ iff } \Box_1, \ldots, \Box_n / \Box \text{ is valid in classical logic.}
\]
3.4 Shortcomings

Hans Rott:

“As far as I know, dynamic semantics up to now has taken into account only revisions by formulae that are consistent with the prior beliefs of the speaker. In the inconsistent case revisions get more complicated since they are determined not only by the meaning of the input, but also by internal factors like preferences associated with the speaker's mental state. These factors, however, may well be taken part and parcel of the belief states themselves, and are responsible for the fortunate fact that belief-contravening information is not indigestible information. Therefore I see no reason to exclude, from the (determination of the) meaning of [], those changes that [] brings about in belief states with which it is logically incompatible.”

No, no, no....no?

- If, by updating with $\square$, the state $S[\square]$ gets incoherent, several things can happen. Revision is just one option.

- The principle of compositionality.
- Revision only makes sense for sentences expressing a proposition.

- Compare:
  - *If Shakespeare didn't write Hamlet, who did?*
  - *If Shakespeare had not written Hamlet, who would have?*
More shortcomings

Consider a text of the form

\[ \text{might} [], \text{would} [] \]

Updating with \textit{might} [] does not change the state. Given this, it is a riddle how such a sentence can create a proper context for a sentence of the form \textit{would} [].

(Solution: take a closer look at the \textit{process} of interpretation. In interpreting \textit{might} [] in \( S \) one has to consider \( S [\square] \) and examine the outcome. Idea: keep the auxiliary state \( S [\square] \) in memory, as a subordinated state, so that \textit{would} [] or any other sentence in the subjunctive mood can be interpreted in the right context).

\textit{might} [] gives second-order information. It tells the hearer something about the information of the speaker.

See for example:


The eliminative set up does not do justice to the fact that we sometimes \textit{discover} that a certain possibility is still open.
4.1 The case of Bill

Bill is 34 years old. He is intelligent, but unimaginative, compulsive and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Please, rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

- a) Bill is a physician who plays poker for a hobby
- b) Bill is an architect
- c) Bill is an accountant
- d) Bill plays jazz for a hobby
- e) Bill surfs for a hobby
- f) Bill is a reporter
- g) Bill is an accountant who plays jazz for a hobby
- h) Bill climbs mountains for a hobby
Various groups of respondents, differing in statistical sophistication, were confronted with this test. Almost 90% ranked the compound target (g) below the simple target (c), and above simple target (d).

Tversky’s & Kahneman’s explanation:

• The participants rank the eight target categories by the degree to which Bill is representative of that category.
• More generally: people tend to base their probability judgments on the notion of representativeness.
• The relation of representativeness has a logic of its own, which departs systematically from the logic of probability.

Challenge:
What exactly is this logic of representativeness?

“Human reasoning cannot be adequately described in terms of context independent formal rules”

(op. cit., 499)

Basic idea:

Bill is more representative of the class A&J of accountants who play jazz for a hobby than of the class J of people who play Jazz for a hobby because given the information at hand

a) some of of his properties, though exceptional for someone playing Jazz for a hobby, are normal for an accountant who does so.

b) none of his properties is normal for people who play Jazz for a hobby but exceptional for accountants who do so.

To make this more precise we have to get to grips with the logic of sentences of the form

\[ P's \text{ normally have the property } Q \]

Compare:

\[ \begin{align*} 
\text{premise 1:} & \quad P's \text{ normally are } Q \\
\text{premise 2:} & \quad P's \text{ normally are } R \\
\text{conclusion:} & \quad P's \text{ normally are both } Q \text{ and } R \\
\end{align*} \]

\[ \begin{align*} 
\text{premise 1:} & \quad \text{Most } P's \text{ are } Q \\
\text{premise 2:} & \quad \text{Most } P's \text{ are } R \\
\text{conclusion:} & \quad \text{Most } P's \text{ are both } Q \text{ and } R. \\
\end{align*} \]
Sentences of the form

\[ P's \, normally \, have \, the \, property \, Q \]

express default rules.

Roughly what they mean is this. Whenever you are confronted with an object with the property \( P \), you may assume it has the property \( Q \) as well—provided you have no evidence to the contrary.

