Spatio-Temporal Indexing in Database Semantics

Roland Hausser

Universität Erlangen-Nürnberg
Abteilung Computerlinguistik (CLUE)
rrh@linguistik.uni-erlangen.de

Abstract In logic, the spatio-temporal location of a proposition is characterized precisely within a Cartesian system of space and time coordinates. This is suitable for characterizing the truth value of propositions relative to possible worlds, but not for modeling the spatio-temporal orientation of natural cognitive agents.[1] This paper presents an alternative approach to representing space and time. While on the same level of abstraction as the logical approach, it is designed for an analysis of spatio-temporal inferences in humans. Such an analysis is important for modeling natural language communication because spatio-temporal information is constantly coded into language by the speaker and decoded by the hearer. Starting from the spatio-temporal characterization of direct observation in cognitive agents without language, the speaker's coding of spatio-temporal information into language is analyzed, followed by the hearer's reconstruction of this location. These procedures of transferring spatio-temporal information from the speaker to the hearer are embedded in the general structure of database semantics.

Overview

The goal of database semantics is a computational model of natural qcommunication. The transfer of information from the speaker to the hearer is based on representing the knowledge of speaker and hearer, respectively, in the form of databases. Natural language communication is successful if a certain database content, coded by the speaker into natural language signs, is reconstructed analogously in the database of the hearer.

A special aspect of analogous reconstruction is the coding and transfer of spatio-temporal information. This aspect is integrated deeply into the system of database semantics, for which reason the general structural principles for transferring propositional content are described first (Section 1–3).

Then, the spatio-temporal indexing of direct observations in cognitive agents without language is analyzed (Section 4). In this analysis, the cognitive representation of time is based on the order in which propositions enter the cognitive agent's database. The chain of incoming propositions is structured naturally by observations of cyclical events like

[1] The analysis of time in tensed predicate calculus (TPL) is critically reviewed in Kamp & Reyle 1993, Vol. 2, p. 483–510. Their own approach, called discourse representation theory (DRT), however, has "no conclusive answers to these questions, and presumably there are no clear, non-stipulative answers to be had" (op. cit. p. 666). From the viewpoint of database semantics, this is not surprising because DRT is itself a variant of model-theoretic semantics.

day and night. Such observations serve as landmarks which provide the basis for spatio-temporal reasoning.

Finally, the transfer of spatio-temporal information from the speaker to the hearer is addressed. This raises two questions: (i) How can the spatio-temporal location of a propositional content in the speaker's database be inferred by the hearer (Section 5) and (ii) how should this location, once it is inferred, be represented in the database of the hearer (Section 6)?

It is shown that the answers depend on the distinction between fleeting and permanent signs and between immediate and mediated reference. These distinctions result in four basic types of natural language communication (Section 7).

1 Intuitive outline of database semantics

In theoretical linguistics, tree structures are assigned to individual sentences in accordance with certain intuitions (constituent structure). Roughly similar intuitions are also used in database semantics, specifically with regard to the functor-argument structure of simple sentences and the composition of complex sentences from simple ones.

However, in database semantics the representation of linguists' intuitions by means of trees turned out to be an unsurmountable obstacle for realizing a functional theory of natural language communication.Instead, the grammatical structures needed for the transfer of information from the speaker to the hearer must be coded in a distributed manner into an abstract data structure.

This data structure may be explained intuitively by the following analogy:

> Consider two girls who store their books in different houses. They are avid readers with an ample budget and continuously add new books to their libraries. Furthermore, they each always finish their current book before they start reading the next one.
> To keep track of their reading, they have the habit of writing the author and title of the current book, the book read previously, and the book to be read next on a card and putting it into the current book. Once in a while, they also note the date on the card, thus anchoring the sequence to a temporal landmark. The books with the cards are stored in their libraries in the alphabetical order of the authors' names.
> When one of the girls goes on holiday, she keeps up her card-making habit. Furthermore, each time she goes to a new place, she notes it on the card of the current book, thus anchoring the sequence to a spatial landmark. After reading an interesting sequence of books, she sends copies of her cards to her friend. Her friend obtains the books, puts the copied cards into them, and sorts the books alphabetically into her library – keeping the first book in the sequence for immediate reading. When it is finished, she uses its card to go to the next book, etc.

The communication of reading sequences between the two girls is based on (i) structure-based storage and retrieval, (ii) internal chaining, and (iii) landmark anchoring. These principles may be described as follows:

1.1 Structural principles for solving the update task

1. **Structure-based storage and retrieval** is used in well-organized private libraries. The books in the shelves are ordered alphabetically according to the author's name.

Several books by the same author are ordered in their sequence of acquisition. In this way, the storage position of new books is determined uniquely, while the retrieval position of old books is sufficiently narrow.

