ITK Research Report No. 15
19 March, 1990

DIT - Dynamic interpretation
in text and dialogue

Harry C. Bunt


ISSN 0924-7807

©1990. Institute for Language Technology and Artificial Intelligence, Tilburg University, The Netherlands
1 Introduction

This paper is concerned with the development of a theory of meaning, inspired by the analysis of dialogues. More specifically, of dialogues with the sole purpose of exchanging factual information, so-called information dialogues. Such dialogues occur at information desks in railroad stations, in travel agencies, on airports, with the telephone information service, etc. They are of particular interest both from a theoretical and from a practical point of view.

The theoretical interest lies in the fact that information dialogues can be explained in terms of the information available to the participants and their intentions to expand the available information, the way information and intention are conveyed by utterances, and the logical properties of intention and information. Dialogues with more complex purposes than the exchange of information require other concepts for their explanation, but also require the concepts needed for pure information dialogues, since every form of genuine communication relies on conveying intentions and information, be it perhaps of a more complex nature. Information dialogues are therefore the conceptually simplest form of full-blown communication, whose understanding is fundamental to understanding communication in general.

The practical interest of information dialogues is that they form the kind of dialogue that would be of interest with a computer information system. A theoretical model of information dialogues with the explicitness, formality
and efficiency that are required for computer implementation would therefore be of practical value for the creation of useful and interesting dialogue systems.

1.1 Information dialogue and change of information state

One of the most conspicuous features of information dialogues is that in such dialogues participants say something in order to change the prevailing state of information. One may therefore think that, roughly speaking, changes in information of participant A occur because he interprets an utterance $u$, directed towards him by participant B. This is not the case, however, at least not on the usual understanding of 'interpreting' an utterance. I take it that, usually, by interpreting an utterance we mean that a meaning is assigned to the utterance. The following example makes clear that assigning meanings alone does not change a state of factual information.

Suppose speaker A asks partner B what time it is. B says: "It's 3 a.m." A sees the sun shining and, not being close to the polar circle, does not believe B. His state of factual information therefore does not change. A's state of factual information would only change if, in addition to interpreting B's utterance, A would also believe B. Evaluating the factual information in B's utterance, weighing it against other evidence, and deciding whether to accept it, are additional steps without which A's state of factual information cannot change.

Still, though the linguistic interpretation alone does not change the addressee's factual information, it does change something for the addressee, for we know that, if the addressee's evaluation (weighing the factual information in the speaker's utterance against other evidence, and deciding whether to accept it) comes out positively, then he will accept the factual information and thereby change his state accordingly. What the linguistic interpretation changes is the addressee's information about the speaker: from the fact that the speaker is answering that it's 3 o'clock (rather than, for instance, checking it) it follows, given the context of an information dialogue, that the speaker believes that it's 3 o'clock. This information about the speaker's beliefs is new for the addressee. (If A already knew that B knew it's 3 o'clock, his utterance would not be a genuine informative question.)

1This may seem to be a self-evident truth, inherent to the notion of an information dialogue; yet we will have to qualify this (see section 2.2).
1.2 Interpretation and meaning

Linguistic interpretation and meaning are closely related, since interpretation involves the assignment of meaning. One well-known view is that linguistic interpretation takes linguistic expressions as inputs and delivers meaning representations as outputs for further processing. Notably, it has become popular to view these outputs logical formulae, expressing truth conditions in terms of model-theoretic entities.

When we consider linguistic interpretation as a process that changes the addressee's state of information about the speaker, a different picture emerges. The recognition of the meaning of an utterance no longer takes the form of expressing truth conditions, which can be evaluated against a model of the discourse domain, but takes the form of expressing how to update the state of information. Such a process may be construed by building an intermediate formal representation of the update operations to be performed, using again logical formulae but this time not with a semantics in terms of truth conditions but one in terms of update conditions (Veltman, 1989). The meaning of an utterance, as determined by such an interpretation process (with or without explicit representation at an intermediate stage), is the way in which an information state has to be updated. This is a more 'dynamic' view on meaning than the traditional one in terms of truth conditions relative to static models.

The dynamic view on meaning comes very naturally for utterances in an information dialogue, which are obviously meant to change something. But it is also attractive from other points of view. Groenendijk and Stokhof (1987, 1989) have developed a dynamic semantics for the purpose of resolving anaphoric expressions, as an alternative to such approaches as Discourse Representation Theory. They formulate the idea of dynamic interpretation as follows:

"The meaning of a sentence does not lie in its truth conditions, but rather in the way it changes the (representation of the) information of the interpreter. The utterance of a sentence brings us from a certain state of information to another. The meaning of a sentence lies in the way it brings about such a transition." (Groenendijk and Stokhof, 1987, p.1)

Upon closer consideration, it turns out that Groenendijk and Stokhof's dynamic semantics is in fact concerned with different phenomena than the dynamic interpretation theory that we are pursuing here. For one thing, Groe-
nendijk and Stokhof are concerned with *sentences* where we are concerned with *utterances*; secondly, Groenendijk and Stokhof associate sentence semantics with changing factual information, where we associate utterances with *speaker* information; and thirdly, Groenendijk and Stokhof limit the dynamics in their models to the values of variables, whereas I think it is essential to allow the interpretation of *constants* to be dynamic. In section 8 below we will come back to the differences between dynamic semantics and dynamic interpretation theory.

If we follow the lead that the meaning of a linguistic expression is the way it may change the (representation of the) information of an interpreter, our first task is to investigate in what ways linguistic expressions do in fact change interpreters' information. This will be the subject of the next section.

2 Expressions and changes of information

2.1 Expressions: sentences and utterances

As mentioned above, a fundamental difference between Groenendijk and Stokhof's Dynamic Semantics (DS) on the one hand, and our Dynamic Interpretation Theory (DIT) on the other, is that DS is concerned with sentence meaning and DIT with utterance meaning. The preoccupation with utterance meaning is natural in the context of dialogues, since in dialogues one finds one and the same expression used in different ways, with different effects on an interpreter's information. The following example illustrates this.²

```latex
(1) 1 A: Schiphol information
2 B: Good afternoon. This is Van I. in Eindhoven. I would like
to have some information about flights to Munich.
When can I fly there between now and ... next Sunday
3 A: Let me have a look. Just a moment
4 B: Yes
5 A: O.K., there are ... three flights every day, one at nine
      fifty,
6 B: Yes,
7 A: one at one-forty ... and one at six twenty-five
8 B: Six twenty-five ... These all go to Munich
9 A: These all go to Munich
```

² Schiphol is the name of Amsterdam Airport
10 B: And that's on Saturday too
11 A: And that's on Saturday too, yes
12 B: Right ... Do you also have information about the connections to Schiphol by train?
13 A: Yes, I do.
14 B: Do you know how long the train ride takes to Schiphol?
15 A: You are travelling from Eindhoven?
16 B: That's right.
17 A: It's nearly two hours to Amsterdam ... You change there and then it's another fifteen minutes, so you should count on some two and a half hours
18 B: O.K., thank you
19 A: You're welcome
20 B: Bye
21 A: Bye

Here we see several nice examples where the same sentence is used with different purposes and has correspondingly different effects. The sentence *These all go to Munich* as used in line 8 by B, has on A (among other things) the effect that A knows that B wants to know whether the flights in question all go to Munich, whereas the same sentence used subsequently by A has the effect on B that B now knows that indeed these all go to Munich. These differences in effects on the addressee are due to the fact that the first use of the sentence has the function of a question (more precisely, of a check; it is an example of a so-called *declarative question act*; see Beun, 1989), whereas the second use has the function of an answer (more precisely, of a *confirmation*).

We see that specific information-changing effects can be associated only with a sentence *used with a certain communicative function*, i.e., with an *utterance*. This point is also illustrated by the different occurrences of *"On Saturday too"*, as well as by the different occurrences of *"Yes"*, which have very different functions.

Below, (section 8) I will argue that it is wrong in principle to consider sentence meaning in terms of information-changing effects, as Dynamic Semantics does; only utterance meaning can sensibly be approached that way. This is independent of our concern with dialogues (cf. section 2.3).
2.2 Information-changing effects of utterances

The ways in which utterances in information dialogues do in fact change interpreters' information has been addressed at some length elsewhere (e.g., Bunt, 1986; 1989a); here I will only summarize a few important aspects.

Inherent to the very notion of an information dialogue is that we have two partners A and B with the sole purpose of obtaining and providing factual information. The communicative actions that A and B perform will obviously have effects on their information. So, basic to the representation of the effects of utterances in information dialogues is the information that A and B have available and in what respects they want to expand their information. Instead of saying that A has certain information available, we sometimes say that A 'knows' something. But we must be careful using the term 'knowledge', for two reasons. First, we should perhaps speak of 'belief' rather than 'knowledge', in order to avoid the implication that the available information is necessarily correct. The course of an information dialogue is not determined by what is actually true, but by what the participants believe to be true. What is meant here by saying that A knows that x, is no more than that A has the information x available, without implying any commitment to the truth of x. Until further notice we will use the terms 'know that' and 'believe that' interchangeably, as shorthands for 'to have the information available that'.

Natural information dialogues contain a substantial amount of verification, which indicates that participants in such dialogues often have uncertain knowledge about something, and a desire to get rid of these uncertainties. I will describe the situation where a participant A has some information p available without fully trusting it as A suspects that p.

Not only the information available to the partners is crucial in an informative dialogue, but also the information which is not available and in particular the information that they want to become available. There are two ways in which one may want information to become available: one may want it to become available to oneself or one may want it to become available to the partner. In other words, one may want to know something or one may want to make something known. The things that one may want to know or to make known are, in the first place, specific facts about the discourse domain; that is what the dialogue is for. However, sometimes a participant wants to know or to make known something about oneself, about the other, or about the way the dialogue is developing; in that case we have so-called dialogue control acts (Bunt, 1986).
It can be argued that successful communication is achieved only when an utterance gives rise to mutual knowledge concerning the intentions and beliefs connected to the utterance (see Clark & Marshall, 1981).