In other words, the next inference is valid (in some sense of the word).

\[ \text{premise 1: Accountants normally are unimaginative} \]
\[ \text{premise 2: Bill is an accountant} \]
\[ \text{conclusion: Presumably, Bill is unimaginative.} \]

Whereas this one is not:

\[ \text{premise 2: Bill is an accountant} \]
\[ \text{premise 3: Bill is not unimaginative} \]
\[ \text{conclusion: Presumably, Bill is unimaginative.} \]

Claim
The notion of ‘validity’ at stake here is the notion introduced as validity2

Compare
- Bill is an architect
- Presumably Bill is an architect
- I am hungry / Presumably, you are hungry
- You are hungry / Presumably, I am hungry
4.2 Rules with exceptions
4.2.1 Expectation patterns

Language

We add two new operators to the languages of propositional logic: 

normally and presumably.

Note: normally and presumably can only occur as the outermost operator of a sentence.

The speakers of this language can learn not only that something is in fact the case, but also that something is normally the case. On top of that they are able to decide whether — in view of the information at hand — something is presumably the case.

States

We want the information states to capture two things: (i) the agent's knowledge and (ii) the agent's expectations.

An information state $S$ will be a pair $[s\square]$ 

Here $s$ represents the agent's knowledge of the facts, just like before. It is a subset of the set $W$ of possible worlds.

The agent's expectations are represented by a relation $\Box$ on the set $W$ of possible worlds. This relation encodes the rules the agent is acquainted with. (When a new rule is learnt, $\Box$ changes).
**Definition**

Let $W$ be the set of possibilities. □ is an *expectation pattern* on $W$ iff □ is a pre-ordering of $W$; i.e. □ is reflexive and transitive.

- Instead of ‘$\llbracket i, j \rrbracket \sqsubset^\square$’, we will write ‘$i \llbracket \llbracket j$’
- If both $\llbracket i, j \rrbracket \llbracket \llbracket$ and $\llbracket j, i \rrbracket \llbracket \llbracket$ we write ‘$i \equiv \llbracket j$’.
- Clearly, $\equiv \llbracket$ is an equivalence relation
- It will be clear what $i \llbracket \llbracket j$’ means.

![Diagram](image)

**Picture of an information state □ $s$□

- There are eight possible worlds, $w_0, \ldots, w_7$.
- The worlds constituting $s$ are placed in the area with dashed borders.
- If two worlds belong to the same $\equiv \llbracket$equivalence class, they are placed within the same circle or oval.
- If $w_i \llbracket \llbracket w_j$, the diagram contains a *rightward* path from the $\equiv \llbracket$equivalence class to which $w_i$ belongs to the $\equiv \llbracket$equivalence class to which $w_j$ belongs.
- There are two *normal* worlds, $w_0$, and $w_5$.
- There is one *optimal* world, $w_6$. 
Let $\square$ be a pattern on $W$;

(i) $i$ is a normal world in $\square$ iff $i \square j$ for every $j \square W$;

(ii) $n\square$ is the set of all normal worlds in $\square$;

(iii) $\square$ is coherent iff $n\square \neq \emptyset$.

A pattern $\square$ is coherent if there is at least one possible world in which every proposition considered normally to be the case is the case.

Let $\square$ be a pattern on $W$, and $s \square W$.

A world $i$ is optimal in $\square$ s iff $i \square s$ and there is no $j \square s$ such that $j < \square i$.

An agent in state $\square s$ expects that the real world to be as close to standards of normality as possible. But it would not be very realistic to expect things to be more normal than the ‘facts’ leave room for.
Example

\[
\begin{array}{c}
\begin{array}{cccc}
0 & 1 & 2 & 3 \\
\end{array}
\end{array}
\]

\[
\begin{array}{cccc}
\text{normally } p
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{cccc}
3 & 1 & 2 & 0 \\
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\text{q}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{cccc}
3 & 1 & 2 & 0 \\
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\text{\neg p}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{cccc}
3 & 1 & 2 & 0 \\
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\text{\neg (p \square q)}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{cccc}
1 & 2 & 0 \\
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{cccc}
3 & 1 & 0 \\
\end{array}
\end{array}
\]

\[
W = \{w_p, w_r, w_2, w_3\}, \text{ where } w_3 = \{p, q\}, w_2 = \{q\}, w_1 = \{p\} \text{ and } w_0 = \emptyset.
\]

0 [normally p] |= presumably p
0 [normally p] [q] |= presumably p
0 [normally p] [q] [\Box p] \not\models presumably p
0 [normally p] [normally q] |= presumably p
0 [normally p] [normally q] [\Box p] \not\models presumably p
0 [normally p] [normally q] [\Box p] |= presumably q
0 [normally p] [normally q] [\Box (p \quad q)] \not\models presumably p
0 [normally p] [normally q] [\Box (p \quad q)] \not\models presumably q
4.2.2 The system

Let $W$ be as before.

