

Database Semantics for Talking Autonomous Robots

Roland Hausser

Universität Erlangen-Nürnberg (em.)

rrh@linguistik.uni-erlangen.de

Abstract Database Semantics (DBS) models the cycle of natural language communication as a transition from the hear to the think to the speak and back to the hear mode (turn taking). In contradistinction to the substitution-driven sign-based approaches of truth-conditional semantics and phrase structure grammar, DBS is data-driven and agent-based. The purpose is a theory of semantics for an autonomous robot with language.

Propositions are content in DBS, instead of denoting truth values (Sects. 1–3). Content is built from the semantic kinds of referent, property, and relation, which are concatenated by the classical semantic relations of structure, i.e. functor-argument and coordination. To enable reference as an agent-internal cognitive process, language and nonlanguage contents use the same computational data structure and operation kinds, and differ mostly in the presence vs. absence of language-dependent surface values.

DBS consists of (i) an interface, (ii) a memory, and (iii) an operation component.¹ The interface component mediates between the agent's cognition and its external and internal environment, represented as raw data provided by sensors and activators (Sects. 4–7). The data of the agent's moment by moment monitoring are stored at the memory's now front. As part of the on-board control unit, the now front is the location for performing the procedures of the operation component, resulting in content.

keywords: data structure, data base schema, pattern matching, turn taking, type-token, grounding, sensory and processing media and modalities, reference

1 Building Content in the Agent's Hear Mode

DBS defines a content in terms of concepts like *square* (7.1) or *blue* (7.3) connected with the classical semantic relations of structure such as subject/predicate and conjunct–conjunct. The concepts are supplied by the agent's memory and defined as types. In recognition, they are activated by matching raw data provided by the interface component, resulting in tokens.² In action, a type is adapted to a current purpose as a token and realized as raw data (7.2, 7.4).

For concatenation, the concepts are embedded as core values into nonrecursive feature structures with ordered attributes, called proplets. Proplets serve as the computational data structure of DBS. The semantic relations between proplets are established by address. Consider the following example:

¹ The components correspond roughly to those of a von Neumann machine (Neumann, J.v., 1945): the (i) interface component corresponds to the vNm input-output device, the (ii) memory (database) component corresponds to the vNm memory, and the (iii) operation component performs functions of the vNm arithmetic-logic unit.

² The type-token distinction was introduced by C. S. Peirce (CP 4:537).

1.1 THE CONTENT OF The dog snored.

sur: noun: dog cat: def sg sem: fnc: snore mdr: nc: pc: prn: 24	sur: verb: snore cat: #n' decl sem: past ind arg: dog mdr: nc: pc: prn:24
---	---

The proplets implement the subject/predicate relation by using the noun value **dog** of the first proplet as the **arg** value of the second, and the **verb** value **snore** as the **fnc** value of the first (bidirectional pointering).

In the hear mode, the content 1.1 results from the following derivation:

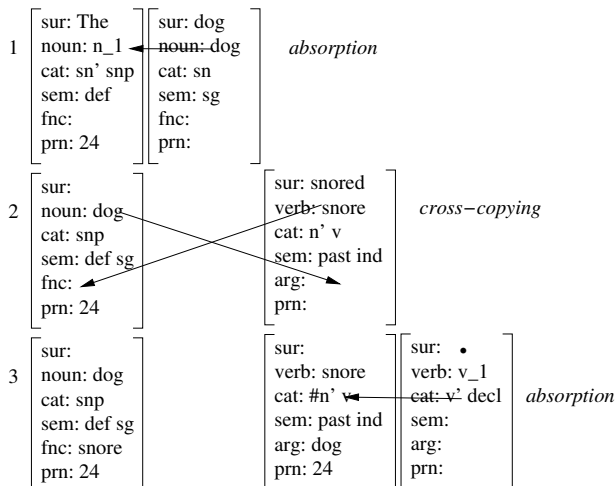
1.2 SURFACE COMPOSITIONAL TIME-LINEAR HEAR MODE DERIVATION

unanalyzed surface

The dog snored •
automatic word form recognition (lexical lookup)

sur: The noun: n_1 cat: sn' snp sem: def fnc: prn:	sur: dog noun: dog cat: sn sem: sg fnc: prn:	sur: snored verb: snore cat: n' v sem: past ind arg: prn:	sur: • verb: v_1 cat: v' decl sem: arg: prn:
---	---	--	---

