Files in Focus

Comments on Enric Vallduví: 'The Dynamics of Information Packaging'

Paul Dekker	Herman Hendriks
ILLC	OTS
Amsterdam	Utrecht

Before we start to nag (and introduce clumsy abbreviations), we would like to stress that we enjoyed 'The Dynamics of Information Packaging' (henceforth: DIP) very much. The issues discussed in DIP are theoretically and empirically quite interesting, and the paper brings together several related lines of current research involving phonology, grammar architecture, discourse structure and dynamic semantics. We think Enric Vallduví formulates a perspective on the phenomena which certainly deserves more detailed and extensive study. The paper also raises new questions, some of which we want to touch upon in these comments. The present text is based on an earlier draft of the paper as it appears in this deliverable.

The paper starts with a general characterization of the way in which sentences package information. Typically, three parts can be distinguished: the *link*, which serves to direct the interpreter to a location in his or her information structure; the *tail*, which signals a certain mode of information update; and the *focus*, which encodes the information to be added in the information structure at the appropriate location and in the appropriate mode. The latter, non-optional part of a sentence S is said to represent *the information of* S, I_s, which is metaphorically described as the amount of propositional information expressed by S, minus the information already present in the hearer's model.[1] Together, link and tail constitute what is also called the *ground*, and the *link* corresponds approximately to the *topic* in the traditional *topic/comment* distinction. In a picture:

<table>
<tr><td></td><td>topic</td><td colspan="2">comment</td></tr>
<tr><td>≫</td><td>link</td><td>tail</td><td>focus</td></tr>
<tr><td></td><td colspan="2">ground</td><td>focus</td></tr>
</table>

Since sentences may lack links and tails, the following four sentence types can be distinguished:

1. More 'formally': $I_s = \phi_s - K_h$.

1. focus
2. link–focus
3. focus–tail
4. link–focus–tail

We will pass over the grammatical and phonological realization of these structures, and concentrate instead on the four instruction types with which they are associated.[2]

Links are associated with so-called GOTO declarations. In the framework adoped by DIP, the target locations of such declarations are 'file cards' *fc*. A tail points at an information record on such a file card, record(*fc*), and indicates that it has to be *modified* (or further specified) by the focus information I_s of the sentence. Normally this record is a (possibly underspecified) condition on a file card. The indicated instruction type is called UPDATE-REPLACE. In the absence of a tail, the focus information I_s of a sentence is simply *added* at the current location. The associated instruction type is called UPDATE-ADD. Summing up, the above four sentence types are associated with the following (compound) instruction types, respectively:

1. UPDATE-ADD(I_s) GOTO(*fc*)(UPDATE-ADD*(I_s))* UPDATE-REPLACE*(I_s,record(fc))* GOTO(*fc*)(UPDATE-REPLACE*(I_s,record(fc)))*

In these comments we would like to address some issues concerning the use of file cards. The second section will discuss (links and) GOTO declarations, and in the third section we will briefly comment on (tails and) UPDATE-REPLACE declarations. In the fourth section we sketch a card-less alternative. We conclude this first section with a few general remarks about the information structures employed in DIP: 'files'.

According to DIP, information is organized in files, which are represented as collections of file cards. Thus the metaphor employed in the file change semantics of Heim (1982, 1983) is taken literally. The file cards of which a file consists represent objects which may be left indefinite to a certain extent, and they are filled in with conditions on the objects they can denote.

Vallduví observes that files are 'dimensionally richer' than the discourse representation structures (DRSs) of discourse representation theory (see Kamp (1981), Kamp and Reyle (1993)). Now, this is potentially true to the extent that each file card introduces its own 'representational space' where all records concerning that file card are to be found. In order to be actually richer, nonetheless, files must be adapted to model more than merely individuals having properties and standing in relations at various spatiotemporal locations. Among other things, they should be able to model quantified, negative and disjunctive information. Discourse representation theory allows the construction of conditions from sub-DRSs, and these conditions—by an appropriate semantic interpretation procedure—model precisely such information. Heim, who explicitly speaks

2. Links, tails and foci will be specified by means of $[_{\mathrm{L}}\ldots]$, $[_{\mathrm{T}}\ldots]$ and $[_{\mathrm{F}}\ldots]$, respectively. Accented items in foci and links will be written in small caps and boldface, respectively.