$S$ is an information state iff $S = \bigcup \mathfrak{s}$ and one of the following conditions is fulfilled:

a) $\mathfrak{i}$ is a coherent pattern on $W$ and $\emptyset \neq s \notin W$;

b) (b) $\mathfrak{i} = \{\mathfrak{v}, w \in W\}$ and $s = \emptyset$.

c) $0$, the minimal state, is the state given by $W \times W$, $W \mathfrak{i}$

d) $1$, the absurd state, is the state given by $\{\mathfrak{v}, w \in W\}$, $\emptyset \mathfrak{i}$

Let $S = \bigcup \mathfrak{s}$ be a state and $\mathfrak{f}$ a sentence. $S [\mathfrak{f}]$ is determined as follows:

- If $\mathfrak{f}$ is a sentence of propositional logic, then
  - if $s \notin \mathfrak{f} = \emptyset$, $S [\mathfrak{f}] = 1$;
  - otherwise, $S [\mathfrak{f}] = \bigcup \mathfrak{s} \notin \mathfrak{f}$.

  (Here $\mathfrak{f}$ is the proposition expressed by $\mathfrak{f}$).

- If $\mathfrak{f} = \text{normally} \mathfrak{f}$, then
  - if $n \mathfrak{f} = \emptyset$, $S [\mathfrak{f}] = 1$;
  - otherwise, $S [\mathfrak{f}] = \bigcup \mathfrak{s} \notin \mathfrak{f}$, $s \notin$ where
    $\mathfrak{i}, j \notin \mathfrak{f} \c \mathfrak{f}$ iff $\mathfrak{i}, j \notin \mathfrak{f}$ and either $j \notin \mathfrak{f}$ or $i \notin \mathfrak{f}$.

- If $\mathfrak{f} = \text{presumably} \mathfrak{f}$, then
  - if $m_S \mathfrak{f} = \mathfrak{f}$, $S [\mathfrak{f}] = S$;
  - otherwise, $S [\mathfrak{f}] = 1$.

  (Here $m_S$ is the set of optimal worlds in $S$)
The refinement operation is put to work when a new rule is learnt. Think of this operation as follows: Suppose \( i \approx j \). Then \( i \) conforms to all the default rules that \( j \) conforms to — at least in so far as the rules encoded in \( e \) are concerned. Now, a new default rule comes in: \( \text{normally } f \). Let \( \boxed{f} \) be the set of worlds in which \( f \) is true. If the new rule is accepted, the new pattern becomes \( e \cdot \boxed{f} \). That is: if it happens to be the case that \( i \in \boxed{f} \) and \( j \in \boxed{f} \), one has to remove the pair \( i, j \) from \( e \). Given the new rule, it is no longer the case that \( i \) conforms to all the rules that \( j \) conforms to.

Updating with \( \text{presumably } f \) amounts to taking a test. All an agent in state \( S \) has to do when told that \( f \) is presumably the case check whether \( f \) is accepted in the optimal part of \( S \). If so, the agent accepts the sentence \( \text{presumably } f \). Otherwise, \( \text{presumably } f \) is unacceptable.
4.2.3 *normally* as a modal operator

We find:

\[
\text{normally} \Box, \text{normally} \Diamond \models_{1,2,3} \text{normally} (\Box \Diamond \Box)
\]

\[
\text{normally} \Box, \text{normally} \Diamond \models_{1,2,3} \text{normally} (\Box \Diamond)
\]

If \models \Box, then \models_{1,2,3} \text{normally} \Box

We already saw that generally speaking,

\[
\text{normally} \Box \not\models \Box
\]

What we have instead is the much weaker principle

\[
\text{normally} \Box \models_{2} \text{presumably} \Box.
\]

More surprising perhaps, it is not generally so that

\[
\text{normally} \Box \models_{1,2,3} \text{normally} (\Box \Diamond)
\]

*Compare.*

- Normally it rains. It is not raining now. So, presumably it is snowing.
- Normally it rains or it snows. It is not raining now. So, presumably it is snowing.
4.3 Rules for exceptions

The system devised above lacks expressive power. We cannot say when exceptional circumstances are to be expected and what one can expect when they obtain.