syntactic-semantic parsing



result

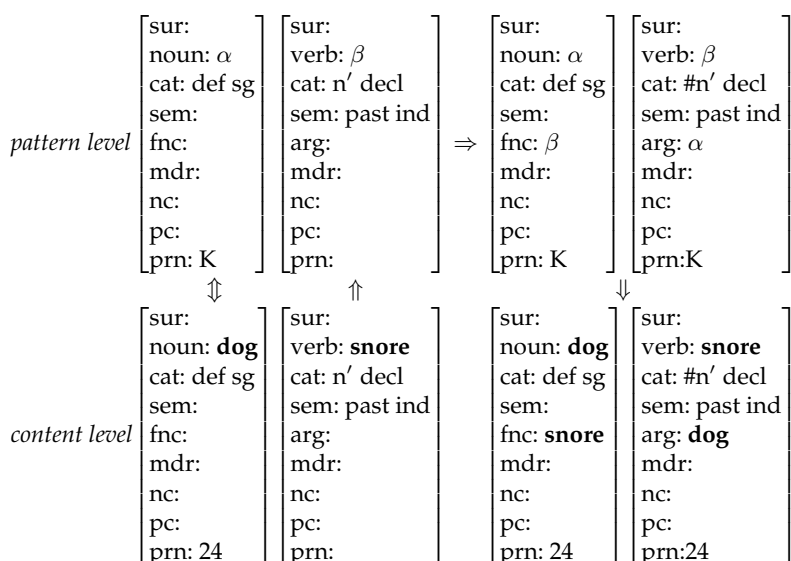
sur: noun: dog cat: snp sem: def sg fnc: snore prn: 24	sur: verb: snore cat: #n' decl sem: past ind arg: dog prn: 24
---	--

In DBS, hear mode derivations are time-linear and surface compositional.

There are three kinds of hear mode operations: (1) cross-copying, (2) absorption, and (3) suspension. Cross-copying encodes the semantic relations of structure such as SBJ×PRED (line 2). Absorption combines function words with content word such as DETUCN (line 1). Suspension such as ADV~NOM (TE_{EX} 3.1.3) applies if no semantic relation exists for connecting the next word with the content processed so far, as in *Perhaps ~ Fido (is still sleeping there now)*.

DBS operations consist of (i) an antecedent, (ii) a connective, and (iii) a consequent. Defined as proplet patterns, DBS operations are data-driven in that they are activated by content proplets matching part of the pattern as input. Consider the hear mode operation SBJ×PRED as it applies in line 2 of 1.2:

1.3 HEAR MODE OPERATION SBJ×PRED APPLYING BY CROSS-COPYING



In the hear mode, the second proplet at the content level, here *snore*, resulted from automatic word form recognition. By *snore* matching the second input pattern at the pattern level (↑), the operation is triggered to look for a content proplet matching its first input pattern (⇕) at the now front (2.2), here the special case of a sentence start consisting of a single lexical proplet. By binding α to *dog* and β to *snore*, the consequent produces the output as content proplets (↓).

2 Storage and Retrieval of Content in the On-Board Memory

Contents derived in the hear mode and activated in the think-speak mode (Sect. 3) have in common that they are defined as sets of self-contained proplets, concatenated by proplet-internal address. As sets, the proplets of a content are order-free, which is essential for their storage in and retrieval from the agent's A-memory (formerly called word bank). The database schema of A-memory is defined as follows:

2.1 TWO-DIMENSIONAL DATABASE SCHEMA OF A-MEMORY

– *horizontal token line*

Horizontally, proplets with the same core value are stored in the same token line in the time-linear order of their arrival.

– *vertical column of token lines*

Vertically, token lines are in the alphabetical order induced by the letter sequence of their core value.

The arrival order of the member proplets is reflected by (a) the position in their token line and by (b) their prn value. The (i) *member proplets* are followed by a free slot as part of the column called the (ii) *now front*, and the (iii) *owner*:³

2.2 A-MEMORY BEFORE INCREMENTAL STORAGE OF 1.1 AT THE NOW FRONT

<i>(i) member proplets</i>	<i>(ii) now front</i>	<i>(iii) owner</i>
... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: dog} \\ \dots \\ \text{prn: 3} \end{array} \right]$... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: dog} \\ \dots \\ \text{prn: 6} \end{array} \right]$	dog
... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: snore} \\ \dots \\ \text{prn: 5} \end{array} \right]$... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: snore} \\ \dots \\ \text{prn: 7} \end{array} \right]$	snore
...	...	

The owners equal the core values in their token line and are used for access in storage and retrieval. Proplets provided by current recognition, by A-memory, or by inferencing are stored at the now front in the token line corresponding to their core value:

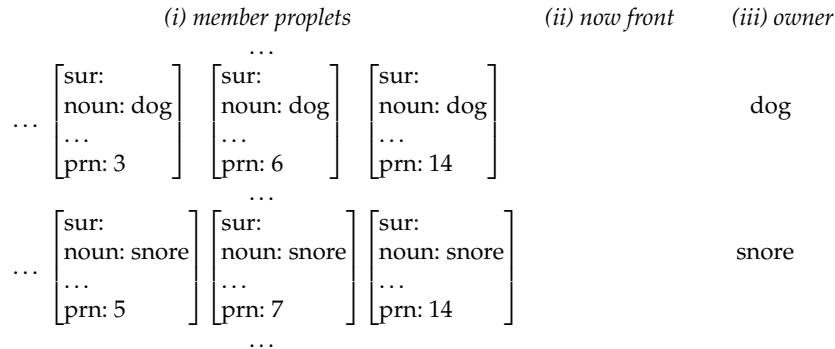
2.3 STORAGE OF 1.1 AT THE NOW FRONT OF A-MEMORY