2. **Internal chaining** is a method for establishing relations between books based on keeping notes in them. Each note describes the book in which it is contained and its relations to other books. By specifying local relations between books A and B, B and C, C and D, etc., arbitrarily long chains may result. By specifying both the preceding and the following book, these chains may be followed forward as well as backward.
3. **Landmark anchoring** is based on (2) internal chaining and makes it possible to connect several items to the same landmark by marking only the initial item in the sequence. For example, if an item in an internal chain is connected to landmark X, e.g. a certain place, then all subsequent unmarked items are assumed to be connected to landmark X as well. As soon as an item in the chain is marked for another landmark Y, however, the following items are connected to Y, etc.

Structure-based storage and retrieval has the advantage that no index has to be constructed and maintained. This is in contrast to catalog-based indexing, as used in public libraries.[2] Database semantics solves the update task of natural language communication without the use of a catalog-based index in order to be compatible with the cognitive evolution of natural agents.

2 Structure-based storage and retrieval

The copies of the cards that are passed between the girls stand for the content words (surfaces), while the corresponding books stand for their meanings. The relation between words and their meaning is established by the lexicon. The meanings in the database are structurally connected into *concatenated propositions*.

Technically, a proposition is represented as a set of feature structures, called proplets. As feature structures, proplets are special in that they use internal chaining to express intra- and extrapropositional relations between proplets.

The intrapropositional relations are expressed as follows: the proplet of an argument specifies the associated functor, the proplet of a functor specifies the associated argument(s), and similarly for modifiers. Proplets belonging to the same proposition share a common proposition number.

The extrapropositional relations are of two types, conjunction and identity. The proplets of functors contain the attribute cnj which specifies the functor of another proposition and the conjunctional relation between them, for example temporal order expressed

[2] Such a catalog consists of cards, each describing (i) a book in terms of its author's name and its title, and (ii) the physical storage location of the book in a certain shelf. While the filing cards are ordered alphabetically according to their respective keywords, the choice of the storage locations is free. Once a given book has been assigned a certain location and this location has been noted in the catalog, however, its address is fixed. In this way, the catalog serves as an index which guides the retrieval of the item desired from a specified storage location.

Another example of catalog-based indexing is a relational database, where several indices, called tables, are used in combination.

by ‘>’. The proplets of arguments contain the attribute id which specifies the identity or non-identity with other arguments.

For example, the concatenated propositions *Peter leave house. Then Peter cross street.* are represented in database semantics by the following proplets:

2.1 Representation of two concatenated propositions

$$
\begin{bmatrix} \text{arg: Peter} \\ \text{FUNC: leave} \\ \text{id: 1} \\ \text{prn: 54} \end{bmatrix}
\begin{bmatrix} \text{func: leave} \\ \text{ARG: Peter house} \\ \text{cnj: > cross}_{55} \\ \text{prn: 54} \end{bmatrix}
\begin{bmatrix} \text{arg: house} \\ \text{FUNC: leave} \\ \text{id: 2} \\ \text{prn: 54} \end{bmatrix}
$$

$$
\begin{bmatrix} \text{arg: Peter} \\ \text{FUNC: cross} \\ \text{id: 1} \\ \text{prn: 55} \end{bmatrix}
\begin{bmatrix} \text{func: cross} \\ \text{ARG: Peter street} \\ \text{cnj: leave}_{54} \text{ >} \\ \text{prn: 55} \end{bmatrix}
\begin{bmatrix} \text{arg: street} \\ \text{FUNC: cross} \\ \text{id: 3} \\ \text{prn: 55} \end{bmatrix}
$$

Each proplet is an autonomous, self-contained item which specifies relevant relations to other proplets by means of attributes.[3] The proplets in 2.1 constitute an unordered set.

This coding depends neither on the graphical means of tree structures, as used in theoretical linguistics, nor on the linear ordering of symbols, as used in logic. Therefore proplets can be stored in accordance with the requirements of any type of database.

For structure-based storage and retrieval, each concept type (corresponding to an author’s name) is listed in an alphabetically ordered column. Different proplets with the same concept are sorted in a row behind their concept type. Such a row, called a token line, corresponds intuitively to a shelf with books by the same author.

This is the data structure of a record-based classical network database. It defines a 1:n relation between two kinds of records, the owner records and the member records. The following example shows the sorting of proplets into a network database, whereby the types are the owner records and the proplets are the member records.

2.2 Sorting proplets into a network database

TYPES   PROPLETS

$$
\begin{bmatrix} \text{concept: cross} \\ \text{role: functor} \end{bmatrix}
\quad \ldots \quad
\begin{bmatrix} \text{func: cross} \\ \text{ARG: Peter street} \\ \text{cnj: leave}_{54} \text{ >} \\ \text{prn: 55} \end{bmatrix}
$$

$$
\begin{bmatrix} \text{concept: house} \\ \text{role: argument} \end{bmatrix}
\quad \ldots \quad
\begin{bmatrix} \text{arg: house} \\ \text{FUNC: leave} \\ \text{id: 2} \\ \text{prn: 54} \end{bmatrix}
$$

---

[3] The first specifies the proplet concept, characterizing it as a functor or an argument. The second (in upper case), called continuation predicate, specifies the intrapropositional relation(s). The attributes cnj, id, and prn stand for conjunction, identity and proposition number, respectively.