Example:

*Normally it rains; but when there is an easterly wind, the weather is normally dry.*

We add two new operators to the language of propositional logic: [] and presumably.

‘[] []’ is short for ‘if [], normally []’

\[\text{normally} [] = \text{df} (\text{[]} \text{ []} \text{ []} \text{ []})\]

It must be possible for a proposition to be a rule in a given domain without it automatically being a rule in all its subdomains. This means we need to be able to assign different patterns to different subsets of \(W\).

Definition

Let \(W\) be as before. A *frame on \(W\)* is a function [] that assigns to every subset \(d\) of \(W\) a pattern [] on \(d\).

The definition of a frame allows for the possibility that in every two different subsets \(W\) of different default rules hold. But, of course, not anything goes. If we make too many exceptions, our expectation frame gets incoherent.
Examples

• Normally it rains, but if there is an easterly wind, the weather is usually dry. And also if there is no easterly wind, the weather is usually dry.
• Normally it rains, and normally the temperature is between 10°C or higher. But when it rains, the temperature is normally below 10°C.

Here is the state of an agent who accepts the general rule *normally* $p$, but who wants to make an exception for the case that $q$ holds.

[$p$] is a default in $\square W$:

```
\begin{array}{c}
1 & 0 & 2 \\
\end{array}
```

[$\square p$] is a default in $\square[q]$:

```
\begin{array}{c}
2 & 3 \\
\end{array}
```

The basic idea here:

Exceptions to exceptive clauses do not count as normal.

Now, suppose the agent also wants to make an exception for the case that $q$ does not hold. Intuitively, this is too much. The resulting frame $\square$ is the same as the frame above except that $\square\{w_0, w_1\}$ looks like this:

```
\begin{array}{c}
0 & 1 \\
\end{array}
```

This means that in the domain $W$ there are no normal worlds left!
Definition

Let $\Box$ be a frame on $W$, $d \subseteq W$, and $i \subseteq d$.

a) $i$ is a normal $d$-world in $\Box$ iff for every $d' \subseteq d$ such that $i \subseteq d'$ it holds that $i \Box j'$ for every $j \subseteq d'$;
b) $\Box d$ is the set of all normal $d$-worlds in $\Box$; c) $\Box$ is coherent iff for every nonempty $d \subseteq W$, $\Box d \neq \emptyset$.

Let $S = \Box$, $s \subseteq$ be an information state. The frame $\Box$ encodes the rules an agent in state $S$ is acquainted with, and $s$ his knowledge of the facts. Now, what will an agent in state $S$ expect?

The crucial notion here is the notion of applicability: If you want to know what an agent in state $\Box$, $s \subseteq$ expects, you will have to sort out which of the rules encoded in $\Box$ apply to which parts of $s$.

Let $d'$ be a subdomain of $d$, and suppose that $e$ is a default rule in $d$, i.e. $\Box d \bullet [\Box] = \Box d$. The $d$-rule $e$ may very well apply to $d'$. When will this be? Roughly, the answer is this: If the rule $e$ can be coherently used as a rule in every domain between $d'$ and $d$.

Examples

For each of the states $S_i = \Box$, $s_i \subseteq$ we want to know which rules apply in $s_i$.

(i) $S_1 = 0 [normally p] [q \Box p] [q]$
(ii) $S_2 = 0 [normally p] [q \Box p] [q \Box r]$
(iii) $S_3 = 0 [normally p] [q \Box p] [(q \Box r) \Box p] [q \Box r]$
Definition

Let $S = \Box, s\square$ be a coherent information state and assume that $e_1,\ldots, e_n$ are defaults in $\Box d_1,\ldots, \Box d_n$ respectively.