<i>(i) member proplets</i>	<i>(ii) now front</i>	<i>(iii) owner</i>
... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: dog} \\ \dots \\ \text{prn: 3} \end{array} \right]$... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: dog} \\ \dots \\ \text{prn: 6} \end{array} \right]$	$\left[\begin{array}{l} \text{sur: chien} \\ \text{noun: dog} \\ \dots \\ \text{prn: 14} \end{array} \right]$ dog
... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: snore} \\ \dots \\ \text{prn: 5} \end{array} \right]$... $\left[\begin{array}{l} \text{sur:} \\ \text{noun: snore} \\ \dots \\ \text{prn: 7} \end{array} \right]$	$\left[\begin{array}{l} \text{sur: ronfler} \\ \text{noun: snore} \\ \dots \\ \text{prn: 14} \end{array} \right]$ snore
...	...	

³ The terminology of member proplets and owner values is reminiscent of the member and owner records in a classic network database (Elmasri and Navathe 1989¹–2017⁷), which inspired the database schema of the A-memory in DBS.

Once a content has been assembled as a proposition, the now front is cleared by moving it and the owners to the right into fresh memory space (loom-like clearance, Sect. 3). This leaves the proplets of the current content behind in what is becoming their permanent storage location as member proplets never to be changed, like sediment.

2.4 A-MEMORY AFTER NOW FRONT CLEARANCE

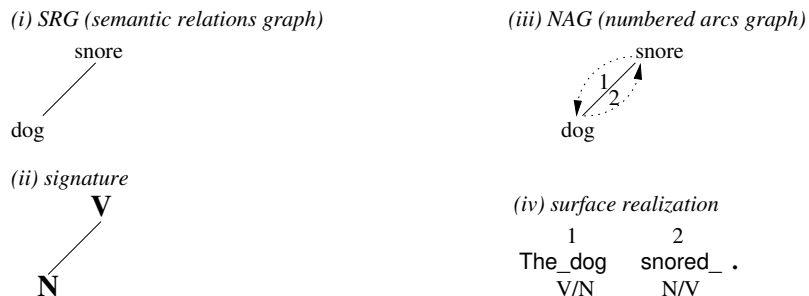


Clearance of the current now front is triggered when its proplets have ceased to be candidates for additional concatenations. This is basically the case when an elementary proposition is completed (formally indicated by the automatic incrementation of the prn value for the next proposition). Exceptions are extrapositional (i) coordination and (ii) functor-argument. In these two cases, the verb of the completed proposition must remain at the now front for cross-copying with the verb of the next proposition until the extrapositional relation has been established (strictly time-linear derivation order).

3 Speak Mode Riding Piggyback on the Think Mode

The classical semantic relations of structure between the order-free proplets defined by address and stored in A-memory may be shown graphically as follows:

3.1 SEMANTIC RELATIONS GRAPH UNDERLYING THE CONTENT 1.1



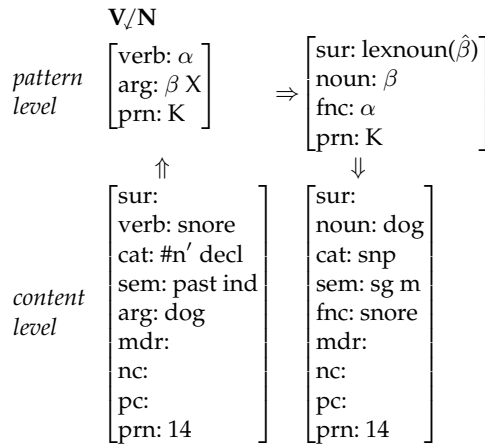
The *(iv) surface realization* shows the language-dependent production and is re-

stricted to the speak mode riding piggy-back on the think mode navigation. The concepts **dog** and **snore** are provided by the agent's interface component as a declarative definition and an operational implementation (grounding, Sect. 7).