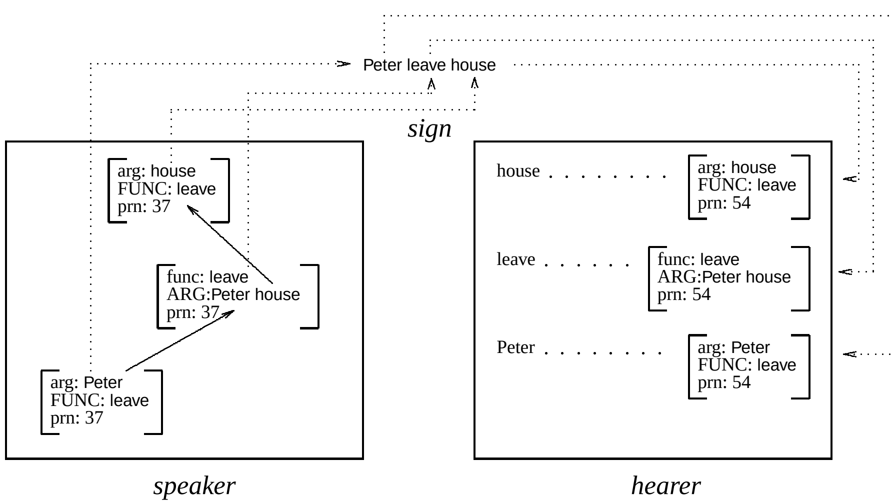
[concept: leave role: functor]	…	[func: leave ARG:Peter house cnj: > cross$_{55}$ prn: 54]	
[concept: Peter role: argument]	…	[arg: Peter FUNC: leave id: 1 prn: 54]	[arg: Peter FUNC: cross id: 1 prn: 55]
[concept: street role: argument]	…	[arg: street FUNC: cross id: 3 prn: 55]	

The structure of a network database supports the filing of proplets in their proper positions and their retrieval without the need of building and maintaining a catalog. Compared to structure-based storage and retrieval in a private library, a network database has the advantage of 'flexible shelves,' i.e., a token line may become arbitrarily long.

A network database for proplets is called a *word bank*. Compared to a standard network database, a word bank is special in the following three respects: First, intrapropositional functor-argument structures and extrapropositional relations between proplets are specified by means of continuation predicates (internal chaining). Second, these continuation predicates support an autonomous navigation through the propositional content, whereby the navigation is powered by a motor algorithm. Third, word banks are embedded in cognitive agents.

A navigation through the speaker's word bank and the appropriate filing of proplets in the hearer's word bank may be illustrated as follows:

2.3 Transfer of information from the speaker to the hearer



The speaker's navigation starts with the proplet *Peter*, which specifies the continuation proplet *leave*. From *leave*, the navigation continues to the proplet *house*. The navigation serves as the conceptualization and basic serialization of language production. The content of the database is coded into language simply by uttering the surfaces of the proplets traversed (cf. *sign*).

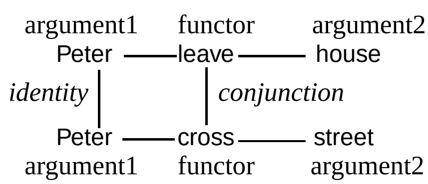
The hearer's interpretation consists in deriving a corresponding set of proplets. This procedure is based on (i) assigning lexical proplets to the sign's surfaces and (ii) copying proplet names into appropriate slots of other proplets during the syntactic-semantic analysis (reconstruction of the internal chaining).

In this way, the word meanings, the functor-argument structure, and the extrapropositional relations of the speaker's content are reconstructed by the hearer. Furthermore, the time-linear order of the sign induced by the speaker's navigation is eliminated, allowing storage of the proplets in accordance with the principles of the data structure in question. Apart from the hearer's prn, the interpretation is exactly analogous to the original structure in the speaker's database.

3 Alternative realizations of propositional content

The representation of content in a word bank is of a universal, language-independent nature. All content is represented in the form of elementary propositions which may be concatenated in only two ways. One is the concatenation of functors via conjunctions. The other is the concatenation of arguments via identity. Using graphical means, this basic structure may be illustrated as follows:

3.1 'RAILROAD SYSTEM' PROVIDED BY TWO PROPOSITIONS



The two propositions in 3.1 correspond to those in 2.1 and 2.2. Language production, based on a navigation through such a universal representation of content, has to take care of the following properties of natural language:

First, there are the language-dependent properties of lexicalization, word order, inflection, and the realization of function words. For example, the underlying navigation *Peter leave house* (cf. 2.3) must be supplied with language-specific surfaces,[4] possibly modified to a different language-dependent word order, provided with inflectional forms (here leave+s), and supplied with function words (here the). The result for English would be Peter leaves the house.