of files and file cards as metaphors (1982, p. 276 and p. 302ff.), spells out quantified, negative and disjunctive information in purely semantic terms, i.e., in terms of the domains and satisfaction sets of files. However, it is not clear how such information must be expressed in the representational set-up adopted in DIP. For one thing, on what file card(s)—if any—should the information expressed by sentences such as *Every man walks*, *No man walks*, *Either John walks or Mary talks* be put? For another, how should it be put? One might think of using sub-files, but then, where must these be put? Are they attached to a main file, or must they be attached to a main file's file card? Which one? We leave it at these questions, in the hope that the subject will be given attention in due course.[3]

Vallduví's first two construction types can be illustrated with the following examples (the first one is DIP's example (6a)):

(1) [$_{\mathrm{L}}$The **president**][$_{\mathrm{F}}$hates the Delft CHINA SET]
(2) [$_{\mathrm{F}}$He always uses DISPOSABLES]

The first example is a link-focus construction, and therefore associated with a GOTO(fc)(UPDATE-ADD(I_s)) instruction. The link *the president* specifies a locus of update fc, viz., the card representing the president—say, card #125. The focus *hates the Delft china set* specifies the propositional information I_s that has to be added to this card. Suppose that the Delft china set is represented by card #136. Then, passing over some formal details, the UPDATE-ADD(I_s) instruction associated with the focus *hates the Delft china set* amounts to adding the condition 'hates(136)(125)' to the locus of update, i.e., the presidential card #125. Moreover, the record '$\boxed{\rightsquigarrow 125}$', a Hypercard-style link to the locus of update, is added to card #136, rendering the condition 'hates(136)(125)' 'accessible' from that card. Vallduví says that this linking mechanism, which designates a unique location for content update, is 'much more efficient' than straightforward multiple recording of information.

The second example is an all-focus construction, associated with a simple UPDATE-ADD(I_s) instruction. Here, this instruction involves the addition of the focus information I_s that the value of the current card always uses disposables. That is, if it is interpreted immediately after the preceding example and if we leave its modified transitive verb phrase unanalyzed for simplicity, it amounts to adding the condition 'always uses disposables(125)' to card #125.

3. In her 1983 paper Heim raises similar questions, and remarks that (p. 184) "(. . .) quantified and negated propositions are (. . .) puzzling if we are so ambitious as to want to say what exactly the file clerk does in response to them". For Heim this is less of a problem since: "(. . .) we have already resigned ourselves to characterizing file change only as far as the domain and satisfaction set are concerned", and, "[U]nder the modest aspect of domain and satisfaction set change, however, they [quantified and negated propositions, D. and H] pose no problem". It must be noted here that such a 'modest' position cannot be retained in DIP, because in this paper the entities that are the subject of update must be *files*, or representations, *not* their domain and satisfaction sets.

DYANA's first objective is 'to determine those aspects of the structure of information states which are relevant to natural language interpretation' (ANNEX (1992), p. 7). In keeping with this, DIP has it that:

> (...) a proper understanding of information packaging, i.e., of the actual strategies used by human agents in effecting information update by linguistic means, will help us gain further insight into the structural properties of the cognitive states these dynamic strategies manipulate. (DIP, §6)[4]

In fact, DIP argues that information packaging instructions contribute in two ways to the optimization of information update, since they provide means to

- designate a file card as the locus of information update and hence circumvent the redundancy of multiple update (§ 2.1); and
- identify the information of the sentence and its relation to information already present in the hearer's model (§ 3.1).

(Recall that the information of the sentence, I_s, is expressed by the focus, and that the ground has an ushering role with respect to I_s: links indicate where I_s goes, and tails indicate how it fits there.) So, as regards the information states exploited by information-packaging instructions, DIP argues that they have at least the internal structure of a system consisting of file cards connected by Hypercard-like links.

We feel that the presented arguments are intuitively quite appealing, but we still want to argue that, strictly speaking, they are not as compelling as they seem. Somehow, DIP seems to be begging the question: 'talking about ushering I_s to a location in the hearer's model K_h [...] does not make much sense unless one assumes some sort of rich internal structure for K_h' (§ 2). However, does this assumption of 'some sort of rich internal structure' make sense of anything besides the ushering function of links?

If Hypercard-style files are assumed, the instruction types that DIP distinguishes do appear to contribute to efficient information exchange. And if this assumption is warranted, it may even serve as an explanation of the fact that we do appear to find these ways of packaging information in a variety of languages. Nevertheless, the more theoretical question is whether this assumption itself is warranted, and whether the organization of linguistic information exchange really presupposes such information states. After all, ushers are very useful, but there are also halls that have unnumbered seats.