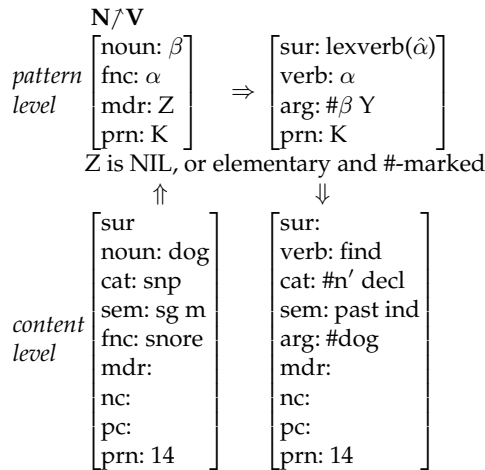
The arc numbers of the NAG are used for specifying (i) a think mode navigation and (ii) a think-speak mode surface production as shown by the (iv) *surface realization*. In the think mode, there are 14 kinds of traversal operations: (1) predicate/subject, (2) subject/^predicate, (3) predicate\object, (4) object\^predicate, (5) noun\adnominal, (6) adnominal^noun, (7) verb\adverbial, (8) adverbial^verb, (9) noun→noun, (10) noun←noun, (11) verb→verb, (12) verb←verb, (13) adnominal→adnominal, and (14) adnominal←adnominal.

The think mode operations driving the traversal of the NAG in 3.1 are V/N and N/V, and apply as follows:⁴

3.2 NAVIGATING WITH V/N FROM *snore* TO *dog* (arc 1)



3.3 NAVIGATING WITH N/V FROM *dog* BACK TO *snore* (arc 2)

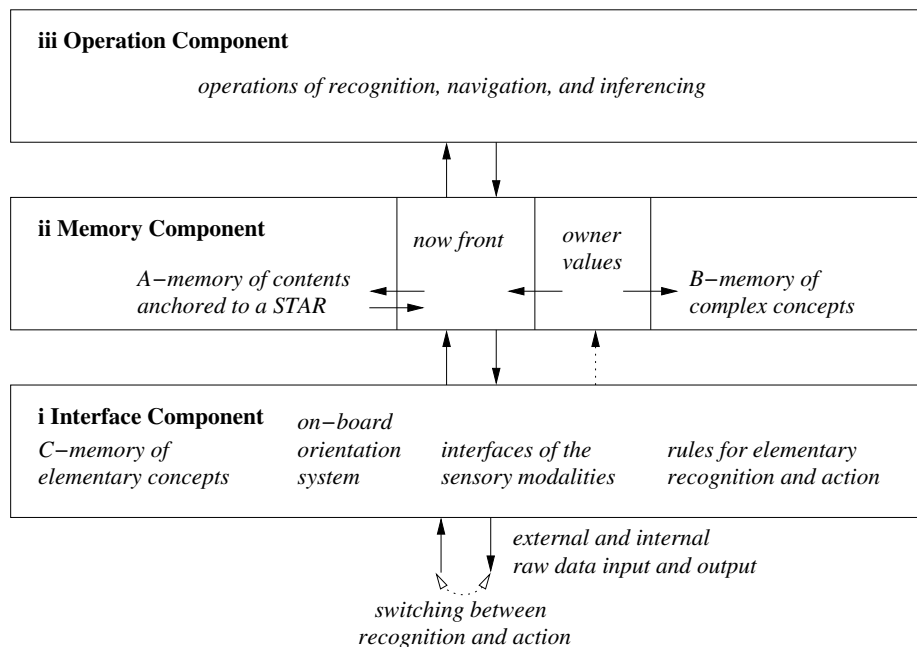


If the lexnoun rules in the sur slot of the output patterns are switched on (as assumed in the surface realization of 3.1), they generate a language-dependent surface using relevant values of the output proplet.

4 Graphical Summary of the DBS Component Structure

The component structure of a DBS agent may be summarized graphically as follows:

4.1 TWO-DIMENSIONAL LAYOUT OF DBS COGNITION COMPONENTS



Cognitive content is processed at the now front. It gets proplets (a) from the interface component (aided by the owners) and (b) from A-memory. For processing, the now front provides proplets as input to (iii) the operations, which either replace the input with their output or add their output to the input. As the now front is cleared in regular intervals by moving into fresh memory space (Sect. 2), the processed proplets are left behind in A-memory like sediment (loom-like clearance). Processing may also result in blueprints for action, which may be copied to the interface component for realization.

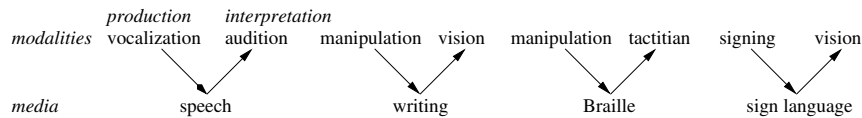
³ Because the Xfig graphics editor used here does not provide a satisfactory representation of arrows in the linear notation of speak mode operation names, the arrow heads are omitted in the (iv) surface realization. Nevertheless, the direction of a traversal is specified unambiguously by the arc number written directly above in the top line. For further detail see NLC 6.1.4 ff.; CLaTR Chap. 7.