On the basis of an analysis of the flow of information in actual information dialogues, in Bunt (1986) a taxonomy of communicative action types has been developed using three major categories: questioning, informing, and answering acts. In each of these categories a variety of communicative functions is distinguished, which are characterized in terms of packages of appropriateness conditions. The functions within one category share an appropriateness condition that expresses the intention motivating the act. For questioning acts this is the condition that the speaker wants to know something (namely the value of the semantic content of the utterance), for informing acts it is that the speaker wants to make something known to the addressee, and for answering it is the speaker's knowledge that the addressee wants to know something.

The details of this are of no special concern here; important is that each utterance is supposed to realize at least one communicative act, which conveys a certain package of beliefs and intentions on the part of the speaker. In describing these packages, it is useful to make the well-known distinction between semantic content and communicative function, where the communicative function characterizes the way in which the belief- and intention attitudes of the dialogue partners are involved. The belief- and intention attitudes most relevant for information dialogues are the following:

(2) \begin{align}
& \text{to know that} \ldots \\
& \text{to suspect that} \ldots \\
& \text{to want to know} \ldots \\
& \text{to want to make known that} \ldots
\end{align}

In the Introduction we have already seen that the effects of an utterance on an addressee are the result of the addressee not only interpreting the utterance, but also evaluating the factual information it contains against other available information. The effects of an utterance on an interpreter thus depend on the available information. This dependence can be taken into account by construing the effects, brought about by interpreting an utterance, through the application of a state-changing function with the set of all possible states of information as its domain.

It also depends, of course, on whether the addressee perceives and interprets the utterance correctly and reasons correctly; if the addressee misunderstands or misinterprets what was said, the effects on him do not cor-
respond to the intended meaning of the utterance. A theory of meaning should not be burdened with such complications. We therefore assume the addressee to be an ideal rational understander, who does not misunderstand or misinterpret, and who applies the rules of logic perfectly.

Technically, a notion of utterance meaning which applies to any possible information state can take the form of the set of all pairs \(<\text{initial state, final state}>\) such that an addressee in an initial state \(S_i\) may come in the state \(S_f\) by interpreting \(u\). This leads, for instance, to saying that "It's 3 o'clock" has as meaning something like (2):

\[
(2) \quad \{<S_i, S_j> | S_j = S_i + \text{'addressee knows that speaker intends hearer to know that it's 3 o'clock'}\}
\]

Alternatively, the meaning of \(u\) can be construed as a function from initial states to final states. This approach is preferred in DIT; it leads, for the same example, to meanings like:

\[
(3) \quad \lambda s : s + \text{'addressee knows that speaker intends hearer to know that it's 3 o'clock'}
\]

I think the use of functions as meanings is to be preferred to the use of state pair sets, since it makes state transitions deterministic. This is not to say that natural language utterances are viewed as unambiguously taking an interpreter from a given initial state of information to a unique final state, but that an utterance with an unambiguously determined communicative function and factual information should achieve this. We will return to this in section 8.

One of the most conspicuous features of natural language utterances is that they are often vague or ambiguous semantically (in their factual information) or pragmatically (in their communicative function) or both. Can this be accounted for in the dynamic approach to meaning? Ambiguity can be reconstructed in terms of multiple possible effects on an interpreter. Can we account for vagueness in effects on interpreters? Or should all vagueness be resolved before we can supply meanings as state changes? I think it would be extremely unattractive if this were required. Semantic ambiguity and vagueness can be dealt with in that part of the theory that takes care of determining the factual information in an utterance. If one allows ambiguity or vagueness in this information, as in the DIRT approach (Kálmán, 1989) or in Two-level Semantics (Bunt, 1985), then this can be inherited by DIT as ambiguity or vagueness in the beliefs about the speaker that constitute
the effects of utterances. Pragmatic ambiguity and vagueness are different. A case of pragmatic ambiguity occurs when a declarative sentence is used, without it being clear to the interpreter whether it is meant as an assertion or as a question. Since this makes an essential difference for which aspects of the addressee’s state are changed, there seems to be no other possibility than assigning two different meanings to such an utterance and require the interpreter to make a choice. A case of pragmatic vagueness occurs when an utterance is interrogative, but it is not clear whether the speaker is making an assumption about the answer or not, i.e., it isn’t clear whether he is ‘just’ asking a question, or verifying something. Vagueness of this kind can be handled by using communicative action types which differ in specificity (see Bunt, 1986; Beun, 1989) - for instance, a check can be defined as a specific type of question.

Another important kind of pragmatic vagueness has its origin in possibly implied indirect interpretations. Many utterances have certain standard indirect interpretations, such as “Can you tell me whether X” with intended interpretation “Please tell me whether X”; however, there is often some uncertainty as to whether indirect interpretations are actually intended. An indirect interpretation can be viewed as more specific than the direct one, since satisfaction of the intention of the indirect interpretation standardly implies satisfaction of that of the direct interpretation. Preserving vagueness of this kind, which means that no choice is made between any of the more specific interpretations, comes down to interpreting the utterance directly, with the possibility to add more specificity through indirect interpretation as supportive evidence for this becomes available.

2.3 Dynamic interpretation of texts

We have argued above that the interpretation of utterances in dialogues necessarily takes the functions of utterances into account. It may seem that this is a particularity of dialogues. This is not so, however; the interpretation of sentences in written texts also has to take the function of the text elements into account. A striking example is provided by the text printed on road signs at railroad crossings in France:

(4) Un train peut en cacher un autre.

Literally, this text merely asserts that one train may hide another. A rather trivial assertion at that, as in general an object may hide another one of
the same kind. Of course, the text is intended as a warning not to cross the railroad immediately after a train has passed, because there may be a train from the other direction that you didn’t see because it was behind the first train. (Interestingly, the corresponding signs in Holland say all that explicitly, in a text headed by ‘WARNING!’, which makes the communicative function of the text explicit.) Clearly, the effect that (4) has on its interpreters involves the recognition of the communicative function of the sign (the ‘written utterance’, so to speak). Interpreted as an assertion, it would change the interpreter’s information in a totally different way. We will return to this point in section 8.

3 Requirements on models

From the above observations we can derive certain requirements on the kind of models that are needed to represent an interpreter’s changing information. The analysis we present here is specific for information dialogues, but when ‘belief’, ‘want’ and ‘knowledge’ are replaced by or supplemented with other propositional attitudes, such as ‘hope’, ‘fear’, or ‘dislike’, relevant in other types of communication, we can expect to apply a similar analysis and find solutions with similar methods.

The first and foremost requirement on interpreter models is that they should be capable of representing epistemic intentions and recursive information about such intentions as well as about both certain and uncertain information (‘knowledge/belief’ and ‘suspicion’).

Second, this information is typically incomplete at all levels: one usually has only limited information about what the dialogue partner intends, knows and does not know. So the model should represent incomplete information adequately.

Third, closely related to the previous point, the model should be computationally attractive for operating on incomplete information. Typically, the information one dialogue partner has about the other’s information state is very limited at the beginning of the dialogue; as the dialogue goes on, more and more information becomes available. To handle this in a computationally attractive way, these models should not have an eliminatory character, as standard possible-worlds models do, but should be incremental in character.

Fourth, the models should be suited to updating. The interpretation of a given utterance typically affects only certain ‘dimensions’ of an interpreter
model, namely those corresponding to the attitudes with which the factual information in the utterance is coupled, according to the utterance's communicative function. Therefore, they should preferably be organized in a modular fashion with respect to the relevant propositional attitudes.

Fifth, the model should take the chronology of the interpreter's changes to some extent into account. An interpreter in a dialogue may have to suspend certain intentions, to discard certain information which he later may want to reconsider, etc.

In the following sections we will outline a formalism for defining models which satisfy at least the first four requirements; the chronological aspect will not be considered here.

4 Incremental partial models

4.1 Possible worlds and partial models

The standard approach for dealing with an agent's knowledge (or belief) is the possible-worlds formalism, where knowledge is expressed in terms of the alternative worlds that the agent considers possible. For an agent $S$, those worlds are distinguished among the set of all logically possible worlds by taking part in the relation 'accessible for $S$'. When $S$ knows that $p$, this is represented by $p$ being true in all worlds accessible for $S$; when $S$ knows that not $p$, this is represented by $p$ being false in all $S$-accessible worlds. When $S$ does not know whether $p$, there is at least one $S$-accessible world where $p$ is true and one where it is false.

This approach is too inefficient for computer implementation, since $S$ not knowing the truth of a fact $q$ which has not been considered before, is modelled by adding to each $S$-accessible world one where $q$ is true and one where $q$ is false. Therefore, the less an agent knows, the more worlds have to be represented. Moreover, the facts whose truth an agent does know have to be represented as such in every one of these worlds. All in all, this has the effect that for a realistic domain of discourse, with a large number of potential facts, the representation of an agent's incomplete knowledge involves an astronomic number of sets of facts; moreover, each of these sets is large, since possible worlds are complete: every atomic proposition must have a truth value in every world.

Ideally, one would like to model an agent's knowledge by representing only the facts he knows, and to represent these only once. This leads us to \textit{partial models}, where truth values are assigned to only those propositions
whose truth is known. Such models get complicated, however, because the modelling of disjunctive, negative and conditional knowledge, as well as of knowledge about absence of knowledge, is not straightforward.

We have developed an approach which constructs modular partial models using structured clusters of (partial) valuation functions, somewhat similar to Fagin, Halpern and Vardi's 'knowledge structures' (Fagin, Halpern & Vardi, 1984; Fagin & Vardi, 1985). For a finite domain of discourse, each valuation function can be implemented as a miniature data base containing its extension. A structured cluster of valuation functions can be used to describe the information relating to one particular combination of agents and attitudes (such as S's knowledge about U's knowledge about S's knowledge) and implemented as a small cluster of miniature data bases; we refer to those clusters (or to the function clusters they implement) as data modules. The entire model can be viewed as a network of data modules; we therefore call our approach the Data Module Net (DMN) approach.

We summarize here the minimal nontrivial use of DMN models, where two agents S and U are considered, one propositional attitude 'to have the information available that' (for brevity also referred to as 'to know' or as 'to believe', and where the embedded logic is just propositional logic. The corresponding logical language is called 2a-DpL: Two-agent Doxastic propositional Language. The reader is referred to Bunt (1989b) for more about DMN-models.