Second, there is the distinction between forward and backward navigation through a given propositional content. For example, an intrapropositional forward navigation through the first proposition is realized as Peter leaves the house, while a corresponding backward navigation is realized as The house is left by Peter. Similarly, a forward

[4] The proplet names in 3.1 as well as 2.1 – 2.3 are really language-independent.

navigation through the two propositions illustrated in 2.1, 2.2, and 3.1 is realized as Peter leaves the house. Then Peter crosses the street., while a corresponding backward navigation is realized as Peter crosses the street. Before that Peter leaves the house.

Third, there is hypotactic (embedding) navigation – in addition to the paratactic (conjoining) navigation illustrated above – from one proposition to the next. A hypotactic navigation based on a conjunction results in an adverbial clause, as in Peter, after he leaves the house, crosses the street. or Peter, before crossing the street, leaves the house. A hypotactic navigation based on identity, on the other hand, results in a relative clause, as in Peter, who leaves the house, crosses the street.

Fourth, there is repeated reference – when a navigation returns to a referent previously traversed. This may be reflected in some languages by using an appropriate pronoun, as in Peter leaves the house. Then he crosses the street., and in others by omitting the repeated referent completely.

Fifth, there are abbreviating navigation types through propositional content. One is reduced coordination as in Peter leaves the house and crosses the street. Another is ellipsis and telegraphic style, as in Leaves! Crosses!

Six, there is metaphor, as when the speaker refers to Peter as the old fox.

In successful communication, these different ways of the speaker's navigating through a propositional content and realizing the navigation in a given natural language must all be reversed back to the original propositional content by the hearer. For example, proper understanding requires that the old fox is identified with Peter, that Leaves! is provided with the arguments intended by the speaker, that a hypotactic navigation must be disentangled, etc. The more completely the language sign is reversed by the hearer into the original language-independent content, the better the understanding.

The different possibilities of coding a propositional content are utilized by the speaker to aid the hearer's decoding. When the hearer changes into a speaker, he or she must in turn select one of the possible types of formulation to maximize the transfer of the content in question. This choice is made relative to the current rhetorical goal and the current utterance situation. Sometimes the proper interpretation has to be negotiated between the speaker and the hearer in a dialog.

4 Spatio-temporal indexing of direct observation

The description of database semantics so far has concentrated on the transfer of propositional content without taking its spatio-temporal location into account. Turning now to the analysis of the latter, we begin with direct observations.

Direct observation is a cognitive agent's keeping track of recognitions and actions in the task environment. Direct observation is one of two methods of reading propositional content into a word bank, the other being based on media such as language.

In terms of evolution, the reading-in of propositional content by means of direct observation is prior to the reading-in by means of language. Furthermore, communicating direct observation is one of the basic uses of natural language.

The modeling of direct observation in an artificial cognitive agent requires that static propositional content be related to a continuously evolving present moment and location.[5] For example, *fridge contain cold can of beer* is a static fact stored in memory, but its relevance for the cognitive agent depends on its temporal (is it there now?) and spatial (can I get to it?) relation to the present.

The temporal indexing of direct observations requires no more than the system's keeping track of the sequencing of incoming observations.[6] This is expressed by the proposition numbers, which are used as a parameter for temporal order (in addition to expressing which proplets belong to the same proposition). Pauses between two subsequent observations may be expressed by propositions characterizing the pause.

Based on internal chaining, the system can navigate forward, e.g., 54, 55, 56, etc., and backward, e.g. 54, 53, 52, etc., through the propositions with these numbers. This forward and backward navigation is the basis of temporal inferencing. As an example consider the two propositions *one can of beer put in fridge* and *one can of beer consumed*. Depending on their order in the database, they render different conclusions as to whether or not there is currently one can of beer available.

The sequencing of propositional content is structured by cyclical external events stored in memory, such as the change of day and night, the phases of the moon, the changing seasons, the daily round of the milkman, etc. These cyclical events are used as temporal landmarks and serve as the basis for further inferences.

For example, if a content's proposition number, e.g. 2054, is positioned between two propositions with the content *It is night* and the proposition numbers, e.g. 2045 and 2075, respectively, and no other propositions of this content are intervening between 2045 and the present, then the direct observation expressed by proposition 2054 occurred yesterday. After one more day, there will be an additional night's proposition number, e.g. 2098, between the content in question and the present moment, leading to the conclusion that the observation now occurred the day before yesterday, etc.[7]

Cyclical events are also used for structuring the future, as when telling a child: *You have to sleep three more nights and then it will be your birthday.* Applying temporal inferencing learned from analyzing the past, the child can extrapolate into the future.

What has been said about temporal indexing and inference applies mutatis mutandis to spatial inference as well. The current location of the cognitive agent is indicated by the observation of appropriate landmarks. These observations appear in the sequence of propositions and serve as the basis of spatial inferences.

For example, if a sequence of propositions contains the observations *Going from kitchen to living room* and *Going from living room to kitchen*, then all direct observations between them are located in the living room. Thus spatial inference is analyzed

[5] In this sense, our approach corresponds to McTaggart's B-series. See P. Ludlow 1999.