5 Sensory Media, Processing Media, and their Modalities

The functional equivalence required between the artificial agent and its natural prototype is defined at a level of abstraction which is above the distinction between different processing media, such as natural, mechanical, and electronic processing. Functional equivalence is shown, for example, by the basic operations of arithmetic: $3+4$ equals 7 no matter whether the calculation is performed by (i) a human,⁴ (ii) a mechanical calculator, or (iii) a computer.

In addition to the processing media there are the sensory media. In natural language communication, there exist four, each of which has two sensory modalities.⁵ For example, if the speaker chooses the medium of speech, the only sensory modality for production is vocalization (\searrow), which leaves the hearer no other option than using the sensory modality of audition (\nearrow). This asymmetry of modalities holds also for the other sensory media of natural language, namely writing, Braille, and sign language:

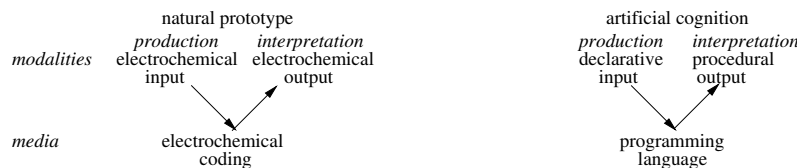
5.1 SENSORY MEDIA AND THEIR MODALITIES IN NATURAL LANGUAGE



In terms of human evolution, the primary sensory medium is speech.

While the sensory media must be the same in the natural prototype and the artificial counterpart, as required by functional equivalence, the processing media are fundamentally different between the two. For the natural prototype, neurology suggests an electrochemical processing medium, though much is still unknown.⁶ In artificial DBS cognition, in contrast, the processing medium is a programming language; its processing modalities are (i) the declarative specification of commands for interpretation by the computer and (ii) their procedural execution by the computer's electronic operations.

5.2 PROCESSING MEDIA AND THEIR DUAL PROCESSING MODALITIES



⁶ The operations of arithmetic as they are processed by the human brain are described by Menon (2011).

⁶ In the literature, the term modality has a multitude of uses, such as the temperature (Dodt and Zotterman 1952), the logical (Barcan Marcus 1961), and the epistemic (Kiefer 2018) modalities.

⁶ For an early overview see Benson (1994).

Utilizing a programming language as the processing medium of an artificial agent requires an interface component capable of efficiently mediating between raw data and an alphanumeric representation in recognition and action.

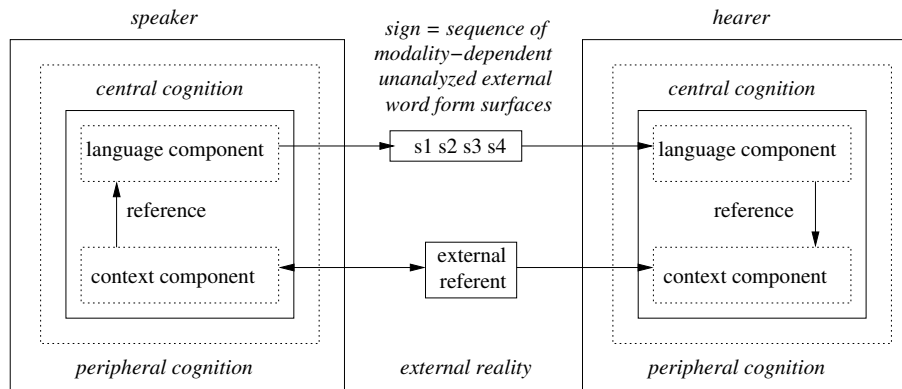
6 Reference as a Purely Cognitive Process

Sign-based philosophy defines reference as a relation between language (referring part) and the world (referred-to part).⁷ Agent-based DBS, in contrast, confines reference to nouns (Hausser 2019, 1.5.3, 12.3.3) and distinguishes (1) between referring nouns with and without external surfaces and (2) between referred-to nouns with and without external⁸ counterparts. The two distinctions may be characterized by the binary features $[\pm\text{surface}]$ and $[\pm\text{external}]$, whereby $[\text{+external}]$ reference is called *immediate*, while $[\text{-external}]$ reference is called *mediated* (FoCL 4.3.1).

For example, identifying “the man with the brown coat” (Quine 1960) with someone seen before, or identifying an unusual building with an earlier language content, e.g. something read in a guide book or heard about, are $[\text{-surface } \text{+external}]$. Talking about Aristotle or J.S. Bach, in contrast, is $[\text{+surface } \text{-external}]$.