4.2 Partial valuations and incomplete information

A consequence of considering the beliefs of two different agents is that the valuation of propositional terms must be agent-dependent. Moreover, this dependency cannot take the form of an 'agent' coordinate with just two values, since we also have to deal with the beliefs of one agent about those of the other (and so on). So if $V(s, p)$ denotes the truth value of $p$ according to $S$, and $V(u, p)$ that according to $U$, we also need something like $V_{su}(p)$ for the truth value that $S$ believes that $U$ believes $p$ to have. And so on. We will return to the agent-dependency of the valuation below, but first consider some properties of the valuation $V(\alpha, \ldots)$ for a fixed value $\alpha$ of the agent-coordinate. We will use the notation $V_\alpha$ to indicate this subfunction of the valuation.

The ideal of only representing the facts that an agent knows, has the consequence that, if agent $\alpha$ does not know whether the atomic proposition $p$ is true or false, $V_\alpha$ should be undefined for $p$, i.e. $V_\alpha$ should be a partial
function. However, just making the valuation functions partial is insufficient for dealing with incomplete information in general. If agent $\alpha$ knows that $p$ or $q$, but not which of the two, this cannot be represented by the partiality of $V_\alpha$ alone. What we do in this case is introduce two 'alternative extensions' of $V_\alpha$, one that makes $p$ true and one that makes $q$ true. Calling these extensions $V_{\alpha_1}$ and $V_{\alpha_2}$, we have, formally:

$$
(5) \begin{align*}
V_{\alpha_1}(x) &= V_\alpha(x) \text{ for } x \neq p, \text{ and } V_{\alpha_1}(p) = 1 \\
V_{\alpha_2}(x) &= V_\alpha(x) \text{ for } x \neq q, \text{ and } V_{\alpha_2}(q) = 1 \\
V_\alpha &\subseteq V_{\alpha_1}; V_\alpha \subseteq V_{\alpha_2}
\end{align*}
$$

This implements the view that, if $S$ knows that $p$ or $q$ but knows neither that $p$ nor that $q$, upon extending his knowledge $S$ will obtain the knowledge that $p$ or the knowledge that $q$ (or both).

Agent $\alpha$'s knowledge that $p$ or $q$ can now be modelled by $V_\alpha$ having these alternative extensions $V_{\alpha_1}$ and $V_{\alpha_2}$, and the truth condition (6):

$$
(6) \text{ $\alpha$ knows that } p \lor q \iff V_\alpha(p) = 1 \text{ or } V_\alpha(q) = 1 \text{ or in all alternative extensions } V_{\alpha_i} \text{ of } V_\alpha:\nV_{\alpha_i}(p) = 1 \text{ or } V_{\alpha_i}(q) = 1
$$

(This will be made more precise below.) Figure 1 gives a pictorial representation of a model which represents, according to (6), that $S$ knows (only) that $p$ or $q$.

---

<table>
<thead>
<tr>
<th>p</th>
</tr>
</thead>
</table>
---

---

<table>
<thead>
<tr>
<th>q</th>
</tr>
</thead>
</table>

---

Figure 1. $S$ knows that $p$ or $q$

We see here the beginning of the use of 'function clusters' (or 'data modules'): it is the valuation $V_\alpha$ with its alternative extensions $V_{\alpha_i}$ that represents $\alpha$'s information about the domain of discourse.
Note that, by the stipulations \( V_\alpha \sqsubseteq V_{\alpha_1} \); \( V_\alpha \sqsubseteq V_{\alpha_2} \), the valuations \( V_{\alpha_1} \) and \( V_{\alpha_2} \) indeed constitute extensions of the knowledge represented by \( V_\alpha \). In other words, the knowledge of \( V_\alpha \) is persistent in alternative extensions. We do not want to copy the extension of \( V_\alpha \) in its alternative extensions, though; that would go against our aim to represent the facts an agent knows only once. We therefore define alternative extensions slightly differently by (7) as valuating only the additional facts, and we will define the truth conditions of 2a-DpL expressions in such a way that they ensure the persistence of information in alternative extensions.

(7) \( V_{\alpha_1}(x) \) is undefined if \( V_\alpha(x) \) is defined, and \( V_{\alpha_2}(p) = 1 \)
\( V_{\alpha_1}(x) \) is undefined if \( V_\alpha(x) \) is defined, and \( V_{\alpha_2}(q) = 1 \)

'Negative information' gives rise to similar complications. Suppose agent \( \alpha \) has very little information about the nationalities of the persons in a certain domain of discourse, not knowing anybody who's Italian, for instance, but knowing that John is not Italian. So \( \alpha \)'s information about the predicate constant ITALIAN is restricted to that the individual john does not belong to \( V_\alpha(ITALIAN) \). This 'negative information' can be modelled by introducing, in addition to \( V_\alpha \), a 'negative part' \( V_{\alpha neg} \) that expresses which entities do not belong to the extensions of the terms.

In the case of 2-agent doxastic propositional logic the need to represent negative information does not present itself at the level of terms, since the knowledge that a propositional constant \( p \) is not true is equivalent to the knowledge that it is false. However, we need to be able to express negative complex information. This brings us back to the agent-dependency of the valuation.

Using the subfunction \( V_{su} \) for \( S \)'s beliefs about \( U \)'s beliefs, \( V_{su}(p) = 1 \) represents that \( S \) believes that \( U \) believes that \( p \), and \( V(su,p) = 0 \) that \( S \) believes that \( U \) believes that not \( p \); we also need a way to represent that \( S \) believes that \( U \) does not believe that \( p \). To this end we introduce the additional valuation \( V_{sueg} \), which expresses \( S \)'s negative information about \( U \)'s information.

Alternative extensions, as introduced above, can take care of disjunctive factual knowledge. The agent-dependency of the valuation must also take alternative extensions into account, because an agent may have disjunctive information involving the other agent. For instance, if all that \( S \) knows is that if \( U \) knows that \( p \) then \( q \), we need the alternative extensions \( V_{s_1} \) and
\( V_{\alpha_2} \) such that \( V_{\alpha_1}(q) = 1 \) and \( V_{\alpha_2}(p) = 1 \). Figure 2 gives a pictorial representation of this model.

Figure 2. \( S \) knows that if \( U \) knows that \( p \) then \( q \)

## 5 DMN models

### 5.1 Formal definition of DMNs

A DMN model can be viewed as a collection of modules (function clusters) linked together to form a network through inter-speaker connections. If the model is to represent the information state of agent \( \alpha \), it contains a module with index \( \alpha \) which functions as an 'entry' to the network (like \( S \) in Fig.2). The presence of alternative extensions depends on which disjunctive knowledge is to be represented; the specification of the extensions which are present will therefore be part of the model. This leads to the following definition of a DMN model (or 'DMN': Data Module Net) as a tuple consisting of 3 elements: (1) a set of agent-dependent partial valuations; (2) the particular 'entry point' valuation describing the domain knowledge of the agent whose state of information is modelled; (3) the specification of alternative extensions.
DEFINITION. A DMN model for 2a-DpL is a triple

\[(8) \quad M = < F_a, \mathcal{F}, A >, \text{ where:} \]
\[- F_a \in \mathcal{F}; \]
\[- \mathcal{F} \text{ is an indexed set of partial functions from the} \]
\[- \text{propositional letters of 2a-DpL to truth values;} \]
\[- A \text{ is a partial function from } \mathcal{F} \text{ into } \mathcal{P}(\mathcal{F}) \text{ (specifying the} \]
\[- \text{alternative extensions present in } M) \]

In what follows we will mostly use ‘S’ as a name of the agent whose information states are modelled, and U as that of the other dialogue participant. The indices occurring in the indexed set of functions \(\mathcal{F}\) are defined as follows.

DEFINITION. For a model \(M = < F_a, \mathcal{F}, A >\) the set \(I_M\) of indices is the smallest set such that:

\[(9) \]
\[1. \ s \in I_M; \]
\[2. \text{if } i \in I_M \text{ then } is \text{ and } iu \in I_M; \]
\[3. \text{if } i \in I_M \text{ and } i \text{ is not of the form } j \text{ neg}, \text{ then } i\neg \in I_M; \]
\[4. \text{if } i \in I_M \text{ and } F_k \in A(F_i) \text{ then } i_k \in I_M. \]

Stipulating that a certain index \(i\) belongs to \(I_M\) is in fact a way of stipulating that the function \(F_i\) belongs to \(\mathcal{F}\). Since the definition of the set of indices of a model is the same for every 2a-DpL model, we keep it for the sake of readability outside the model definition (8).

Note that the specification of the inter-speaker connections and negative parts does not form part of the model. This is because we assume that every valuation has both a positive and a negative part and has connections to both agents; however, many of these connections will connect up to ‘empty’ parts, which contain no information. Such empty parts are defined formally below.

At this point we can also make precise what we mean by a ‘function cluster’ (or ‘data module’).

DEFINITION. A function cluster or data module in a DMN-model \(M_i\) for 2a-DpL is the information contained in a valuation \(F_i\), for some index \(i \in I_M\), plus that in the negative part \(F_{in\neg}\) and that in the alternative extensions \(F_{i\neg k}\), for \(i_k \in I_M\).

The index \(i\) of the valuation \(F_i\) which forms the heart of a function cluster
will be called the index of the function cluster.

A **DMN submodel** is a (sub-)DMN with an ‘entry point’ corresponding to a particular complex agent/attitude combination \( \alpha \), such as \( S \) believes \( U \) believes that \( S \) does not know that (where the ‘entry’ index would be \( su \) neg \( s \)). We will write \( M_i \) to denote the sub-DMN with entry index \( i \).

Formally, a sub-DMN of a DMN-model \( M = \langle F_\alpha, \mathcal{F}, \mathcal{A} \rangle \) is a triple \( \langle F_i, \mathcal{F}', \mathcal{A}' \rangle \) where \( F_i \in \mathcal{F}' \), \( \mathcal{F}' \subseteq \mathcal{F} \), and \( \mathcal{A}' \) is \( \mathcal{A} \) restricted to \( \mathcal{F}' \).