[6] This is in contrast to postulating events as basic entities of one's semantic ontology. For example, in their attempt to avoid the problems of TPL, Kamp & Reyle 1993 introduce events as basic entities, but admit that events treated in this way are problematic as well: "Their identity criteria are ill-defined; it is difficult to know under which conditions there exists, for two or more distinct events, a single new 'sum' event which has those events for its parts; and it is not clear how events are related to other entities, such as, for instance, times." (op. cit. p. 666). The crucial role of ontology for semantic interpretation is shown in Hausser 2001.

[7] Another inference applies to cooccurrence. For example, if events A and B both have proposition numbers which are lower than last night and higher than the night before, then A and B must have occurred on the same day, namely yesterday.

just like temporal inference on the basis of landmarks which are accessed by navigating forward or backward through the sequence of propositions. The similarity between temporal and spatial reasoning is in concord with the fact 'that prepositions of time are on the whole identical to spatial expressions and that temporal prepositional phrases are attached to sentences in the same way as prepositional phrases of location.'[8]

In modern society, temporal inference is supported by an external system of clocks and calendars. A realistic model of human cognition should not postulate an absolute internal clock, however, which would provide functor proplets with a time stamp, e.g. [Thursday, Oct. 19 2000, 9:45:24]. This is shown by situations in which the usual cyclical events are missing. For example, being locked up in an empty cell without a watch and without access to daylight causes a painful loss of temporal orientation in humans.

Similarly, spatial inference is supported by an external system of signs and road maps characterizing the relations and distances between places. However, just as temporal inferencing can do without an absolute internal clock, spatial inferencing can do without an absolute positioning system which constantly updates the cognitive agent's location on an internal map representing the external environment.

5 Reconstructing spatio-temporal location

In natural language communication, the spatio-temporal location of the speaker's propositional content must be reconstructed by the hearer. This reconstruction is based on three potentially distinct spatio-temporal locations. They are

1. the ST_{prop}, i.e. the spatio-temporal location of the propositional content,[9]
2. the STAR point, i.e. the parameters of the sign's origin in terms of Space, Time, Author, and intended Recipient,[10] and
3. the ST_{inter}, i.e. the spatio-temporal location of the interpretation.

The hearer's analogous reconstruction requires (i) inferring a content's ST_{prop} in the speaker's word bank and (ii) expressing this ST_{prop} in the hearer's word bank.

Inferring a content's ST_{prop} consists of two parts. First, the hearer must proceed from the ST_{inter} to the sign's STAR point. Second, the hearer must interpret the sign by decoding the sign's content and reconstructing its ST_{prop} relative to the STAR point.

[8] Jackendoff 1983, p. 189.

[9] The T-value of ST_{prop} corresponds to the 'point of the event' by Reichenbach 1947, p. 288.

[10] For characterizing the origin of a sign's token, spatio-temporal indexing is necessary, but not sufficient. In addition, the author and the intended recipient of the sign must be defined. The author may be distinct from the speaker, as when a letter from far away is being read aloud to the whole family. The intended recipient may be distinct from the hearer, as when a person is forced to overhear other people's conversation on a train.

The STAR point is an extension of the 'point of speech' by Reichenbach 1947, p. 288, which refers to the temporal parameter T only and is terminologically restricted to spoken language. The STAR point was first described in Hausser 1989, p. 274 f. See also Hausser 1999, p. 93 f.

In line with current practice in the field, we will continue to use 'speaker' in the sense of author and 'hearer' in the sense of intended recipient or unintended interpreter, unless clarity requires the correct terms. The current, medium-specific terminology may be understood charitably as emphasizing the primacy of the spoken over the written medium.

For example, a speaker observes on Monday *Peter leave house* (T-value of ST_{prop} = Monday) and expresses this fact on Tuesday (T-value of STAR point = Tuesday) by writing Yesterday Peter left the house on a piece of paper. A hearer reads the sign on Wednesday (T-value of ST_{inter} = Wednesday).

The communication is successful if the hearer's word bank stores the sign's content with an ST_{prop} relating to Monday. For this, the hearer must reconstruct the ST_{prop} of the sign's content (i) by inferring the sign's STAR point and (ii) by interpreting the sign's tense and temporal adverb relative to this STAR point.

While the speaker's ST_{prop} of a sign's content and the STAR point of a sign's origin are defined once and for all, there may be many possible ST_{inter} of a sign's interpretation. For example, the above note might be read again by another hearer on Thursday (T-value of ST_{inter} = Thursday). The two different interpretations of the sign are successful only if they use the same STAR point.

The relation between the hearer's ST_{inter} and the sign's STAR point is characterized by the distinction between fleeting and permanent signs.[11]

5.1 Fleeting and permanent signs

> [+fs] is the use of fleeting signs, which requires the hearer to be present[12] during the sign's production. This presence causes the spatio-temporal location of the interpretation point to equal the STAR point,
> formally written as [+fs]: STAR point = ST_{inter}.
>
> [-fs] is the use of permanent signs, which allows the hearer to be absent during the sign's production. This absence causes the spatio-temporal location of the interpretation point to differ from the STAR point,
> formally written as [–fs]: STAR point $\neq$ ST_{inter}.