4. Cf. also: 'in discourse, speakers not only present information to their interlocutors, but also provide them with detailed 'instructions' on how to manipulate and integrate this information. The use of these instructions reveals that speakers treat information states as highly structured objects and exploit their structure to make information update more efficient for their hearers' (DIP, 'Abstract'), and: 'We agree with Heim that there has to be some additional internal structure in the hearer's model of the common ground that plays an important role in natural language interpretation, even if this internal structure is of tangential relevance in truth value computation. It is the internal structure of information states which is, in fact, crucially exploited by the different information-packaging strategies used by speakers in pursuing communicative efficiency' (DIP, §2).

Maybe links really make no sense without files, but, for that matter, maybe we simply fail to understand what links do. The notion of 'ushering I_s to a location' may be just as metaphorical as the notion of 'file card collection'.[5]

Let us now, for the sake of argument, imagine an utterance made by Irene, a speaker who organizes her utterances on the basis of the assumption that her audience uses Hypercard-style files to store information, to Hans, a hearer who in fact employs discourse representation structures. Clearly, it would be inappropiate to say that Irene uses links to usher I_s to a location in the hearer's model K_h, since there is no sensible notion of location in Hans's representation structures. Still, this does not at all preclude Hans from updating his discourse representation with the proposition that Irene attempts to get through. And worse, from a bookkeeping perspective he has got an easier job than a hearer who uses Hypercard-style files. For, consider the following link-focus example:

(3) $[_{\mathrm{L}}\textbf{Frank}_5][_{\mathrm{F}}$flew from Amsterdam$_9$ to Oslo$_8$ via STUTTGART$_2]$

Neglecting all details, if a file clerk is to update her file in order to represent the information expressed by example (3) in the way indicated in DIP, she has to carry out the following sequence of instructions:[6]

GOTO(5)(UPDATE-ADD(flew(8)(2)(9)(5)))
GOTO(8)(UPDATE-ADD($\boxed{\rightsquigarrow 5}$))
GOTO(2)(UPDATE-ADD($\boxed{\rightsquigarrow 5}$))
GOTO(9)(UPDATE-ADD($\boxed{\rightsquigarrow 5}$))
GOTO(5)

Hans, on the other hand, only has to carry out the following instruction:

UPDATE-ADD(flew(8)(2)(9)(5))

Notice that discourse representation structures can model precisely the same information as Hypercard files, except for one small difference. The only thing

---

5. Cf.: 'Our opinion, for whatever it is worth, is that the calculating mind is a metaphor rather than a model. It is a powerful metaphor, no doubt, on which many branches of 'cognitive' science are based, and sometimes it can be helpful, even insightful. But it remains a way of speaking, rather than a true description of the way we are' (Groenendijk and Stokhof (1991), p. 97). Our opinion, for whatever it is worth, is that one may substitute 'file card collection' for 'calculating mind' in this quote.

6. Assuming that establishing links to the locus of update is done via packaging instructions—of course, these links have to be established *some*how. Note, by the way, that the file clerk's task would not be made easier by structure sharing (something suggested by Enric Vallduví (p.c.)), since also the structure sharing will have to be established somehow—e.g., in the following way:

GOTO(5)(UPDATE-ADD($\boxed{1}$ flew(8)(2)(9)(5)))
GOTO(8)(UPDATE-ADD($\boxed{1}$))
GOTO(2)(UPDATE-ADD($\boxed{1}$))
GOTO(9)(UPDATE-ADD($\boxed{1}$))
GOTO(5)

that discourse representations lack is a marked discourse referent corresponding to the Hypercard file notion of current locus of update. If we assume that discourse representation structures have a marker for indicating such target discourse referents, then the two systems differ only in the way in which they display their information: in one big box, or on several cards.

The above may serve as an indication that none of the data discussed precludes the use of, say, discourse representation structures instead of Heimian Hypercard files. Clearly, there may be evidence for assuming there to be Hypercard files at work, and one of the last things we would like to claim is that people organize their information in simpler systems than Hypercard files (or discourse representation structures, for that matter). On the contrary. The only point is that the use of Hypercard files does not appear to be imperative so far.