A sub-DMN is called **empty** if it does not contain any information. Formally:

**Definition.** An empty sub-DMN is a sub-DMN \( M_i = \langle F_i, \mathcal{F}', \mathcal{A}' \rangle \) such that \( F_j = \emptyset \) for every \( F_j \in \mathcal{F}' \).

We now turn to the semantics of 2a-DpL-expressions of the form \( S \models X \) ('\( S \) has the information that \( X \)'), i.e., we define under what conditions a DMN, meant to model the information state of \( S \), represents that \( S \) knows that \( X \).

### 5.2 Modularity and locality in DMN models

One of the requirements on models that we formulated above is that of **modularity.** In order to determine whether \( S \) knows that \( U \) knows that \( p \), in the possible-worlds approach one has to consult all the worlds which are \( U \)-accessible from a world that is \( S \)-accessible. \( S \)'s knowledge about \( U \)'s knowledge is, so to say, spread out over the entire model, intertwined with information relating to other propositional attitudes (such as what \( U \) believes that \( S \) believes that \( U \) knows). The computational complexity of evaluation and update operations may greatly benefit from a more modular organization, where one consults only the relevant ‘modules’ of the model. DMN-models have been designed to meet this requirement: to determine, for example, whether (10) is true in a DMN-model one consults the function cluster with index \( su \) (the module containing what \( S \) knows that \( U \) knows).

\[ (10) \quad S \text{ knows that } U \text{ knows that } p \]

In very simple cases, evaluation and update involve only one function
cluster. In general, however, the consultation of several 'modules' is required. This is in fact even the case for (10), as the following model illustrates. The model of Fig. 8 represents the information (11):

\[(11) \quad S \text{ knows that } U \text{ knows that } (p \text{ or } q) \]
\[S \text{ knows that } U \text{ knows that } \neg p\]

Since (10) can be deduced from (11), we want the evaluation of $S$ knows that $U$ knows that $p$ to come out true.

![Figure 8](image)

**Figure 8.** $S$ knows that $U$ knows that $(p \text{ or } q)$;
$S$ knows that $U$ knows that $\neg p$

This model says that, according to $S$, $U$ entertains two possibilities, corresponding to the submodels $M_{su_1}$ and $M_{su_2}$. However, in both cases $U$ holds the belief that $\neg p$, in view of the persistency of information in alternative extensions. So the two possibilities that $U$ entertains correspond in fact to the totalities of information available at the index sets $\{su, su_1\}$ and $\{su, su_2\}$.

When we consider the first of these sets, we are dealing with a situation where, according to $S$, on the one hand $U$ knows that $\neg p$ since $F_{su_2 neg}(p) = 1$, but on the other hand $U$ knows that $p$ since $F_{su_1}(p) = 1$. Therefore, the submodel $M_{su_1}$ can be viewed as an inconsistent part of the DMN. The only possibility that $S$ can entertain consistently corresponds to the index set $\{su, su_2\}$. We therefore define a DMN model to represent a belief on the part of $S$ only if that belief is true according to the relevant consistent
sets of valuations.

It turns out that the collections of valuations which are relevant in the truth-conditional semantics of 2a-DpL are of a the same kind as the collections that we see in this example. Such collections of function clusters form small parts of DMN-submodels, which we call truncated submodels. Their precise definition is as follows.

**Definition.** A truncated submodel of a submodel \( M_\alpha = \langle F_\alpha, \mathcal{F}, A \rangle \) is a triple \( T = \langle F_i, \mathcal{F}', A' \rangle \), where \( A' \) is \( A \) restricted to \( \mathcal{F}' \) and \( \mathcal{F}' \) is the smallest subset of \( \mathcal{F} \) defined by:

(i) \( F_i \in \mathcal{F}' \);
(ii) if \( F_j \in \mathcal{F}' \) then \( F_{j \text{ neg}} \in \mathcal{F}' \);
(iii) if \( F_j \in \mathcal{F}' \) and \( A(F_j) \neq \emptyset \) then there is exactly one \( F_k \in A(F_j) \) such that \( F_k \in \mathcal{F}' \);
(iv) if \( F_j \in \mathcal{F}' \), where \( j \) is a complex index decomposable as \( j = gh \), with \( h \) not complex, and if \( A(F_g) \neq \emptyset \) then there is exactly one \( F_k \in A(F_g) \) such that \( F_k \in \mathcal{F}' \).

The index \( i \) of the truncated submodel \( \langle F_i, \mathcal{F}', A' \rangle \) will be called the index of the truncated submodel.

The set of all consistent truncated submodels with index \( i \) of a given (sub-)model \( M_\alpha \) will be denoted by \( CTS(i) \).

In words, a truncated submodel with index \( i \) is that part of the submodel that contains the information locally available at index \( i \), plus what is available at one of the alternative extensions at that index; moreover, if the index is complex, say \( i = su_s \), we also take the last step back in building up \( i \), take one alternative extension at the shorter index, say \( su_3 \), and add to that the operation used to build \( i \), so in the example we get \( su_3s \).

The fact that the semantics of 2a-DpL can be formulated in terms of truncated submodels emphasizes that DMN-models to a great extent work in a modular and local fashion.

**5.3 Truth in a DMN model**

The definition of when a 2a-DpL-expression of the form \( S \models \phi \) is true and when it is false in a DMN model will be given recursively in terms of the truth in certain consistent truncated submodels.
A DMN model $M$ as a whole behaves classically in that any formula of the form $S \models \phi$ is either true or false in the model. Truncated submodels, on the other hand, have a three-valued logic: given a consistent truncated submodel $T$, a well-formed 2a-DpL formula $\phi$ can be true in $T$, false in $T$, or undefined in $T$. We will use the notation $T \models \phi$ in the first case and the notation $T \vDash \phi$ in the second.

The definition starts off at the level of $M$, as follows:

**Definition.** The truth of a 2a-DpL formula expressing a belief on the part of $S$ in the DMN-model $M$ is defined by:

$$M \models S \vdash \phi \iff T \models \phi \text{ for every } T \in \text{CTS}(s), \text{ and } \text{CTS}(s) \neq \emptyset.$$

In defining the truth conditions relative to truncated submodels the notion 'consistent U-extensions' of a truncated submodel is used. A $U$-extension of a truncated submodel with index $i$ is a truncated submodel with index $i_\sharp$; we use the notation $\text{CUX}(T)$ to denote the consistent $U$-extensions of $T$; similarly for $S$-extensions.

**Definition.** The truth of a 2a-DpL formula in a truncated submodel $T_\sharp$ is defined as follows.

(12) For any truncated submodel $T = < F_i, \mathcal{F}, \mathcal{A} >$:

1. If $p$ is a propositional constant:

   $$T \models p \iff F_i(p) = 1 \text{ for some } F_i \in \mathcal{F}, \text{ and }$$
   $$\text{for all } F_i \in \mathcal{F}_T: F_i(p) \neq 0$$

   $$T \vDash p \iff F_i(p) = 0 \text{ for some } F_i \in \mathcal{F}, \text{ and }$$
   $$\text{for all } F_i \in \mathcal{F}_T: F_i(p) \neq 1$$

The remaining clauses apply to any (consistent) truncated submodel $T$ and 2a-DpL-expressions $\phi, \psi$. 

20
2. $T \models \phi \land \psi \iff T \models \phi \text{ AND } T \models \psi$

$T \models \phi \land \psi \iff T \models \phi \text{ OR } T \models \psi$

3. $T \models \phi \lor \psi \iff T \models \phi \text{ OR } T \models \psi$

$T \models \phi \lor \psi \iff T \models \phi \text{ AND } T \models \psi$

4. $T \models \neg \phi \iff T \models \phi$

$T \models \neg \phi \iff T \models \phi$

5. $T \models U \models \phi \iff$ for all consistent $U$-extensions $T'$ of $T$:

$T' \models \phi$, and $CUX(T) \neq \emptyset$

$T \models U \models \phi \iff$ for all consistent $U$-extensions $T'$ of $T$:

$T' \models \phi$ and $CUX(T) \neq \emptyset$

5'. Similarly for $T \models S \models \phi$ and $T \models S \models \phi$.

### 6 Building and updating DMN models

Above, we have defined what it means for a DMN model $M$ to satisfy a formula. When we want to define the update function $u$, for updating a model with a formula known to be true, we are faced with the problem that the model may already verify or falsify that formula. If it already verifies the formula, nothing needs to be done, but what if the model falsifies the formula? It would seem to depend on 'tactical' considerations what to do in this case; this is not so much a matter of logic. I therefore leave this open here, defining the update function only in case the model does not contradict the formula. This update function thus covers, besides the trivial case where the model already verified the formula, only the case where the model was underdetermined with respect to the formula. In the practice of information dialogues, this is the most important case.

The update function $u$ takes two arguments, a (sub-)model $M_i = <F_i, \mathcal{F}, \mathcal{A}>$ and a formula $\phi$, and delivers a new (sub-)model $M'_i = <F'_i, \mathcal{F}', \mathcal{A}'>$ that satisfies $\phi$. We will use the notation $M_i[\ldots]$ to denote the submodel that is equal to $M_i$ except (at most) for the properties stipulated inside the square brackets.

**Definition.** The update of a model $M_i = <F_i, \mathcal{F}, \mathcal{A}>$ with a 2a-DpL formula $\phi$ is defined as follows.
For any propositional constant \( p \) and index \( i \):

1. \( u(M, p) = M_i[F'_i = F_i \cup \{< p, 1>\}] \)

The remaining clauses apply to any DMN-submodel \( M_i \) and 2a-DpL-expressions \( \phi, \psi \).

2. \( u(M_i, \phi \land \psi) = u(u(M_i, \phi), \psi) \)

3. \( u(M_i, \phi \lor \psi) = \) if \( M_i \) has alternative extensions for index \( i \), then for each of these, with index \( j \), do
   \( u(M_j, \phi \lor \psi) \); else \( M'_i = M_i[A(F_i) = \{F'_i, F_i\}] \), and do
   \( u(u(M'_i, \phi), \psi) \)

4. \( u(M_i, \neg \phi) = u'(M_i, \phi) \) (See below for the function \( u' \).)