When the hearer has no direct access to the situation in which a permanent sign originated ([–fs]), the STAR point must be explicitly coded into the sign's type in order to ensure successful communication.

For example, a standard letter specifies the values of S and T in the dateline, of A in the signature, and of R in the greeting. If different copies are made of such a letter, either by writing or by a copying machine, these copies are additional tokens of the letter's type. Each of these tokens may be interpreted correctly because the STAR point has been explicitly coded into the associated type.

The relation between the speaker's ST_{prop} and the sign's STAR point is characterized by the distinction between *immediate reference* and *mediated reference*,[13] formally expressed by the binary feature [$\pm$ ir].

[11] Permanent signs are used in written language. Recordings of fleeting signs are permanent signs as well. Therefore, recordings of language usually indicate the STAR point explicitly, either by stating place, date, author, and intended recipient at the beginning of the recording, or by specifying this information on the cassette.

[12] An exception created by modern technology is telephone conversations. Their signs are fleeting, yet STAR point and interpretation point differ in their S values, necessitating that spatial information be interpreted accordingly.

[13] See also Hausser 1999, p. 73 f.

[+ir] is the speaker's reference to objects in the current task environment, which causes the spatio-temporal location of the propositional content to equal the sign's STAR point,
formally written as [+ir]: ST_{prop} = STAR point.

[–ir] is the speaker's reference to objects not in the current task environment,[14] which causes the spatio-temporal location of the propositional content to differ from the sign's STAR point,
formally written as [–ir]: $ST_{prop} \neq$ STAR point.

The pivot of natural language communication is a sign's STAR point. It stands between the speaker's ST_{prop} and the hearer's ST_{inter}. The speaker characterizes the contents ST_{prop} relative to the sign's STAR point and the hearer reconstructs the contents ST_{prop} relative to it. The sign's type and the STAR point of the sign's token jointly determine the ST_{prop} of the sign's content for arbitrary ST_{inter}.

6 Expressing spatio-temporal location

The next question is how to express a content's ST_{prop} in the hearer's word bank once it has been inferred. More specifically, given the hearer's spatio-temporal system based on proplets stemming from direct observation, where should the proplets stemming from language content be stored?

Proplets stemming from direct observation have a two-fold temporal characterization. One is by means of internal chaining which relates propositional content to temporal landmarks (cf. Section 4). The other is by means of the proplets' order in their token lines, which reflects the temporal sequence of their arrival.

Consider a proposition P which is stored in the speaker's word bank as a direct observation, but in the hearer's word bank as a language-based content. Does analogous reconstruction require that the storage of P in the hearer's word bank reflect temporal order in exactly the same way as the storage of P in the speaker's word bank? In other words, should proplets stemming from language-based content be sorted into the hearer's token lines to express their temporal location in terms of the token-line-ordering? This would have several disadvantages.

First, it would be difficult to find the position in the token line which would correctly express the temporal location of the language proplet in question. Second, the content of a word bank would have to be constantly restructured to accommodate the temporal interpretation of incoming language proplets. Third, this method cannot be extended to express the spatial location of language-based content.

---

[14] For example, J.S. Bach (1685–1750) as as a database referent has no counterpart in the current real world. The same is true when a speaker talks about future events.

In mediated reference, the ST_{prop} is usually indicated by a suitable spatio-temporal landmark, e.g. a specific date and place. Such landmarks serve as the 'point of reference' in the sense of Reichenbach 1947, p. 288, providing a fixed point for the time structures coded into the language sign.

Therefore, language proplets are sorted in at the end of the token lines in their order of arrival, just like proplets stemming from direct observation. In other words, the ordering of language proplets in a token line characterizes the temporal location of their ST_{inter}, and not of their STAR point or their ST_{prop}. The spatio-temporal location of language-based content is expressed solely by means of internal chaining, which relates the propositional content to suitable landmarks.

For example, based on direct observation the hearer's word bank contains the proposition *It's Monday* with the proposition number, e.g., 23. On Wednesday, the hearer interprets the sign Yesterday Peter left the house, deriving the content *Peter leave house* with the proposition number, e.g., 54. Based on the ST_{inter} and the STAR point, the hearer determines that the ST_{prop} of proposition 54 has Monday as its T-value.

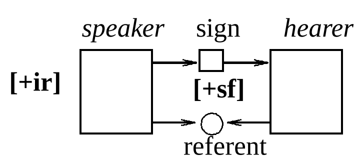
To express this ST_{prop} in the hearer's word bank, proposition 54 is related to the temporal landmark constituted by proposition 23. This is done by extending the leave-proplet of proposition 54 to the attribute cnj: [54 $=_{time}$ 23], meaning that propositions 54 and 23 occurred at the same time. The spatial location of a language-based propositional content is expressed in a similar manner by anchoring it to a spatial landmark.