In this section we want to raise some questions dealing with DIP's analysis of tail-containing sentences and the idea that information is organized in Hypercard files. Let us first summarize what tails are supposed to accomplish. The presence of a tail in a sentence is taken to signal a mode of update different from the straightforward UPDATE-ADD(I_s) instruction. A tail indicates that a (possibly underspecified) record on a file card has to be replaced (or specified further). Here the material in the tail serves the purpose of determining *which* record. Consider the following example, (16b) in DIP, which is presented as a reaction to the statement *Since John is dead, we can now split his inheritance*:

(4) I hate to spoil the fun, but [$_{\mathrm{L}}$**John**][$_{\mathrm{T}}$is][$_{\mathrm{F}}$NOT][$_{\mathrm{T}}$dead]

The idea is that the speaker, after directing the hearer to his file card for John (card #15, say), instructs this hearer to replace the record saying that John—the value of card #15—is dead by one saying that John is not dead. In short, the tail serves to highlight a condition on file card #15, the one saying its value is dead, which is then modified using the material in the focus.

Something similar is assumed to be going on in the following example. A newly appointed White House butler asks the presidential aid whether it was a good idea to buy an additional item for the president's china set. The presidential aid reacts:

(5) Nope. [$_{\mathrm{L}}$The **president**][$_{\mathrm{F}}$HATES][$_{\mathrm{T}}$the Delft china set]

The idea is that the butler has an underspecified record on his card for the president, which says that the president has some attitude towards the Delft china set. The lack of information about the nature of this attitude is reflected by a blank in the record, an it is this blank which, after hearing the presidential aid's reaction, is replaced by 'hates'.

Intuitively, the cases of UPDATE-REPLACE discussed in DIP are dealt with in a relatively satisfactory way. However, we expect complications when it comes to giving an explicit formalization of the instructions associated with tails. Any attempt to give an appropriate and fully general definition of these instructions will have to confront various questions. Thus, how exactly do you know which record has to be replaced or specified further? Is there guaranteed to be such a record? Is there a unique one, and what happens if there are more? Is it always one record that has to be replaced, or do we sometimes need to replace a group of records? What kind of match must there be between the material in a tail, and the material in the target record? We will now look at two more specific complications, stemming from the fact that the replacement of records takes place on file cards.

First, assume that someone told A that Kasparov beat Karpov in Linares, and that this person A has dutifully recorded this error by writing 'beat(1)(2)' on his file card #2 for Kasparov and '$\boxed{\leadsto 2}$' on his file card #1 for Karpov. Then A meets B and the following dialogue unfolds:

A: Karpov was beaten by Kasparov, so he may also be beaten by Timman
B: I hate to spoil the fun, but [$_{\text{L}}$**Karpov**][$_{\text{T}}$was][$_{\text{F}}$NOT][$_{\text{T}}$beaten by Kasparov]

B's reaction should lead A to correct the beat-record on card #2 by a non-beat-record. But a simple-minded UPDATE-REPLACE instruction would lead A to look for a record 'beaten_by(2)(1)' on file card #1 for Karpov, which need not exist.

Second, consider the following example, in which we see a project manager characteristically indulge in exotic conditionals:

(6) If Rintje Ritsma did not win, the Oslo workshop will be in Heerenveen

A Nordic site manager soberly replies:

(7) I hate to spoil the fun, but [$_{\text{F}}$Johan Olav KOSS][$_{\text{T}}$won] (of course)

Clearly, the site manager here instructs the project manager to replace his record according to which Rintje won[7] by one according to which Johan Olav did. But, presumably, this should not (only) be done on the card for Rintje.

The two examples show that replacement is complicated by the use of file cards. But in view of the tough nuts that have to be cracked when it comes to coming to theories of belief revision, the author of DIP has probably been wise in not too hastily engaging in the issues involved. Of course, they remain challenging questions which hopefully will be addressed in the future.

If, as we have argued, the use of Hypercard files does not appear to be imperative, then we face a question: what purpose *do* links serve if they do not serve to

7. We may safely assume that the project manager will eat his hat if Oslo and Heerenveen turn out to be one and the same town. By the way, exactly the same point is illustrated by a much more simple example:
Frank Veltman: Rintje Ritsma won
Jan Tore Lønning: No, [$_{\text{F}}$Johan Olav KOSS][$_{\text{T}}$won]

point at locations? What does 'ushering to a location' mean if representations do not come with locations? We would like to suggest the following tentative answer, which we take to carry less presuppositions than the file metaphor.

