Sentence meaning as argumentative dialogues

Davide CATTA (Univ. Montpellier, LIRMM)
Alda MARI (IJN, CNRS, ENS, EHESS)
Christian RETORÉ (Univ. Montpellier, LIRMM)

Principles In formal semantics the meaning of a sentence is usually defined as the situations in which the sentence is true and usually formalised with possible worlds semantics. Let us twist this classical view into a dialogical one:

The meaning of a sentence A asserted by speaker P is defined as the set of all possible justifications of A, which are argumentative dialogues starting with A won by speaker P.

where:
An argumentative dialogue *is a sequence of what we call* utterances*, namely* assertions *(*!-mode *prefixed sentences) or as* questions *(*?-mode *prefixed sentences). More precisely an argumentative dialogue for a sentence A is a finite alternate sequence $S_O = !A, S_1, \ldots S_N$, where even utterances (including the initial utterance !A that is the assertion of the initial sentence) are told by the proponent (P) and odd utterances are told by the opponent (O).*

There are answering rules *often referred to as* attack *and* defence *rules defining how O (resp. P) may answer an utterance U_{2p} by P (resp. an utterance U_{2p+1} by O) according to the mode, assertion or question, and the logical structure of the answered utterance.*

An argumentative dialogue is won by *P if and only it the last utterance is an assertion made by P, in which all question asked by O have been successfully answered by P, and O cannot reply anymore according to the answering rules. The argumentative dialogue is* won by *O otherwise.*

Readers accustomed with dialogical logic [6] will recognise that our informal definition has a dialogical logic flavour. Here are two examples of answering rules:

C: Conditional Rule When a conditional (*if A then B*) is asserted by a speaker the other one questions it by asserting *A* and asking for a justification of *B*. In other words *?(if A then B)* is the same as *!A, ?B*.

At: Atomic Rule *P* may affirm an atomic proposition *q* only if *q* was already affirmed by *O* earlier in the dialogue.

(every logical connective has answering rules, they are not included for lack of space)

Argumentative dialogues can be recursively enumerated. Indeed, argumentative dialogues are among the alternate sequences of sentences (which can be enumerated), and argumentative dialogues are the alternate sequences matching the answering rules.

Observe that such a view of meaning is internal to linguistic activity: both sentences and dialogues are natural language objects.

Proceedings of the 22nd Workshop on the Semantics and Pragmatics of Dialogue, November 8-10, 2018, Aix-en-Provence, France.

Here is an example of an argumentative dialogue:

0. P: ! $(S_1 \to S_2) \to (S_2 \to S_3 \to (S_1 \to S_3))$
1. O: ! $S_1 \to S_2$, ? $S_2 \to S_3 \to (S_1 \to S_3)$
2. P: ! $S_2 \to S_3 \to (S_1 \to S_3)$
3. O: ! $S_2 \to S_3$, ? $S_1 \to S_3$
4. P: ! $S_1 \to S_3$
5. O: ! S_1, S_3
6. P: ! S_1, ?S_2
7. O: ! S_2
8. P: ! S_2, ?S_3
9. O: !S_3
10. P: !S_3

where:

S_1: John kills Mary,.
S_2: John will go to jail.
S_3: John will pay for his crime.

This argumentative dialogue is won by *P*.

Formalisation and computability Could this view be formalised and implemented? Are there lexicons, grammars and algorithms computing the argumentative dialogues associated with a sentence? As usual in formal and computational linguistics, feasibility depends on knowledge representation and existing linguistic resources, hence on the context. Below are two extreme cases:

When the considered language fragment is *natural logic* [10] the correctness of an argumentative dialogue is easily checked, and it is even possible to effectively compute all the argumentative dialogues starting with a given sentence *S*; this set is, according to the view of the present paper, the semantics of the sentence. Indeed, in natural logic, sentences can be mapped, automatically and unambiguously, to formulas of a decidable fragment of (an extension of) first order logic (similar to description logic). Natural logic also provides completely formalised answering rules. Hence in natural logic, the argumentative dialogues starting with *A* and won by *P* correspond to the dialogues of dialogical logic [6] starting with *A*, and they are easily computed. A difference is that the ultimate defences of *P* may consist in axioms which are hitherto unknown — they are learnt that way.

In *ordinary conversation* a complete and computable formalisation is much more problematic. There does exist wide scale syntactic and semantic analysis systems (e.g. [9]) that map sentences to logical formulas (using compositionality and λ-DRT [11], and DRT anaphora resolution). In order to verify and enumerate the possible argumentative dialogues justifying a sentence, the systems needs at least to know the axioms encoding lexical meanings, as well as the axioms describing the situation under discussion and the proponent beliefs. In general the later resources are not available, hence argumentative dialogues are hard to check or enumerate automatically. But, when resources are available, the set of argumentative dialogues, i.e. the semantics, can be computed.