7 Four basic types of natural language communication

Because the [± ir] distinction characterizes the relation between the speaker's ST_{prop} and the sign's STAR point, and the [± fs] distinction characterizes the relation between the sign's STAR point and the hearer's ST_{inter}, there result four basic types of natural language communication. These may be characterized as follows:

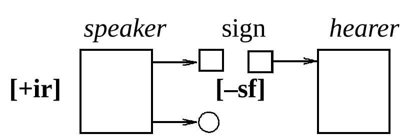
7.1 Basic relations between ST_{prop} and ST_{inter}

(i): [+ir, +fs]: ST_{prop} = STAR point = ST_{inter}



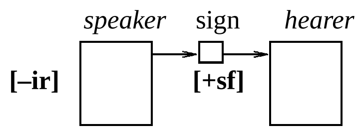
Speaker and hearer are together and talk about objects in their common task environment. This allows the use of fleeting signs and ensures that the ST values of the propositional content agree with the sign's STAR point and the hearer's ST_{inter}. Consequently, the language-based information may be stored by the hearer just like direct observation, without any need to establish spatio-temporal relations to stored landmarks.

(ii): [+ir, –fs]: ST_{prop} = STAR point $\neq$ ST_{inter}



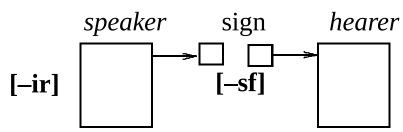
The ST values of the propositional content agree with the sign's STAR point, as in (i). However, the situations of utterance and interpretation are distinct, requiring the use of permanent signs. Therefore, the hearer must determine the STAR point and relate the propositional content to landmarks corresponding to the STAR point.

(iii): [–ir, +fs]: $ST_{prop} \neq$ STAR point = ST_{inter}



Speaker and hearer are present in the same situation, allowing the use of fleeting signs. However, because of mediated reference the ST values of the propositional content do not agree with the sign's STAR point. Therefore, the hearer must determine the spatio-temporal relation of the content in the speaker's database to the STAR point, and relate the content to suitable landmarks in his or her own database.

(iv): [–ir, –fs]: $ST_{prop} \neq$ STAR point $\neq ST_{inter}$



The situations of utterance and interpretation are distinct, requiring the use of permanent signs. Furthermore, the ST values of the propositional content do not agree with the sign's STAR point. Therefore, the hearer must determine both the STAR point and the relation of the content to the STAR point. This jointly determines the spatio-temporal location of the content for the hearer.

The four types constitute only the most basic distinctions. They are based solely on whether the ST values of propositional content, STAR point, and interpretation point are equal or not. There are many kinds of nonequality, however.

One set of additional distinctions arises from nonequality between the ST values of the propositional content and the STAR point ($ST_{prop} \neq$ STAR point). The nonequality may consist in reference to events in recent past, distant past, near future, distant future, to objects which are spatially near, spatially distant, above, below, in front of or behind the speaker or a certain landmark,[15] to structures outside of time and space, like mathematical laws, etc. All these distinctions are reflected in natural language and must be decoded by the hearer with reference to the STAR point.

Another set of additional distinctions arises from non-equality between the ST values of the interpretation and the STAR point (STAR point $\neq ST_{inter}$). This requires

[15] For a summary of ego-centered, object-centered, and environment-centered frames of spatial description see MacWhinney 1999, p. 221 f.

the hearer to determine the STAR point, which may be specified explicitly by reference to external landmarks, like date and place in a letter, or left implicit, necessitating inference. In the case of the latter, the question arises as to how far the ST values of interpretation and STAR point are apart, how precisely the STAR point has to be inferred for proper understanding, etc.

8 Spatio-temporal questions and answers

Corresponding to the distinction between fleeting signs ([+fs]) and permanent signs ([–fs]) there is the distinction between dialog and text. Thereby dialog is type i and iii, while text, including books, letters, notes, emails, etc., is of type ii and iv.

In text, the relation between speaker and hearer is STAR point $\neq ST_{inter}$. Therefore, reactions of the hearer (reader) are either not fed back to the speaker (author) at all, or take a considerable amount of time and effort, as in a reply letter. Furthermore, common ground cannot be negotiated between the speaker and the hearer. Instead, it is up to the hearer to infer the intention of the speaker on his or her own.

In dialog, the relation between speaker and hearer is STAR point = ST_{inter}. Therefore, the intended recipient as well as third parties present in the utterance situation[16] can react immediately by taking over the role of speaker (turn taking), including the possibility of negotiating the common ground between the parties involved. This is particularly relevant for type iii communication, where $ST_{prop} \neq$ STAR point.

Even if the possibility of immediate reaction may seem to favor questions, answers, and commands in dialog, while the absence of this possibility may seem to favor statements in text, these different uses of language all occur in both dialog and text. For example, a dialog may contain statements – just as a text, e.g. a letter, may contain questions, answers, and commands. Apart from the distinction between STAR point $\neq ST_{inter}$ and STAR point = ST_{inter}, the production and interpretation of statements, questions, answers, and commands function the same in texts and dialogs.