5. \( u(M_i, U \models \phi) = u(M_{i\phi}, \phi) \)
   \( u(M_i, S \models \phi) = u(M_{i\phi}, \phi) \)

The function \( u' \) builds up the 'negative parts' of a DMN model, and is defined as follows.

(13b) 1. For propositional constant \( p \):
   \( u'(M_i, p) = u(M_{i\neg p}, p) \)

2. \( u'(M_i, \phi \land \psi) = \) if \( M_i \) has alternative extensions for index \( i \), then for each of these, with index \( j \),
   \( u(M_j, \neg \phi \lor \neg \psi) \); else \( M'_i = M_i[A(F_i) = \{F'_i, F_i\}] \), and do
   \( u(u(M'_i, \neg \phi), \neg \psi) \)

3. \( u'(M_i, \phi \lor \psi) = u'(u'(M_i, \phi), \psi) \)

4. \( u'(M_i, \neg \phi) = u(M_i, \phi) \)

5. \( u'(M_i, U \models \phi) = u(M_{i\neg p}, \phi) \)
7 DMN models and dynamic interpretation

7.1 Beliefs and intentions

When we try to use DMN models for representing the changes in an interpreter's information state, and thereby for formalizing meaning, we immediately see that the models defined in the previous sections are too simple. This is for two reasons.

First, our 2a-DpL models are suitable only when we use the language of propositional logic for representing the factual information ('propositional content') of utterances. So we have to replace the 'pL' of 2a-DpL by a more powerful language, with predicate-argument structures, generalized quantifiers, modal operators, provisions for non-individuating expressions (mass terms, collectives), etc.

Second, we have argued earlier in this paper that the changes in an interpreter's information involve different kinds of beliefs and intentions. Besides the kind of belief that corresponds to to have the information that we must also take weak belief ('suspicion') into account, and mutual belief. Furthermore, the perhaps most important information about the speaker that an interpreter should distill from the utterance is the speaker's intentions. In an information dialogue, these intentions always relate to obtaining or providing information, factual as well as communicative information. So far, in 2a-DpL we have only considered the attitude to have the information that; we have not yet considered the other belief attitudes and intentions.

The 'suspicion' attitude can, evidently, be handled in much the same way as the attitude to have the information that, which from now on we will consequently call to know. Since our interpretation of know is nonfactual, it has much the same logical properties as suspect. There are only small differences, such as the positive introspection axiom: for know this is (14); for suspect this should not be the corresponding formula (15), but rather (16):

\[
(14) \quad S \text{ knows that } p \implies S \text{ knows that } S \text{ knows that } p
\]
\[
(15) \quad S \text{ suspects that } p \implies S \text{ suspects that } S \text{ suspects that } p
\]
\[
(16) \quad S \text{ suspects that } p \implies S \text{ knows that } S \text{ suspects that } p
\]

Properties like (14) have not been built into the truth definitions given in the previous section, but they might be.\(^3\)

\(^3\)In fact, some properties like this one should be built in, but there are several choices one can make here, just like there are many axiomatisations of epistemic and doxastic
The addition of the suspect attitude would be straightforward, were it not for the interactions between the know and the suspect attitudes as well as with other attitudes we will have to add. An interpreter’s suspicions should not contradict his knowledge:

\[(17) \ S \text{ knows that } p \implies S \text{ does not suspect that not } p\]

This should be built into the truth definitions. Furthermore, one can want to know something, but one does not want to suspect something. We will not pursue the addition of the suspect attitude any further here, since it mainly has the effect of making DMN-models larger and truth- and update conditions longer, without adding much of interest.

Mutual knowledge can be viewed as a particular form of knowledge. To say that $S$ and $U$ mutually know that $p$, is to make a statement about both $S$’s and $U$’s state of information. Such a statement can form part of a description by an external observer of the dialogue, but not of the information of a participant. For a participant, mutual knowledge can only occur as a belief on his part. By definition, if $S \text{ knows that it is mutually known by } U \text{ and him that } \phi$, then also $S \text{ knows that } U \text{ knows that } \phi$. This illustrates that the facts $S$ knows to be mutually known by $U$ and him form a subset of the facts $S$ knows that $U$ knows, as well as of the facts $S$ knows that $U$ knows that $S$ knows, etc. We can thus add mutual knowledge to 2a-DpL and its models by adding additional valuations with the appropriate indices (such as $F_{\mu, \mu}$ for the facts $S$ knows that $U$ knows that $S$ and $U$ mutually know), provided that we make sure to express in the truth definitions that mutual knowledge is a form of nested ordinary knowledge.

We now turn to the modelling of intentions, which is both conceptually and technically of more interest.

### 7.2 Intentions

The logic of intentions, goals, wants, desires, etc. has not been well developed yet. These attitudes come in subtly varying forms, and present tricky logical problems. When we restrict ourselves to information dialogues, however, we can focus on particular forms of intention of which a formalization appears to be quite feasible.

In general, an important difference between intentions and goals on the logic one can choose; I have formulated DMN-models here in a very sober way, in this respect.

24
one hand, and wants and desires on the other, is that the latter can conflict with one's beliefs, whereas the former cannot. One can very well desire to become the king of France, even though one does not believe this to be possible; to intend or to have the goal to become king of France is not well possible, however. Also, one can have different, conflicting desires, but it does not seem possible to have conflicting intentions. Another difference is that intentions (more clearly than goals) can only relate to situations over which one has control. One can not intend that the sun shines, but one can desire it. For more about the nature of intention see Bratman (1989).

Not all intentions are of the same character. Intentions can be 'static', such as the intention to stay home tonight; or they can be 'dynamic', such as the intention to bring something to somebody's attention. Intentions of the latter kind may be called 'goals'; it seems inherent to the nature of a goal that it is something that has not yet been achieved: as soon as a goal has been achieved, it is no longer a goal. Dynamic intentions (goals) give rise to actions, and are of interest here because the interpretation of the communicative actions in a dialogue involves the understanding of the speaker's intentions. We assume that, at least in an information dialogue, every communicative action owes its existence to some underlying intentions, and that the 'logic' of dialogues can largely be reduced to the logic of these intentions.

What sort of intentions can one have in an information dialogue? Since the participants in such a dialogue by definition are supposed to have no other purposes than exchanging certain factual information, we can d priori identify two kinds of possible goals for a participant: the participant has obtained certain factual information, or he has achieved that the partner possesses certain factual information. However, we have also seen earlier in this paper that the activity of participating in a dialogue gives rise to situations where one wants to clarify, verify, explain, etc., i.e. where one wants to obtain or to provide information which is not of a factual nature (not relating to the domain of discourse), but relates to the communication, leading to dialogue control acts. So, the kinds of goals that arise for an agent in an information dialogue are:

A the agent possesses certain information, factual or communicative, which he did not possess before;

B the dialogue partner possesses certain information, factual or communicative, which he did not possess before.
These intentions are not specifically communicative in nature: one may very well have such intentions (notably the first one) in other than dialogue situations. This observation has led Edwards & Mason (1989) to reject A and B as intentions in a dialogue, I think not quite correctly, for the following reason. Consider the situation where agent S has a goal of type A, such as to know the way to the theatre. He is standing on a street corner, a map in his hands, looking puzzled, while another person U is approaching. It so happens that U is in a hurry. Now as long as S does not ask U for help, perhaps because he notices that U is in a hurry, U has the possibility to ignore S (maybe he didn't notice S clearly, because he is in a hurry). But as soon as S has asked U for help, U is morally obliged to do something about it. Now does S, in the case where he accosts U, have a different goal than in the case where he doesn't? Edwards & Mason would appear to say yes: in the former case S has the goal to obtain the information he wants through a dialogue with U. Therefore he addresses U.

Edwards & Mason's analysis suggests that in fact every dialogue act is motivated by an intention with an 'appendix' of the form ".. through a dialogue with ..". I'm afraid that, first of all, this analysis can be extended ad libitum to include other features of the action besides the fact that it forms part of a dialogue with a particular partner. When S in the above example says: "Excuse me, what is the way to the theatre please?", one may, following the line of Edwards & Mason, construe S's intention as: to know the way to the theatre through a spoken dialogue beginning with the utterance "Excuse me, what is the way to the theatre please?", spoken clearly by me with a polite and pleasant intonation, and addressed to the person who is currently approaching, when he is at a distance of one meter, after I have taken three steps in his direction. All that matters for the interpreter, however, is that the speaker wants to know the way to the theatre; the rest is either irrelevant or follows from the standard conventions for communication. Since the 'appendix' of the intention, which is highly arbitrary in its detail anyway, does not have to be acknowledged explicitly by the interpreter, it is justified to consider it as not forming part of the meaning of the utterance, and to restrict the intentions to be taken into account to those of types A and B.

When we try to formalize such intentions in an extension of 2a-DpL, we should take into account that expressions like S intends that φ are well-formed only if φ expresses one of the two types of intention just mentioned. So an expression like S intends that ~ITALIAN(john) is incorrect. Since an intention operator may only occur in combination with an agent obtain-
ing information or making something known, it is technically simplest to introduce attitudes for these combinations: the attitudes \textit{wanting to know something} and \textit{wanting to make something known}. The first of these can be construed mathematically as an \textit{extension} of what the agent in question knows, as one can only want to know something one doesn’t know; the second as a \textit{restriction} of what one knows if we are to model that one can only make something known which one knows.\footnote{The assumption that one can only make something known which one knows would be reasonable if we restrict ourselves to information dialogues. The possibility to distinguish a lie from a sincere inform is in conflict with this assumption, though. Another way of saying the same is that this assumption means that the complex attitude \textit{want to make known that} is not quite the same as the complex attitude \textit{want that the addressee believes that}.}

We will use the following notations for these attitudes:

\begin{align}
S \leftarrow X & \overset{\text{def}}{=} S \text{ wants to know the value of } X \\
S \longrightarrow X & \overset{\text{def}}{=} S \text{ wants to make known that } X 
\end{align}

We will call the language 2a-DpL, extended with these intentional operators: 2a-DpL \textit{(Two-agent Doxastic-Intentional propositional Language)}.