(i) A world $w$ complies with $\{e_1,\ldots, e_n\}$ iff $w \square e_i$ for every $i$ such that $w \square d_i$ (1 $\leq$ $i$ $\leq$ $n$).

(ii) The set of defaults $\{e_1,\ldots, e_n\}$ applies within $s$ iff for every $d \supseteq s$ there is some $w \square n \Box d$ such that $w$ complies with $\{e_1,\ldots, e_n\}$.

(Instead of saying ‘the set $\{e_1,\ldots, e_n\}$ applies within $s$’, we often say ‘$e_1,\ldots, e_n$ jointly apply within $s$’).

If a set of defaults applies within a given context $s$, the effect will be that worlds not complying with these defaults do not count as normal $s$-worlds any more.

So, when does a set of defaults apply within $s$? If for no domain $d$ extending $s$, the set $n \Box d$ of normal $d$-worlds consists entirely of worlds not complying with the defaults in question. Because otherwise, if the defaults did apply, the frame would get incoherent.
Example

Consider $\Box, s=0 [ p \Box r ] [ q \Box \neg r ] [ p \Box q ]$.

<table>
<thead>
<tr>
<th>index</th>
<th>world</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>1</td>
<td>$p$</td>
</tr>
<tr>
<td>2</td>
<td>$q$</td>
</tr>
<tr>
<td>3</td>
<td>$q, p$</td>
</tr>
<tr>
<td>4</td>
<td>$r$</td>
</tr>
<tr>
<td>5</td>
<td>$r, p$</td>
</tr>
<tr>
<td>6</td>
<td>$r, q$</td>
</tr>
<tr>
<td>7</td>
<td>$r, q, p$</td>
</tr>
</tbody>
</table>

We are dealing here with a set $W=\{w_0, \ldots, w_7\}$ of eight possible worlds described in the table on the left.

The set $s=\{w_3, w_7\}$.

$\Box$ is the following frame:

If $d \neq \{w_1, w_3, w_5, w_7\}$ and $d \neq \{w_2, w_3, w_6, w_7\}$, $\Box d = d \times d$.

$\Box[p]$ looks like this:

$\Box[q]$ is this:


Definition
Let $S = [\square, s]$ be a coherent information state and assume that $\{e_1, \ldots, e_n\}$ are defaults in $[\square d_1, \ldots, \square d_n]$.

a) $\{e_1, \ldots, e_n\}$ is a \textit{maximal applicable set} in $S$ iff $e_1, \ldots, e_n$ jointly apply within $s$, and for every $e_{n+1}$ and $d_{n+1}$ such that $e_{n+1}$ is a default in $[\square d_{n+1}]$, and $e_1, \ldots, e_n, e_{n+1}$ jointly apply within $s$ it holds that $e_{n+1} = e_i$ and $d_{n+1} = d_i$ and for some $i \leq n$.

b) A world $w$ is \textit{optimal in $S$} iff $w \square s$ and $w$ complies with a maximal applicable set of defaults. The set of optimal worlds is denoted by $m_S$.

c) $S[\text{presumably } \square]$ is determined as follows:
   - If $m_S \square [\square] = m_S$, then $S[\text{presumably } \square] = S$
   - Otherwise, $S[\text{presumably } \square] = 1$.

4.3 Predicate logical interpretation

Think of $p, q, r$ etc. as atomic predicates rather than atomic sentences. Each well-formed expression of propositional logic now specifies a boolean combination of properties.

Think of $W$ as the set of \textit{possible (types of) objects}.

Like before, the set $s$ figuring in an information state $[\square, s]$ represents an agent's knowledge, only this time it is not the agent's knowledge about the real world, but about \textit{some} real object.