As an example, consider a dialog in which a user asks the word bank 2.2 When did Peter cross the street? Because STAR point = ST_{inter}, problems of narrowing the relevant temporal interval do not arise. In order to respond, the system can go immediately through the following three steps:

First, the system's syntactic-semantic interpretation assigns to the sign a 'query proplet' as its semantic representation:

8.1 QUERY PROPLET REPRESENTING When did Peter cross the street?

$$\begin{bmatrix} \text{func: cross} \\ \text{ARG: Peter street} \\ \text{cnj: } ?_{time} \\ \text{prn: } ? \end{bmatrix}$$

[16] A special case are telephone conversations, where third parties may take part in the utterance situation, in the interpretation situation, or in both.

A query proplet contains variables (here characterized by '?') which turn it into a pattern. The pattern is used to find suitable variable values in the word bank.

Second, the system uses the query proplet pattern to search the associated token line, here that of cross. Beginning at the end, the pattern is tried on one proplet after the next until it finds the first (and here only) one matching the arguments Peter and street. The matching proplet binds the variable of the prn to 55.

Third, starting from proposition 55, the system navigates either forward, e.g. to 56, 57, 58, etc., or backward, e.g. 54, 53, 52, etc. Given that the proposition numbers of proplets stemming from direct observation reflect temporal order, any one of them would serve as a suitable basis for an answer.

In our example, the preceding proposition is *leave Peter house* with the prn 54. Because 54 < 55, the system answers After Peter left the house.

This is the kind of vague answer one might expect from a human. As a reaction, the user could ask When was it exactly? or What time was it? A word bank with more propositional content than 2.2 might try a more precise response by looking for a temporal landmark like It was before I looked at my watch and saw that it was 10am.

Searching through propositions in order to adequately answer adverbial wh-questions requires that the system can recognize temporal landmarks in the case of temporal questions, spatial landmarks in the case of spatial questions, causal landmarks in the case of causal questions, etc. However, given proper lexical analysis, this is not a difficult problem, especially when inferences are supplied in addition.

Technically, the navigation through the token line and subsequently through the propositional chain is handled like all navigation in database semantics, namely by suitable LA-grammars. The detailed formal treatment of the navigations realizing the search initiated by when- and where-questions is similar to that of yes/no- and who-questions presented in Hausser 1999, p. 491f.

9 Conclusion

In our approach, the spatio-temporal location of propositional content is specified without the usual reference to a precise, objective, external Cartesian coordinate system. Instead, spatio-temporal location is characterized cognitively by the order of direct observations entering the database of a cognitive agent. The sequence of propositions is structured by observations of cyclical events, serving as temporal landmarks, and by observations of the environment, serving as spatial landmarks.

This analysis works for cognitive agents with and without language. Modeling the former, however, poses the additional task of communicating the spatio-temporal location of propositional content (in addition to the propositional content itself) from the speaker to the hearer. For this, a propositional content's spatio-temporal location is characterized in the database of the speaker, (ii) coded into the language sign, (iii) reconstructed by hearer, and (iv) and represented in the hearer's database.

Spatio-temporal inferences are handled like all inferences in database semantics, namely on the basis of navigating through the concatenated propositions. The purpose of navigating forward or backward through the sequence of propositions is finding a suitable spatio-temporal landmark. The result of these inferences characterizing the spatio-temporal location of content is shown to be similar to that of humans.

In artificial agents, spatio-temporal localization can be made more powerful by using technical devices such as on-board clocks, geo-stationary satellite positioning systems, indexing and retrieval tools of databases, etc. Instead of replacing the ‘soft’ system, however, technology should only complement it in a way that maintains the basis for a natural understanding between artificial and human cognitive agents.

References

Hausser, R. (1989) *Computation of Language, An Essay on Syntax, Semantics and Pragmatics in Natural Man-Machine Communication*, Springer-Verlag, Symbolic Computation: Artificial Intelligence, Berlin, New York.

Hausser, R. (1999) *Foundations of Computational Linguistics*, Springer-Verlag, Berlin, New York.

Hausser, R. (2001) ‘The Four Basic Ontologies of Semantic Interpretation,’ in H. Kangassalu et al. (eds.) *Information Modeling and Knowledge Bases XII*, IOS Press Ohmsha, Amsterdam.

Jackendoff, R. (1983) *Semantics and Cognition*, The MIT Press, Cambridge, Mass.

Kamp, H. & U. Reyle (1993) *From Discourse to Logic*, Part 2, Kluwer Academic Publishers, Dordrecht, Holland.

Ludlow, P. (1999) *Semantics, Tense, and Time: An Essay in the Metaphysics of Natural Language*, The MIT Press, Cambridge, Mass.

MacWhinney, B. (1999) “The Emergence of Language from Embodiment,” in B. MacWhinney (ed.) 1999.

MacWhinney, B. (ed.) (1999) *The Emergence of Language*, Lawrence Erlbaum Associates, Mahwah, New Jersey.

Reichenbach, H. (1947) *Elements of Symbolic Logic*, The Free Press, New York.