Let us go systematically through the four kinds of generalized DBS reference,⁹ beginning with the $[\text{+surface } \text{+external}]$ constellation between speaker and hearer:

6.1 IMMEDIATE REFERENCE IN LANGUAGE COMMUNICATION



Agent-externally, language surfaces (shown here as s1 s2 s3 s4) are modality-specific unanalyzed external signs (raw data) which are passed from the speaker

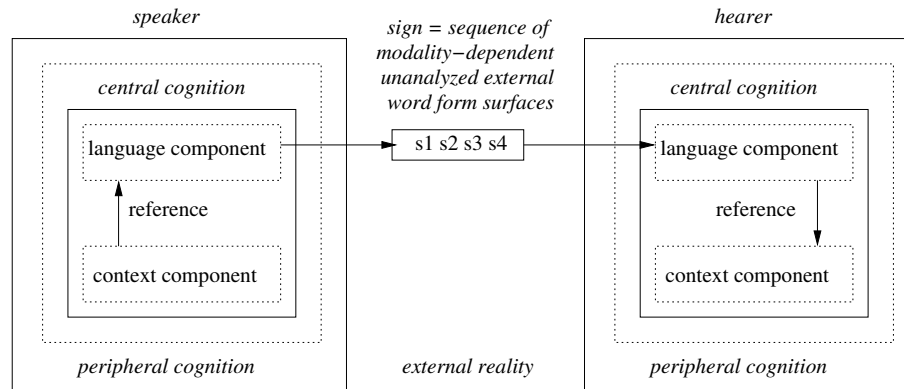
⁷ Reimer and Michaelson (2014) extend the referring part from language to “representational tokens,” which include cave paintings, pantomime, photographs, videos, etc. DBS goes further by generalizing the referring part to content per se, i.e. without the need for any external representation (6.3).

⁸ Newell and Simon call the agent’s external surroundings the *task environment* (Newell and Simon 1972).

to the hearer and have neither meaning nor any grammatical properties, but may be measured by the natural sciences.

The corresponding [+surface –external] constellation between the speaker and the hearer is as follows:

6.2 MEDIATED REFERENCE IN LANGUAGE COMMUNICATION



The reference relation begins with content in the memory of the speaker and ends as content in the memory of the hearer. The mechanisms of assigning surfaces to content in the speak mode and content to surfaces in the hear mode are the same in immediate and mediated language reference.¹⁰

The graphs 6.1 and 6.2 show the speaker on the left, the sign in English writing order in the middle, and the hearer on the right. This is a possible constellation which is in concord with the naive assumption that time passes with the sun from left to right (→) on the Northern Hemisphere. Yet it appears that the first surface *s1* leaves the speaker last and the last surface *s4* arrives at the hearer first, which would be functionally incorrect.

It is a pseudo-problem, however, which vanishes if each surface is transmitted individually and placed to the right of its predecessor, i.e. ((*s1 s2 s3 s4*)). This *left-associative*¹¹ departure and arrival structure allows incremental surface by surface processing, provided the derivation order is based on computing possible continuations, as in Left-Associative grammar (LAG, Hausser 1992).

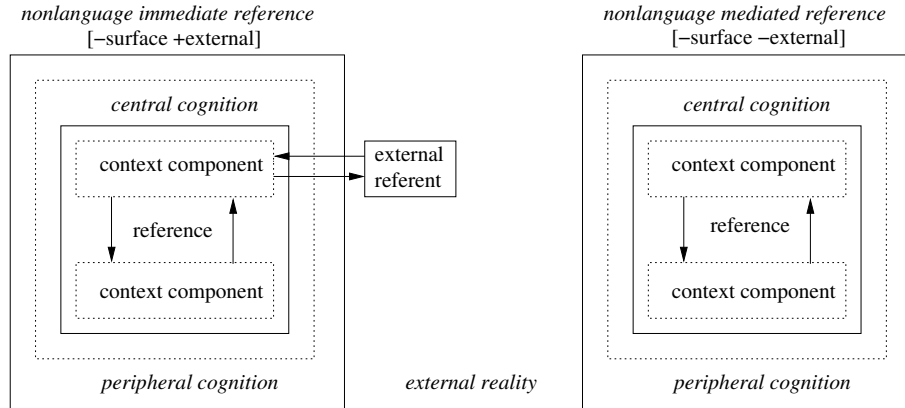
Nonlanguage reference differs from language reference in that it is [–surface]. Thereby nonlanguage immediate reference is [–surface +external] while nonlanguage mediated reference is [–surface –external]:

⁹ The [±surface] and [±external] distinctions are not available in truth-conditional semantics and generative grammar because their sign-based ontology provides neither for cognition nor for cognitive modes.

¹⁰ On the phone, the speaker may use an immediate reference which is mediated for the hearer and vice versa. For example, if the speaker explains to the hearer where to find something in the speaker's apartment, the speaker uses mediated reference and the hearer immediate reference.

¹¹ Aho and Ullman (1977), p. 47; FoCL 10.1.1.