A default in a pattern $[\square d]$ now is a property — a property that the objects in $d$ \textit{normally} have. The formula $[\square [\square] \square]$ can be read as ‘objects with the property expressed by $[\square]$ normally have the property expressed by $\square$’.
Example

Adults normally are employed
Students normally are not employed
Students normally are adults
John is a student

Presumably, John is adult and not employed

This argument is valid. To see why, we have to determine the state

\[ 0 \left[ p \square r \right] \left[ q \square \neg r \right] \left[ q \square \neg p \right] \left[ q \right] = S = \Box, s \]

Let \( W \) be defined as before. Then \( s = \{ w_2, w_3, w_6, w_7 \} \). For the frame \( \square \), we find:

\( \square \{ w_1, w_3, w_5, w_7 \} \) can be depicted as:

```
  7   3
  5   1
```

And this is \( \square \{ w_2, w_3, w_6, w_7 \} \),

```
  7
  3
  2
  6
```

The \[ p \square \] default \[ r \] does not apply within \( s \), because \( n \square s = \{ w_3 \} \square [p] \sim [r] \)

But both the other rules do apply. So, \( m_s = \{ w_3 \} \), which means that

\( S \models presumably (p \square \neg r) \).
Exercise

Prove that

a) $p \Box r, q \Box (p \Box \neg r), p \Box q \models_2 \text{presumably } \neg r$

b) $p \Box q, q \Box r, p \models_2 \text{presumably } r$

c) $p \Box q, \neg q \models_2 \text{presumably } \neg p$

d) $p \Box q, q \Box \neg p, p \models_2 \text{presumably } q$

e) $p \Box q, q \Box \neg p, p \not\models_2 \text{presumably } \neg p$

f) $p \Box q, q \Box \neg p, p \not\models_2 \text{presumably } \neg q$

4.4 Notable differences with other theories:

- Default principles are treated as sentences of the object language. (Cf. Reiter’s format)

- The distinction between defeasible and indefeasible conclusions is made manifest at the level of the object language. (Note that the conclusions start with ‘presumably’)

- Questions of priority, which are likely to arise in the case of conflicting defaults, are decided at the level of semantics. (I don’t have to stipulate things like: ‘More specific rules override more general rules’
4.5 Back to Bill

In state $S$ property $Q$ is considered normal for objects with property $P$ iff ‘$P$'s are normally $Q'$ is accepted in $S$.

In state $S$ property property $Q$ is exceptional for objects with property $P$ iff ‘$P$'s are normally not $Q'$ is accepted in $S$.

In state $S$ the set of properties $X$ is more representative for objects with property $A$ then for objects with property $B$ iff

a. Some properties in $X$ are normal for objects with property $A$ and exceptional fore object with property $B$.

b. No properties in $X$ are normal for objects with property $B$ and exceptional fore objects with property $A$.

Is the conjunction fallacy really a fallacy?

Yes, is the answer of Tversky and Kahneman.

“The reliance on representativeness leads to systematic errors”.(op.cit. p.89)

Side remark

In an argument-endorsement procedure it turned out that most statistically trained respondents (83%) were ready to admit that they made a “stupid” mistake. Naive subjects were much less impressed by normative arguments and many (57%), even those who who endorsed the the conjunction rule in te abstract, remained committed to to their initial responses.
No, is the answer of many logicians working in non-monotonic logic.

“Which arguments should a logical theory admit as valid/invalid? In the case of nonmonotonic logic the criteria are not the same as in the case of classical logic.... In the case of default reasoning the whole notion of what is correct (the notion we are trying to model) is defined in terms of what sort of reasoning people actually engage in and what sorts of reasoning an intelligent agent will have to do to get along in a common sense way. So “mistakes” by common sense reasoners are only possible in the sense that they are different from the majority....Psychologism with respect to default reasoning is correct!”


However:

• The theory I proposed, or any other logic of defaults for that matter, is just as normative as probability theory or classical logic is.

• Both probability theory and the theory I proposed offer a ‘logic’ for taking decisions in circumstances where the facts of the matter are only partly known. Perhaps the theory I proposed is more descriptive in the sense that it better describes the ‘logic’ that people actually use, but from this it does not follow it is a better logic.

• The question which logic is best is open.
4.5 Is there any distinction left?

• *Most A are B.*

Is the following sentence a semantic or pragmatic consequence?

• *Presumably, some A are not B.*

• *John hit Bill. He was severely injured.*

Are the following sentences semantic or pragmatic consequences?

• *John or Bill was severely injured.*

• *Presumably, Bill was severely injured.*

• *John hit the Terminator. He was severely injured.*

Same question about:

• *Presumably, John was severely injured.*

Questions

Cancellability = Defeasibility ?

In the interpretation process, where does linguistic competence end, and world-knowledge begin?