Relation to inferentialism Our dialogical view of semantics is clearly related to the inferentialist view of meaning [5, 1] which has already been developed, but not much, in formal semantics [3, 12, 8, 7].

A positive consequence is that our proposal for semantics is computabl, because inference rules and proofs or dialogues are finite and enumerable. Argumentative dialogues can be checked and even enumerated from some limited and partial knowledge of the situation. This is clearly a cognitive and computational improvement over the hardly enumerable infinity of possible worlds – furthermore, a finite description of a given possible world is itself hardly computable.

But, for our proposal to be part of inferentialism we should respect the main requirement of a *Theory of Meaning* as described in [5, 4]:

> *The knowledge of the sense of a sentence or expression must be — in principle — completely observable and publicly testable.*

Thus, the speaker's knowledge must be observable in the interactions between the protagonists and any speakers' disagreement regarding the meaning of an expression must emerge under some circumstances. This is indeed the case in *argumentative dialogues*: such a disagreement on the interpretation of an expression A will result in incompatible arguments for justifying A, and such conflicts are observable. As an example, let us consider the following simplistic argumentative dialogue:

0. P: John is not a murderer
1. O: John is a murderer, he killed Mary
2. P: I grant that he killed Mary but it was by accident

The opponent consider that*x killed y* entails that *x is murderer* while the proponent refutes this claim by pointing out that *x killed y by accident*. When the meaning of an expression consists in the arguments justifying it, then we can observe that the respective interpretations by the opponent and by the proponent of *x is a murderer* differ. The above dialogue shows that the axioms representing the meaning of the atomic predicate $murderer(x)$ for each of the two speakers are visibly different — according to P *murderer* includes a notion of *deliberateness*.

Future prospects We are presently willing to explore the formal properties of argumentative dialogues but also willing to establish their empirical relevance. For instance do argumentative dialogues bring a finer grained notion of semantics? Do they tell apart expressions that usually gets the same semantic representation?

As the last section suggests, we plan to to characterise manifestability, that is to find hypotheses that would guarantee the emergence in an argumentative of any possible disagreement about word meaning. If the emergence of the disagreement can be triggered, then computing a dialogue exhibiting a disagreement can be viewed as a machine-learning procedure for "axioms".

Unsurprisingly, the practical development of natural language processing tools using such ideas can only be achieved if a very precise topic has been delimited. Indeed, before being developed, tested, improved and evaluated, a prototype would require sophisticated linguistic resources (lexicons, knowledge representation) .

References

[1] Robert Brandom. *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press, 2000.

[2] Stergios Chatzikyriakidis and Zhaohui Luo, editors. *Modern Perspectives in Type Theoretical Semantics*. Springer, 2017.

[3] Stergios Chatzikyriakidis and Zhaohui Luo. On the interpretation of common nouns: Types versus predicates. In *Modern Perspectives in Type Theoretical Semantics* [2].

[4] Cesare Cozzo. *Meaning and Argument: A Theory of Meaning Centred on Immediate Argumental Role*. Stockholm Studies in Philosophy. Almqvist & Wiksell International, 1994.

[5] Michael A. E. Dummett. What is a theory of meaning? In Samuel Guttenplan, editor, *Mind and Language*. Oxford University Press, 1975.

[6] Walter Felscher. Dialogues as a foundation for intuitionistic logic. In Dov M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, pages 115–145. Springer Netherlands, Dordrecht, 2002.

[7] Christophe Fouqueré and Myriam Quatrini. Argumentation and inference a unified approach. In *The Baltic International Yearbook of Cognition, Logic and Communication Volume 8: Games, Game Theory and Game Semantics*, pages 1–41. New Paririe Press, 2013.

[8] Nissim Francez. *Proof Theoretical Semantics*, volume 57 of *Studies in Logic*. College Publication, 2015.

[9] Richard Moot. The grail theorem prover: Type theory for syntax and semantics. In Chatzikyriakidis and Luo [2], pages 247–277.

[10] Lawrence Moss. Natural logic. In Shalom Lappin and Chris Fox, editors, *Handbook of Contemporary Semantic Theory*, chapter 18, pages 561–592. John Wiley and Sons, second edition, 2015.

[11] Reinhard Muskens. A Compositional Discourse Representation Theory. In Paul Dekker and Martin Stokhof, editors, *Proceedings of the Ninth Amsterdam Colloquium*, pages 467–486, Amsterdam, 1994. ILLC, University of Amsterdam.

[12] Aarne Ranta. *Type theoretical grammar*. Oxford University Press, 1994.