6.3 NONLANGUAGE IMMEDIATE VS. MEDIATED REFERENCE

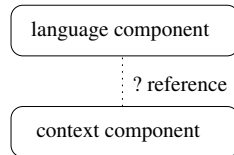


The referring content in the [-surface +external] constellation is a current nonlanguage recognition. In the [-surface -external] constellation of nonlanguage mediated reference, in contrast, the referring content is activated without an external trigger, for example, by reasoning. In both, the referred-to content is resonating (Hausser 2019, Sects. 3.2, 3.3) in memory.

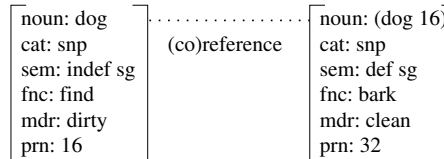
Computationally, the conceptual view of reference as a vertical interaction between two separate components in 6.1–6.3 is implemented as a horizontal relation between two proplets in the same token line:

6.4 COMPARING THE NAIVE AND THE COMPUTATIONAL SOLUTION

Component proposal (preliminary)



Token line solution in A-memory

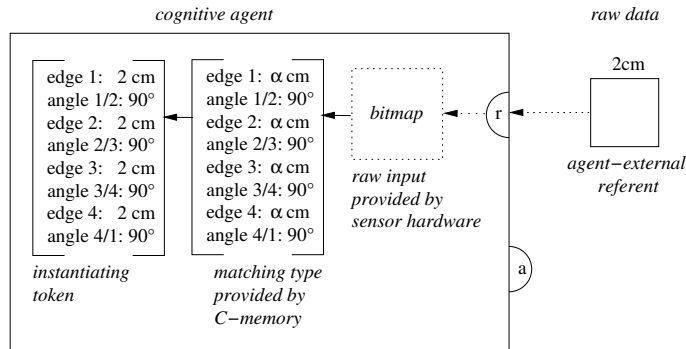


Because the semantic kind of referent is limited to nouns while that of adjs is property and that of transitive verbs is relation, (co)reference is restricted to nouns (Hausser 2019, 6.4.1, 6.4.4–6.4.6). The core value of the referring noun (shadow, copy) at the now front is always an address. The core value of the referred-to noun (referent, original) is never an address. The fnc and mdr values are free (identity in change, Hausser 2019, 6.4.7).

7 Grounding

The semantics of DBS is grounded (Barsalou et al. 2003, Steels 2008, Spranger et al. 2010). In recognition, concept types (supplied by the agent's memory) are matched with raw data (provided by sensors):

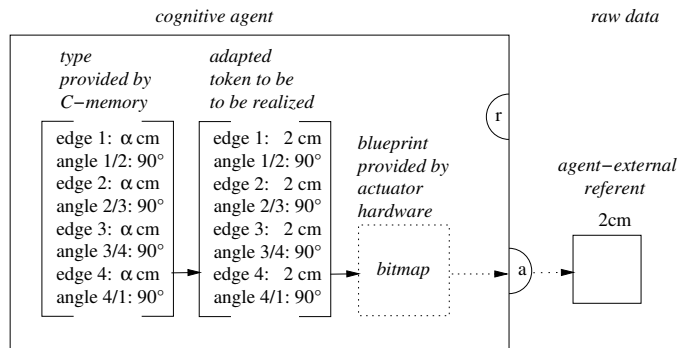
7.1 RECOGNITION OF square



The raw data are supplied by a sensor, here for vision, as input to the interface component. The raw data are matched by the type, resulting in a token.

In action, a type is adapted to a token for the purpose at hand and realized by the agent's actuators as raw data:

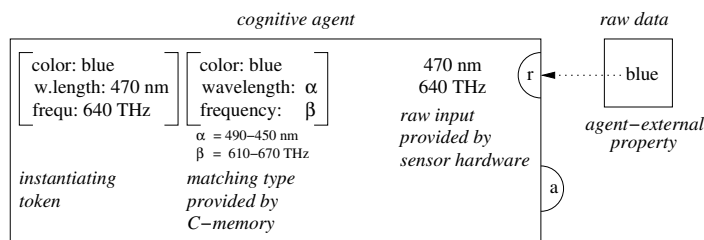
7.2 ACTION OF REALIZING square



The token is used as a blueprint for action, (e.g. drawing a square).