We will not go systematically into the logical properties of these two intentional operators here, but only list a few of the most interesting ones.

\begin{align}
(18) & \quad S \leftarrow p \quad \implies \neg S \leftarrow \neg p \\
& \quad S \longrightarrow p \quad \land \quad S \leftarrow (p \rightarrow q) \quad \implies \neg S \longrightarrow \neg q \\
& \quad S \leftarrow p \quad \implies \neg (S \leftarrow \neg p) \quad \lor \quad S \leftarrow \neg p \\
& \quad S \leftarrow p \quad \land \quad S \leftarrow (p \rightarrow q) \quad \implies S \leftarrow q
\end{align}

7.3 Intentions in DMN models

To add the intentions discussed above to DMN models, we have to extend the definition of models with the relevant extensions and restrictions of information, corresponding to what an agent wants to know and wants to make known. We have to add the corresponding valuations and indices, and make additions to the truth definition and update conditions.

\textbf{Definition.} A \textit{DMN model for 2a-DpL} is a quintuple
(20) \[ M = \langle F_\alpha, \mathcal{F}, A, \Gamma, \Lambda \rangle, \] with:
- \( F_\alpha, \mathcal{F}, \) and \( A \) as before;
- \( \Gamma \) and \( \Lambda \) partial functions from \( \mathcal{F} \) into \( \mathcal{F} \).

The function \( \Gamma \) specifies extensions of available information that a certain agent intends to obtain (according to the other, perhaps); the function \( \Lambda \) specifies restrictions of available information that an agent intends to make known. For instance, if \( \Gamma(F_{ru}) \) is defined, and \( \Gamma(F_{ru}) = F_{ru} \), then \( F_{ru}(p) = 1 \) represents that \( S \) knows that \( U \) wants to know whether \( p \).

Analogous to the consistent \( U \)-extensions \( CUX(T) \), \( S \)-extensions \( CSX(T) \), and negative extensions \( CNX(T) \) of a truncated submodel \( T \) we define the consistent \( \Gamma \)-extensions \( C\Gamma X(T) \) and \( \Lambda \)-restrictions \( C\Lambda X(T) \) of a truncated submodel \( T \).

We will skip the formal extension of the definition of the set \( IM \) of the indices of a model here, and only note that we use the index names \( i_\gamma \) and \( i_\lambda \) to refer to the \( \Gamma \)-extension and \( \Lambda \)-restriction at index \( i \), respectively, when these are defined by the model.

The truth-conditional semantics of the intentional operators is defined as follows:

(21) 6. \[ T \models U \models \phi \iff \text{for every consistent } \Lambda \text{-extension } T' \text{ of a consistent } U \text{-extension of } T: T' \models \phi \]

7. \[ T \models U \models \phi \iff \text{for some consistent } \Lambda \text{-extension } T' \text{ of a consistent } U \text{-extension of } T: T' \models \phi \]

6'. Similarly for \( T \models S \models \phi \) and \( T \models S \models \phi \).

7'. Similarly for \( T \models S \models \phi \) and \( T \models S \models \phi \).

In addition, we may add to the truth definitions of \( U \models \phi \) and \( S \models \phi \) clauses expressing that \( \phi \) is known if the agent wants to make \( \phi \) known.\(^5\)

The update conditions for the new operators are as follows:

\(^5\)Cf. previous footnote.
6. \( u(M_i, U \models \phi) = \)
if \( M_i \) has a \( \Gamma \)-extension for index \( i \),
then do \( u(M_{i\Gamma}, \phi) \);
else \( M_i' = M_i[\Gamma'(F_i) = F_{i\Gamma}] \), and do
\( u(M_{i\Gamma}', \phi) \)

7. \( u(M_i, U \models \phi) = \)
if \( M_i \) has a \( \Lambda \)-extension for index \( i \),
then do \( u(M_{i\Lambda}, \phi) \);
else \( M_i' = M_i[\Lambda'(F_i) = F_{i\Lambda}] \), and do
\( u(M_{i\Lambda}', \phi) \)

7.4 Meaning representations using 2a-DIpL

Now that we have provided the tools to represent the kinds of intentions and beliefs that are most important for describing the flow of information between the participants in an information dialogue, we also have, according to the DIT approach, the tools to represent utterance meanings.

When doing so, we must be careful to distinguish the representation of actual changes in the participants' information from the representations of meanings of the utterances that cause the changes in information. In describing actual flows of information in dialogues one has to take into account such factors as misunderstandings, imperfect use of contextual information, relative social power of the participants, etc. A theory of meaning, on the other hand, should not be burdened with such considerations. Therefore, in a theory of meaning we abstract from certain dimensions of actual communication, and restrict ourselves in the kind of changes in information states that are taken into consideration. In section 2 we already argued that, for this reason, a theory of meaning should only consider the changes in interpreters who are ideal, rational understanders of the language. Similarly, we will also assume that there are no particular power relations between speaker and interpreter that have a distorting influence on the effects of the speaker's utterances.6

A further important restriction is that only those effects of utterances are taken into account that occur as the result of interpreting the utterance, not those effects that result from subsequent evaluation (cf. section 1) or reinterpretation. The latter point brings up the debatable issue of whether interpretation should be restricted to literal interpretation. Full-blown utterance interpretation should take not only literal meanings into account, but

---

6 Perhaps the best way to put this more precisely is that speaker and hearer are assumed to have equal social power.
also metaphors, irony, indirect speech acts, and conventional/conversational implicatures. The question is not so much whether all this should be taken into account, but how. The way this question is answered, is closely related to what is considered to be part of an utterance. A well-known view is that an utterance is a sentence plus its context; this view is sometimes formalized by construing an utterance $u$ as a pair $<s,c>$ where $s$ is a sentence and $c$ a context (see Levinson, 1983). On this view, the entire context is part of the utterance. I find this counterintuitive for various reasons, and prefer the view that an utterance has a number of semantic, pragmatic as well as other properties, which are the result of a complex interaction between properties of the linguistic expression that is used (syntactic, lexical, prosodic, .. properties) and properties of the context (linguistic as well as nonlinguistic context). Among the semantic and pragmatic properties of an utterance we have literal meaning, dialogue control function, indirect interpretations, presuppositions, etc.

I want to focus here on direct interpretation, thereby excluding indirect interpretations and dialogue control function, as explicated in speech act theory in terms of communicative function and semantic content (Searle, 1969). When we consider only those aspects of an interpreter's information that relate to utterance meaning in this restricted sense, we can use the tools of 2a-DIpl and its DMN models to formalize utterance meanings.

Let us now consider three examples, corresponding to the three major classes of communicative action types in information dialogues, distinguished in Bunt (1986): question, inform and answer. The same sentence

7Among the reasons for finding this approach counterintuitive are the observation that an utterance is a concrete, observable event which a pair $<s,c>$ is not, and that it just doesn't seem right that the entire context, which includes such abstract things as the speaker's and hearer's beliefs about each other's intentions, be part of an utterance. Also, we mean by an utterance is in my opinion the result of an interaction between a linguistic expression and a context, an interaction which is more complex than what is captured by the pair-forming operator.

8Following Bunt (1986), the communicative acts in an information dialogue can be roughly divided into factually informative acts (FI acts) and dialogue control acts (DC acts). In fact, every communicative act is best viewed as having both an FI-aspect and a DC-aspect, where one of the two is the primary function of the act. For instance, an answer to a question is an FI-act with the DC-aspect of (implicitly) acknowledging correct understanding of the question. What is meant here by saying that dialogue control function is excluded here is that for FI-acts their DC-aspect is not considered.

By 'direct interpretation' I mean that interpretation which, as far as its communicative function is concerned, is most straightforwardly related to the linguistic features of the utterance (cf. Beun, 1989).
can be used with either of these functions, so let us take for all three cases the sentence:

(23) *John is sleeping*

Let us assume that this sentence is used by *U* and addressed to *S*, with the respective communicative functions *QUESTION*, *INFORM* and *ANSWER*. So we consider the utterance of (23) with the following labelling, indicating function and content:

(24) a. <QUESTION, SLEEP(john)>
b. <INFORM, SLEEP(john)>
c. <ANSWER, SLEEP(john)>

How the meanings of these utterances should be represented in 2a-DIpL depends on the precise analysis of the information conveyed. Minimally, there is the effect the speaker ultimately intends to achieve with his utterance. In addition, there are the speaker's beliefs which, for instance, distinguish a genuine informative question from an exam question or a rhetorical question, and a genuine inform from a lie. Following the analysis in Bunt (1986), this part of the meaning leads, for (24) a, b, and c to the packages of beliefs and intentions represented in the left-hand sides of (25) a, b, and c as 2a-DIpL expressions, with their information changing effects described in the right-hand sides as updates of the DMN-model that represents the interpreter's state of information. The notation *U* \(\models p\) is used here as abbreviation for *U* \(\models p \lor U \models \neg p\) (in words: 'U knows whether p', abbreviating *U* knows that *p* or *U* knows that not *p*), and the notation *U* \(\models \cdots p\) is used for *U* suspects that *p*.

(25) a. \(\models U \models \neg SLEEP(john) \land \neg U \models \neg SLEEP(john)\)

\(\models U \models \cdots S \models \neg SLEEP(john)\) =

\( = u(M, U \models \neg SLEEP(john) \land \neg U \models \neg SLEEP(john) \land U \models \cdots S \models \neg SLEEP(john))\)

b. \(\models U \models SLEEP(john) \land U \models \neg SLEEP(john)\)

\(\models U \models \neg S \models \neg SLEEP(john)\) =

\( = u(M, U \models SLEEP(john) \land \neg U \models \neg SLEEP(john) \land U \models \neg S \models \neg SLEEP(john))\)

c. \(\models U \models \neg S \models \neg SLEEP(john) \land U \models \neg SLEEP(john)\) =

\( = u(M, U \models SLEEP(john)) \land U \models \neg SLEEP(john))\)
According to the full analysis of information transfer in Bunt (1986) there is more than this, since the speaker expects that his utterance will be correctly understood and give rise to mutual knowledge concerning the speaker's beliefs and intentions; moreover, speaker and hearer both know that the speaker expects this - the hearer's knowledge on this point constitutes part of his new information, resulting from the interpretation of the utterance. With the addition of mutual belief to DMN models, as indicated above, this can be accounted for in a similar way.