In all examples in DIP, links specify a card on which a condition gets written which crucially refers to that card. In link-less constructions, moreover, conditions are written on the current card which also refer to that very same card. The point is that these two types of actions also make sense in the framework of discourse representation theory. For, given the assumption that we have a marked discourse referent corresponding to the notion of the current locus of update (cf. the discussion following (3) above), we can also assume that a link-less construction identifies one of the arguments of the construction with the currently marked discourse referent, and that a construction with a link identifies the link argument with a *different* discourse referent and turns this one into the marked discourse referent. If this makes sense, then links serve to link terms with previously introduced discourse referents, i.e., they play a role in determining pronominal (co-)reference. Consider, by way of illustration, DIP's examples (10) and (11):

(10) $[_{\mathrm{L}}$**Sue**$]$$[_{\mathrm{F}}$insulted CATHY] and then $[_{\mathrm{F}}$she HIT her]
(11) $[_{\mathrm{L}}$**Sue**$]$$[_{\mathrm{F}}$insulted CATHY] and then $[_{\mathrm{L}}$**she**$]$$[_{\mathrm{F}}$HIT her]

The first conjunct in both examples contains the link *Sue*. In a Hypercard file this marks the card for Sue as the locus of update. In a DRS it would involve turning the discourse referent for Sue into the currently marked one. Then, in example (10), we find an all-focus continuation. The unstressed pronoun *she* is simply anchored to the currently marked discourse referent. Thus, the conjunction turns out to say that Sue hit Cathy (after insulting her). In example (11), on the other hand, the subject *she* is phonologically realized as a link, and this produces the 'obligatory switch-reference effect' along the lines we just sketched: the link argument *she* is identified with a discourse referent *different* from the currently marked one (and turns this one into the marked discourse referent). As a consequence, this example is interpreted as claiming that Cathy hit Sue (after being insulted by her). In short, the absence/presence of a link serves to associate/dissociate a pronominal subject with/from the marked discourse referent. This is precisely DIP's conclusion about these examples. The only difference is that links shift the locus of update in Hypercard files, whereas they shift the marking of discourse referents in discourse representation structures.[8]

Let us wind up this discussion. We have argued that the data discussed in DIP do not enforce the conclusion that information states have at least the structure of a system of file cards. The phenomena can also be accounted for in terms of discourse representation theory. This does require a slightly different perspective on the function of links which will have to await further scrutiny.

---

8. Vallduví (1992, p. 104) notes that there is no structural restriction on the number of links in Catalan. The possibility of having multiple links might yield an argument in favour of DRSs, since allowing multiple marked discourse referents seems more intuitive and technically less cumbersome than permitting multiple loci of update.

Again, we like to stress that our conclusion does not at all rule out the appropriateness of DIP's Hypercard files—we have only attempted to show that there are no compelling grounds for using them. Finally, we hope that these comments have not completely obscured the merits of DIP, a paper which we appreciate as a substantial contribution of broad interest.

ANNEX (1992). *Dynamic Interpretation of Natural Language (*DYANA*-2), Technical Annex.* ESPRIT Basic Research Project 6852. ILLC, University of Amsterdam.

Groenendijk, J., and M. Stokhof (1991). 'Dynamic Predicate Logic'. *Linguistics and Philosophy* **14**, pp. 39–100.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases.* Dissertation University of Massachusetts at Amherst.

Heim, I. (1983). 'File Change Semantics and the Familiarity Theory of Definiteness'. In R. Bäuerle, C. Schwarze and A. von Stechow (eds.) (1983), *Meaning, Use and Interpretation of Language*, pp. 250–269. De Gruyter, Berlin.

Kamp, H. (1981). 'A Theory of Truth and Semantic Representation'. In J. Groenendijk, T. Janssen and M. Stokhof (eds.) (1981), *Formal Methods in the Study of Language*, pp. 277–322. Mathematical Centre, Amsterdam. Reprinted in J. Groenendijk, T. Janssen and M. Stokhof (eds.) (1984), *Truth, Interpretation and Information. Selected Papers from the Third Amsterdam Colloquium*, pp. 1–41. Foris, Dordrecht.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic.* Kluwer, Dordrecht.

Vallduví, E. (1992). *The Informational Component.* Garland, New York.

Vallduví, E. (1994). 'The Dynamics of Information Packaging'. In E. Engdahl (ed.) (1994), *Integrating Information Structure into Constraint-based and Categorial Approaches.* ESPRIT Basic Research Project 6852, Dynamic Interpretation of Natural Language, DYANA-2 Deliverable R1.3.B. ILLC, University of Amsterdam.