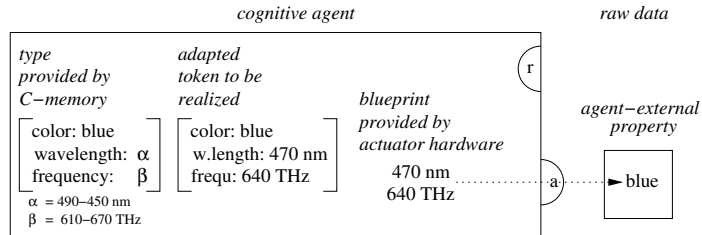
Next consider the recognition of a color, here blue:

7.3 RECOGNITION OF blue



An example of the corresponding action is turning on the color blue, as in a cuttlefish using its chromatophores:

7.4 ACTION OF REALIZING blue



The concept type matches different shades of blue, whereby the variables α and β are instantiated as constants in the resulting token. Recognizing the color blue is a general mechanism which may be applied to all colors. It may be expanded to infrared and ultraviolet, and to varying intensity.¹²

Pattern matching based on the type-token relation applies to nonlanguage items (e.g. 7.1, 7.2, 7.3, 7.4) and language surfaces alike. For example, in the surfaces of spoken language the type generalizes over different pitch, timbre, dialect, and speaker-dependent pronunciation. In written language, the type generalizes over the size, color, and font of the letters. Computational type-token matching is more adequate descriptively than the nonbivalent (Rescher 1969; FoCL Chap. 20.5) and fuzzy (Zadeh 1965) logics for treating vagueness because type-token matching treats the phenomenon at the root (best candidate principle in pattern matching, FoCL Sect. 5.2) instead of tinkering with the truth tables of Propositional Calculus.

Bibliography

- Aho, A.V. and J.D. Ullman (1977) *Principles of Compiler Design*, Reading, Mass.: Addison-Wesley
- Barcan Marcus, R. (1961) "Modalities and Intensional Languages," *Synthese*, Vol. 13.4:303–322
- Barsalou, W., W.K. Simmons, A.K. Barbey, and C.D. Wilson (2003) "Grounding conceptual knowledge in modality-specific systems," *TRENDS in Cognitive Sciences*, Elsevier, Vol. 7.2:84–91
- Benson, D.F. (1994) *The Neurology of Thinking*, New York: OUP
<https://doi.org/10.1002/ana.410360535>
- Biederman, I. (1987) "Recognition-by-components: a theory of human image understanding," *Psychological Review*, Vol. 94:115–147

¹² Complementary approaches from cognitive psychology are prototype theory (Rosch 1975) and composition based on geons (Biederman 1987).

- Boas, F. (1911) *Handbook of American Indian languages (Vol. 1)*, Bureau of American Ethnology, Bulletin 40. Washington: Government Print Office (Smithsonian Institution, Bureau of American Ethnology)
- Dotz, E., and Y. Zotterman (1952) "The Discharge of Specific, Cold Fibres at High Temperatures," *Acta Physiologica Scandinavica*, Vol 26.4: 358–365, December 1952
- Elmasri, R., and S.B. Navathe (2017) *Fundamentals of Database Systems, 7th ed.* (1st ed. 1989), Redwood City, CA: Benjamin-Cummings
- Hausser R. (1992) "Complexity in Left-Associative Grammar," *Theoretical Computer Science*, Vol. 106.2:283-308
- Hausser, R. (2019) *Computational Cognition, Integrated DBS Software Design for Data-Driven Cognitive Processing*, pp. i–xii, 1–237, lagrammar.net
- Kiefer, F. (2018) "Two kinds of epistemic modality in Hungarian," in Guentchéva, Z. (ed.) *Empirical Approaches to Language Typology*, DOI: 10.1515/9783110572261–013
- Menon, V. (2011) "Developmental cognitive neuroscience of arithmetic: implications for learning and education," *ZDM (Zentralblatt für Didaktik der Mathematik)*. Vol. 42.6:515–525
- Neumann, J.v. (1945) *First Draft of a Report on the EDVAC*, *IEEE Annals of the History of Computing*, 27-75
- Newell, A., and H.A. Simon (1972) *Human Problem Solving*, Englewood Cliffs, New Jersey: Prentice-Hall
- Peirce, C.S. (1931–1935) *Collected Papers*. C. Hartshorne and P. Weiss (eds.), Cambridge, MA: Harvard Univ. Press
- Putten, S. van (2019) "Perception verbs and the conceptualization of the senses: The case of Avatime," Accepted for publication in *Linguistics*
- Quine, W.v.O. (1960) *Word and Object*, Cambridge, Mass.: MIT Press
- Reimer, M., and E. Michaelson (2014). "Reference," *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/win2014/entries/reference/>
- Rescher, N. (1969) *Many-valued Logic*, New York: McGraw-Hill
- Rosch, E. (1975) "Cognitive representations of semantic categories," *J. of Experimental Psychology*, General 104:192–253
- Spranger, M., M. Loetzsch, and S. Pauw (2010) "Open-ended Grounded Semantics" <https://csl.sony.fr/wp-content/themes/sony/uploads/pdf/spranger-10b.pdf>
- Steels, L. (2008) "The symbol grounding problem has been solved. so what's next?" in *Symbols and Embodiment: Debates on Meaning and Cognition*, ed., M. de Vega, Oxford: Oxford University Press
- Zadeh, L. (1965) "Fuzzy sets," *Information and Control*, Vol. 8:338–353