Some authors have proposed more complex kinds of intention, according to which a speaker has the intention to achieve a certain effect through the addressee's recognition of that intention. For instance, Beun (1989) ascribes such an intention to questions: (U = speaker, S = addressee)

\[(26) \quad QUE \equiv I_U K_S I_U \overline{K}_U p\]

which says that U intends that S knows that U intends that U knows whether p. Using the intention operators of 2a-DIPL, this can be expressed as:

\[(27) \quad QUE \equiv U \rightarrow U \leftarrow p\]

The update-conditional semantics of the right-hand side of this expression, with the semantic content of the question substituted for p, gives us the desired part of the dynamic utterance meaning.

Other views, such as those of Cohen and Levesque (1985) or Perrault (1989) can accommodated in a similar way. More detailed examples can be found in Bunt (1989).

8 Dynamic theories of meaning: methodological issues

Having presented Dynamic Interpretation Theory (DIT) in outline in the preceding sections, we now return to some of the methodological issues touched upon in the Introduction. I will do this mostly in the form of a discussion of some of the fundamental differences that exist between DIT and Dynamic Semantics (DS), in spite of the common starting point where meaning is viewed as information-changing potential. To this discussion I will add some other considerations which are relevant to the further devel-
opment of a dynamic theory of meaning.

8.1 Factual and 'second-order' information

The starting point of dynamic approaches to meaning being that the meaning of a linguistic expression is the way it may change an interpreter's information, two fundamental question arise:

(i) What do we mean in this context by 'linguistic expression';

(ii) In what ways do linguistic expressions in fact change an interpreter's information?

It was noticed in the Introduction that Dynamic Interpretation Theory answers the first question by: utterances, whereas Dynamic Semantics answers it by: sentences. This difference is perhaps historically explained by the circumstance that DS originated in an attempt to do better than Kamp's Discourse Representation Theory in providing a compositional treatment of anaphora and related phenomena as they occur in short texts, consisting of a few sentences (see Groenendijk and Stokhof, 1987), whereas DIT developed from an attempt to develop a theory of dialogues in terms of the information states of the participants, changing under the influence of utterances (see Bunt, 1986). However, the history is not our primary concern here.

We have argued in section 2.1 that utterances are the linguistic elements that may change an interpreter's information state, rather than sentences, since the same sentence can be used in different contexts as a different utterance, having a different effect on an interpreter. I think this is in fact quite obvious when we consider dialogues. The information-changing effect of a sentence, occurring in a dialogue, depends on the purpose with which the speaker utters the sentence, as perceived by the interpreter. For texts the situation is not really different. When interpreting the sentences in a text the purposes of the author matter equally much: it makes all the difference whether the text is an advertisement, an obituary, a medical prescription, a joke, an exam, etc. In fact, the recognition of this point appears to be dawning already in Groenendijk and Stokhof's programmatic statement, quoted above and repeated here:

"The meaning of a sentence does not lie in its truth conditions, but rather in the way it changes the (representation of the) information of the interpreter. The utterance of a sentence brings us from a certain state of information to
another. The meaning of a sentence lies in the way it brings about such a transition." (Groenendijk and Stokhof, 1987, p.1, my italics.)

The italicized part suggests an awareness that in fact information-state changes are brought about by utterances, rather than by sentences; nonetheless, the quotation clearly expresses a view on sentence meaning, rather than on utterance meaning.

It may be thought that the DS preoccupation with sentences rather than utterances is simply a matter of restricting the attention to declarative sentences, interpreted as making true, sincere informative statements. This is not the case, however, as we can see from the following reply by Groenendijk and Stokhof to Bunt (1989a), where DS was criticized for applying the information-change approach to sentence meaning. First, Groenendijk and Stokhof characterize the difference between DS and DIT as one between `semantic' and `pragmatic' interpretation, where the latter takes the communicative function of an utterance into account:

"How, then, does dynamic semantics relate to Bunt's notion pragmatic interpretation, which distinguishes between semantic content and communicative function? Part of the role of the pragmatic aspect of meaning that is embodied in the notion of communicative function as a function that changes information states, is covered in dynamic semantics in the semantic content of an expression. For as we have seen above, a notion of update of information is an ingredient of the semantic interpretation of sentences in the dynamic semantics framework." (Groenendijk and Stokhof, 1989, p. 484)

What is said here is, in my opinion, rather misleading. What is called the 'pragmatic aspect of meaning', viz. the intentions and beliefs characterizing the communicative function of an utterance, is, contrary to what is claimed, not covered in dynamic semantics in the semantic content of an expression; instead, it is completely neglected! We will see this more clearly below. Also, the 'explanation' that this be so since DS uses a notion of update of information is clearly besides the point, the point being that DS only considers changes in factual information, whereas taking communicative functions into account requires the consideration of changing information concerning intentions and beliefs relating to factual information.

That there is a fundamental issue at stake here becomes especially clear in the following passage from Groenendijk and Stokhof (1989), which is the continuation of the text quoted above.
“However, it should be added right away that this notion of update is more restricted than Bunt would have it. It only presents the update of the information of the addressee with the content of the discourse, but it does not concern what we might dub ‘second order’ information, such as information about the information of other speech participants. And a second difference is that where communicative function comes into the picture of dynamic semantics at all, it is only in the form of one particular function, that of assertion.” (Op. cit., 484, my italics.)

This quotation makes it clear that Groenendijk and Stokhof hold the view that they provide meanings for declarative sentences, used as assertions, in terms of dynamic information states which only contain information about the domain of discourse; information about the intentions and beliefs which distinguish assertions from other speech act types is not present in their information states. It seems to me that the Dynamic Semantics approach is fundamentally inadequate in this respect. If an interpreter’s information states do not contain information about beliefs and intentions, than how can the interpreter possibly interpret an utterance as an assertion, rather than some other type of speech act? What else is the difference in meaning between, say, an open question and a verification (with the same semantic content) than that in the latter case the speaker has an expectation concerning the truth of the semantic content of his utterance, whereas in the former case he hasn’t? In other words, it makes no sense to view meaning in terms of information change while leaving ‘second-order’ information out of consideration; the changes an utterance brings about in an interpreter are always in the first place changes in ‘second-order’ information.

8.2 (Non)determinism and the (in)adequacy of partial models

In the Introduction, we briefly touched upon the fact that in DIT the meaning of an utterance is construed formally as a function from information states to information states, whereas in DS meanings are construed as sets of state pairs, consisting of an input and an output state. This may seem to be a small technical difference, but there is more to it than a technical difference only. To see this, consider the DS interpretation of the sentence A man is sleeping, represented in predicate language as (28):
In DS, this formula is interpreted as denoting the set of input-output pairs such that in the output state the variable $z$ is assigned a value belonging to the extension of the predicate:

$$\exists z : \text{MAN}(z) \land \text{SLEEP}(z)$$

What does this tell us about what happens when an interpreter interprets the utterance *A man is sleeping*, pretending for a moment that this can be accounted for in terms of changing factual beliefs, as DS supposes? Well, let the interpreter be in the state $g_1$ before the utterance is addressed to him. Although DS does not explicitly claim to be a theory about actual changes in an interpreter's information brought about by interpreting a linguistic expression, the programmatic statement quoted above strongly suggests that actual changes do indeed correspond to input-output pairs of DS, at least for interpreters who understand expressions correctly. This means, in the present example, that our interpreter would get from his input state $g_1$ to one of the final states $h_i$, connected to $g_1$ according to (29). Each of the output states $h_i$ is characterized by being identical to $g_1$ except that in the output state the variable $z$ is assigned some man who is sleeping.

Two things are strange about this. Firstly, what is this `z' that suddenly turns up in the output states? Why `z', why not `y', `z', or any other variable? How is this variable related to the sentence *A man is sleeping*? Secondly, which of the possible output states will our interpreter actually find himself in? Given a domain of discourse with three sleeping men, which of those will be associated with `z'? These questions do not seem to have satisfactory answers. For the first question, it is obvious that the variable `z' has no special relation to the sentence whatsoever; one could equally well choose a different variable to express the quantification. In DS, however, the choice of variable is significant for the meaning of the expression. For consider the interpretations of the following formulae:

$$\exists z : P(z) = \{<g,h> | h[z]g \land h(z) \in P \}$$

Clearly, the set of pairs in the right-hand side of (30)a. is not the same as that in (30)b.
So on one hand the choice of variable matters for DS, on the other hand the sentence provides no grounds (nor does DS elsewhere) for choosing one variable rather than another. When criticized on this point in Bunt (1989a), Groenendijk and Stokhof have replied that any time a quantified noun phrase is to be translated, a ‘fresh’ variable be picked. They furthermore suggest that

“A variable can be viewed as the name of an address. Each variable refers to a particular address. Instead of choosing a new variable in order to refer to a new address, we could introduce a pointer to addresses, and use it in the translation of a quantified phrase to ‘the next free address’.” (Op. cit., 471-472.)

This is all quite unsatisfactory, since none of the ideas ‘fresh variable’, ‘address’, ‘pointer’ or ‘fresh address’ is defined within DS. The mechanisms that are suggested here would have to be defined and built into the DS framework, which would undergo rather drastic changes as a result. The DS framework as it stands does not have a good answer to this criticism, and it remains to be seen whether this can be overcome.

These problems came up in relation to the puzzling questions, raised by the use of a particular variable in the representation (or interpretation) of quantified noun phrases in DS. In DIT variables have the same status as in classical logic, where the interpretation of a formula without free variables is independent of the choice of variable names, and so these problems do not arise.

The second question we raised concerns the fact that an interpreter, faced with an existential statement, seems to have to choose one of a collection of possible output states. This has two problems. First, why should an interpreter, upon interpreting a statement like *A man is sleeping* have to choose which man is sleeping, and secondly, how could he make a choice?

It seems to me that our interpreter shouldn’t have to choose at all. All he knows, if accepting that *A man is sleeping*, is that *some* man is sleeping; so his state of information should be changed accordingly. (Again, pretending, as DS has it, that interpretation of expressions changes factual information.) This means that there is only one possible output state, rather than a set among which the interpreter has to choose. There being one output state for each input state, the appropriate formal device to use is a function, therefore in DIT meanings are construed as state-changing functions rather than sets of input-output pairs.

37
In reply to Bunt (1989a), where a similar criticism was worded, Groenendijk and Stokhof have acknowledged that their approach is nondeterministic. Formulae of predicate logic, interpreted dynamically, "are in some cases essentially non-deterministic, relational, not functional. Executing them in a particular state can result in a number of alternative possible ones, where in each of them one or more variables are assigned a certain value." (Op. Cit., 473.) Why this should be a positive feature of DS escapes me; indeed, such examples as where an interpreter of the sentence *A man is sleeping* can (and must) choose between changing into an output state where John sleeps, or one where Tom sleeps, or one where Peter sleeps, seems to me to expose a weak point of DS rather than a strong one.

In DIT, as we have seen, such problems do not arise since we explicitly represent disjunctive information. So in the example just mentioned, an interpreter would (upon believing the speaker) come into the state where he knows that John, Tom, or Peter is sleeping.

In their reply Groenendijk and Stokhof (1989) have accused DIT of not being able to deal with this kind of situation adequately, because partial models, as used in DIT, would not have the same power as the possible-worlds models of DS. As they put it:

"Just using partial assignment functions, or 'partial worlds' instead of total ones, is not sufficient to get an account of the partiality of information. (...) For example, a partial assignment will leave the value of certain variables undefined. If the value of a certain variable is undefined, this can be viewed as a representation of a situation where we have no information at all about this value. And if it does assign a certain value to it, this means that we are completely informed about the value of this particular variable. But what about the case where we have partial information about the value of a particular variable, for example that it is either this object or that, but we don't know which? One single partial assignment cannot account for this situation." (Op. Cit., 460-461.)

Groenendijk and Stokhof are clearly right here, but also very inaccurate. It hasn't been claimed anywhere that incremental partial models should make use of one single partial assignment. In fact, sections 5 and 6 above make clear that truth conditions and update conditions for DMN models

---

9 In particular, this has certainly not been claimed in Bunt (1989a), to which the paper from which this quotation has been taken is a reaction.
require the use of sets of sets of partial valuations.\(^{10}\)

### 8.3 Semi-constant variables and variable constants

The most obvious difference between Dynamic Semantics and other theories is probably the treatment of variables. We have seen above that a quantified variable has a specific meaning of its own, much the same as constants in standard logics. This is precisely the reason why DS permits an elegant treatment of anaphora, using variables in a way which would be called 'free' in standard logics in that they fall outside the scope of a quantifier, but which works as if the variable is 'bound' nonetheless. As Groenendijk and Stokhof (1989) put it: "The choice of variable causes a difference in meaning. In this respect variables (..) behave more like constants than 'ordinary' variables." (Op. cit., 1989, p. 470.)

Variables also play a special role in DS in that their valuation is the only part of the interpretation which is dynamic: nonlogical constants receive a static interpretation. In Bunt (1989a) I have drawn attention to this limitation of DS, arguing that changing information about the extensions of descriptive constants is in fact much more important to take into account in a dynamic theory of meaning than changing value assignments to variables, since the information available to an interpreter is typically expressed in terms of the values of descriptive constants.

Groenendijk and Stokhof (1989) do not deny this, as they say: "Information about the extension of descriptive constants is really what counts in the end, it is what we are really interested in, it is the information we want to keep in memory." (Op. cit., 474) By way of counter-argument, on the other hand, Groenendijk and Stokhof remark that a dynamic interpretation of descriptive constant raises problems, and they suggest that it may not even be feasible, the main reason being that the values of constants can be interdependent whereas the values of variables are typically logically independent:

"It is pretty hard, rather implausible, and perhaps downright impossible to make sense of a notion like that of an interpretation function \( F' \) differing at most from an interpretation function \( F \) with respect to the predicate \( P \) in that \( a \in F'(P) \). How can we be sure that such a unique \( F' \) exists?" (Op.

\(^{10}\)Truth conditions are formulated with reference to sets of truncated submodels, which themselves involve sets of partial valuations; update conditions are formulated with reference to sets of submodels, which also involve sets of partial valuations.
Of course, in general no such unique $F'$ exists, since one change in $F$ may entail another. However, this is not necessarily a problem for DIT. The point is that the functions $F$ and $F'$, interpreting nonlogical constants, in DS represent states of the discourse domain, whereas in DIT the functions involved in dynamic meaning always represent an agent's information concerning the state of the speaker. This is an essential difference, that in fact already turned up when we considered the dynamic meaning in DS and in DIT of the sentence *A man is sleeping*: in a domain of discourse with three men, John, Tom and Peter, there is no single unique state which also makes that sentence true; however, there is one single state of information which corresponds to the sentence, viz. that state where it is known that John, Tom or Peter (or more than one of them) is sleeping. The point is that one state of information generally corresponds to a set of states in the discourse domain.

Still, it remains true that changing one belief may entail changing another, if an interpreter has beliefs about the relations between his beliefs. This complicates the matter, since changing one belief $p$ often entails changing either some belief $q_1$ or some other belief $q_2$, in order to preserve a consistent set of beliefs. This is a technical complication rather than a fundamental difficulty, however, since we have already seen that there can be a single state of information where the interpreter believes that $p$ and either $q_1$ or $q_2$.

So, when we agree that changing information about the values of descriptive constants is what we are really interested in, we should not be disconcerted by the technical complications that it entails and only treat variables dynamically, but we should deal with the technical complications. I would like to offer one suggestion here, which may be useful in reducing the number of alternative possibilities that have to be taken into account when descriptive constants are treated dynamically. This is the use of a heuristic principle, the Economy Principle for minimal changes in an interpreter's information. The idea is that, when the interpretation of an utterance forces interpreter $A$ to change not only the value of a descriptive constant $C_1$ in a certain way, but also that of some other constants in order to maintain a consistent information state, the Economy Principle allows $A$ to not necessarily take all logical possibilities into account for maintaining a consistent state, but only certain possibilities which involve 'minimal' changes in his beliefs. The following example may clarify this.
Consider the interpretation of the assertion:

(31) A male teacher is walking and whistling

given the domain of discourse modelled in (32):

<table>
<thead>
<tr>
<th></th>
<th>MAN</th>
<th>FARMER</th>
<th>SIT</th>
<th>EAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>MAN</td>
<td>TEACHER</td>
<td>SIT</td>
<td>?</td>
</tr>
<tr>
<td>Peter</td>
<td>WOMAN</td>
<td>TEACHER</td>
<td>SIT</td>
<td>?</td>
</tr>
<tr>
<td>Eve</td>
<td>WOMAN</td>
<td>TEACHER</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Billy</td>
<td>WOMAN</td>
<td>TEACHER</td>
<td>WALK</td>
<td>WHISTLE</td>
</tr>
<tr>
<td>Mary</td>
<td>WOMAN</td>
<td>WAITRESS</td>
<td>SIT</td>
<td>EAT</td>
</tr>
<tr>
<td>Chris</td>
<td>?</td>
<td>?</td>
<td>SIT</td>
<td>EAT</td>
</tr>
</tbody>
</table>

According to (32) there is no walking and whistling male teacher to start with, so to accomodate the factual information in (31) the extensions of one or more constants have to be adapted. Let us now consider what would be the minimal changes in the model that would make the proposition in (31) true, and what would be the most plausible changes.

The minimal change in terms of number of predicates involved would be that where Billy changes from woman to man. This is the only adaptation of the model that involves only one operation; all other possibilities involve at least two predicates. Is it also the most plausible change? I think not, since changing sex is a more drastic change than one from eating to whistling or from sitting to walking. Intuitively, the most plausible change seems to be the one where Peter is walking instead of sitting and the missing information about his activities is supplied as whistling.

What would be the most implausible change? This is clearly the one where Mary changes sex and profession, goes walking, stops eating and starts whistling. And of course, still more implausible are those involving more than one individual.

The implausibility of changes involving more than one individual is different in nature than the implausibility of Mary changing sex and profession. It is a matter of logical minimality not to involve more than one individual, in much the same way as the state transitions considered in Dynamic Semantics do not involve irrelevant variables; the other (im)plausibilities are semantic in character. The Economy Principle I have in mind is the conjecture that by default only semantically minimal changes have to be taken
into account. The question is then how the notion of semantically minimal change can be defined.

The example (31)-(32) gives us some indications of the criteria to be considered. First the sheer number of descriptive constants that would change is relevant: the change where Peter starts to walk and to whistle is smaller than the one where John does the same, plus changes profession. Second, the change from unknown value in the constant’s extension to known value is smaller than the change from one value to another: the changes needed to make Chris a walking and whistling male teacher are smaller than those for Mary. Third, and related to the previous one, the strengths of the beliefs represented in the model are relevant: if the belief that Billy is a woman is only a weak belief, while all the other beliefs are strong beliefs, than the change of Billy’s sex is not at all unlikely. Fourth, even if all beliefs are equally strong, some change more easily than others, due to the nature of the concepts named by the descriptive constants: one more easily changes job than sex, and while still more easily changes activities like eating, walking and whistling. Fifth, and finally, the topic-focus articulation in the utterance matters; the present example is not the best illustration of that, but it does exemplify this point a little: if the sentence A male teacher is walking and whistling is spoken with heavy stress on male, to indicate that something is asserted of a certain walking and whistling teacher, it becomes more likely that the model should change in that Billy is assigned male sex, since Billy is the only walking and whistling teacher.

The notion of ‘semantically minimal change’ in the Economy Principle should thus take at least the following criteria into account:

1. the number of descriptive constants involved;
2. the presence or absence of certain value-elements for descriptive constants;
3. the strengths of the beliefs concerning descriptive constants;
4. the believed ‘rigidity’ of the values of descriptive constants;
5. the topic-focus articulation of the sentential structures in the utterances that cause the changes.

Exactly how these, and perhaps other criteria together can be used to define a useful concept of semantically minimal change is an issue for further
study.
References


