Facticity in semantic concept definitions
In SemETAP semantic text analyzer

Ivan Rygaev*

Laboratory of Computational Linguistics,
A. A. Kharkevich Institute for Information Transmission Problems RAS, Moscow, Russia

Introduction

Who are we?
> Developers of SemETAP semantic text analyzer.[1,2]

What do we want?
> To achieve near-human understanding of the text.

What is understanding? How can we measure it?
> By the amount of inferences made out of the text.

How can we test inferences?
> By questioning.

Hence SemETAP is able to:
- Translate an original sentence to a semantic representation in a formal language.[1]
- Apply logical rules to infer new knowledge.[7]
- Answer questions for which there is no direct answer in the text.[2]

Example

Having an input sentence:
> 1) John sold an umbrella to Peter

Humans can easily answer:
> 2) Who bought the umbrella? (Peter)
> 3) What did John give to Peter? (the umbrella)
> 4) What did John get? (money)
> 5) Who owns the umbrella? (Peter)

Where is the knowledge?
> In the meaning of the words
> *sell*, *buy*, *give*, *get*
> and *own*.

We feed the system with word definitions (or rather concept definitions)[3,7] in a formal language.

Concept definition example

```
Giving ?giving ->
?giving
    hasAgent (Agent ?agent)
    hasRecipient (Agent ?agent2)
    hasObject (Thing ?object)
    hasTime (TimeInterval ?givTime)
    hasPrecondition (Own
        hasAgent ?agent
        hasObject ?object
        hasTime (TimeInterval finishedBy ?givTime))
    hasResult (Own
        hasAgent ?agent2
        hasObject ?object
        hasTime (TimeInterval metBy ?givTime))
    hasSyncEvent (Getting
        hasRecipient ?agent2
        hasSource ?agent
        hasObject ?object
        hasSyncEvent ?giving)
```

Facticity

In a semantic graph there can be factual and non-factual nodes. Factual nodes are marked with an epistemic modality[5] concept, which represents the degree of speaker's confidence.

> 5) Peter wanted to buy an umbrella
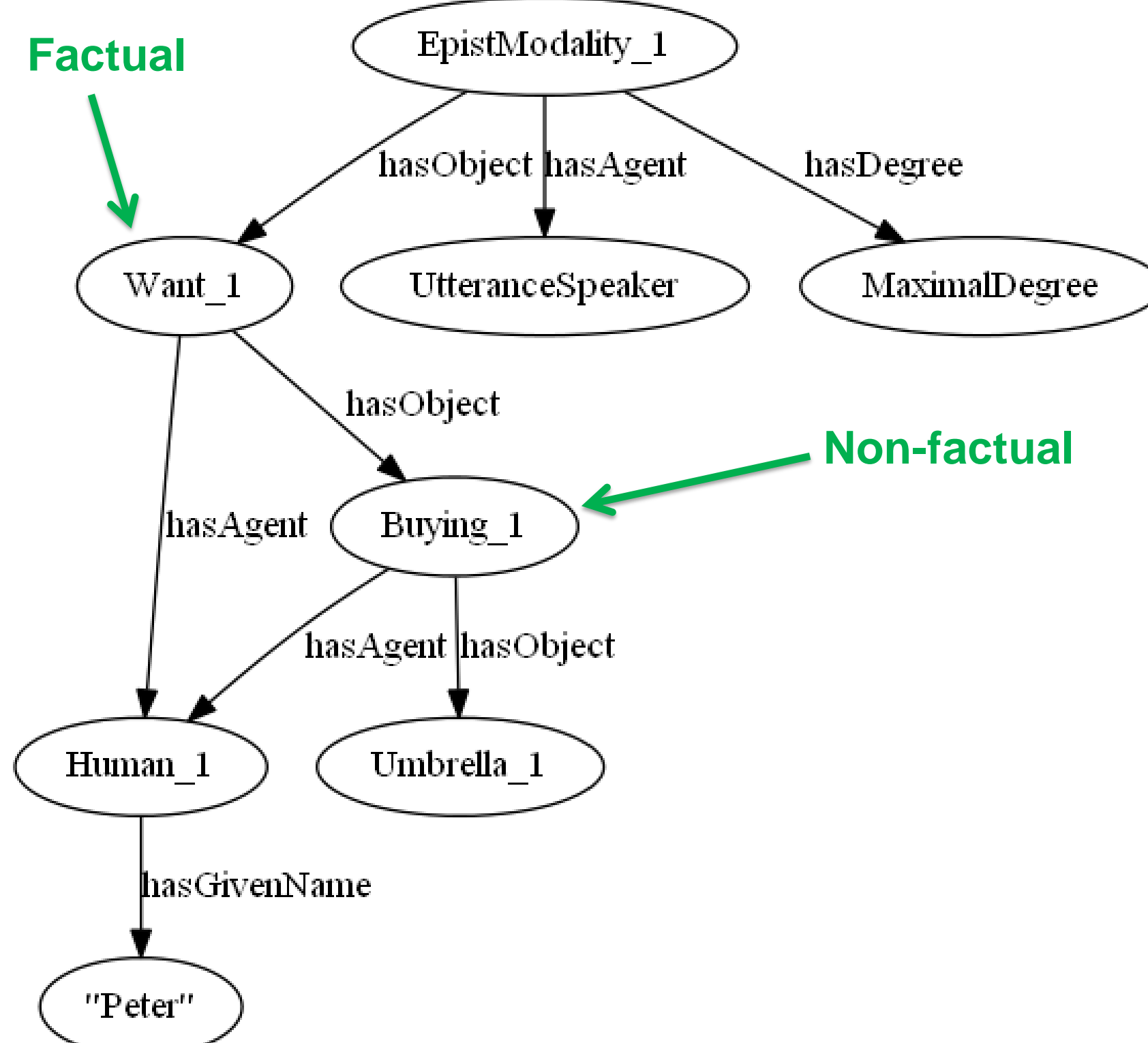
First only the top node of the sentence is marked as factual.

After that facticity can be transferred to lower nodes or inferred concepts by inference rules.

Problem

Rules of facticity transfer depend on:
- A lexical item being defined (*manage* vs *want*):
  > 6) Peter managed to buy an umbrella
- A part of the definition (*own* vs *get*):
  > 7) John did not give an umbrella to Peter =>
  > Peter did not get the umbrella (correct)
  > John did not own the umbrella (wrong)
- Usage context (polarity and aspect):[4,6]
  > 8) Peter bought an umbrella =>
  > Peter owns the umbrella (correct)
  > 9) Peter was buying an umbrella =>
  > Peter owns the umbrella (wrong)

This creates a huge amount of work for a linguist writing definitions. On the picture below arrows show the sequence of inferences. Each arrow needs to be analyzed in four different contexts (negation: yes/no, completion: yes/no).[6]
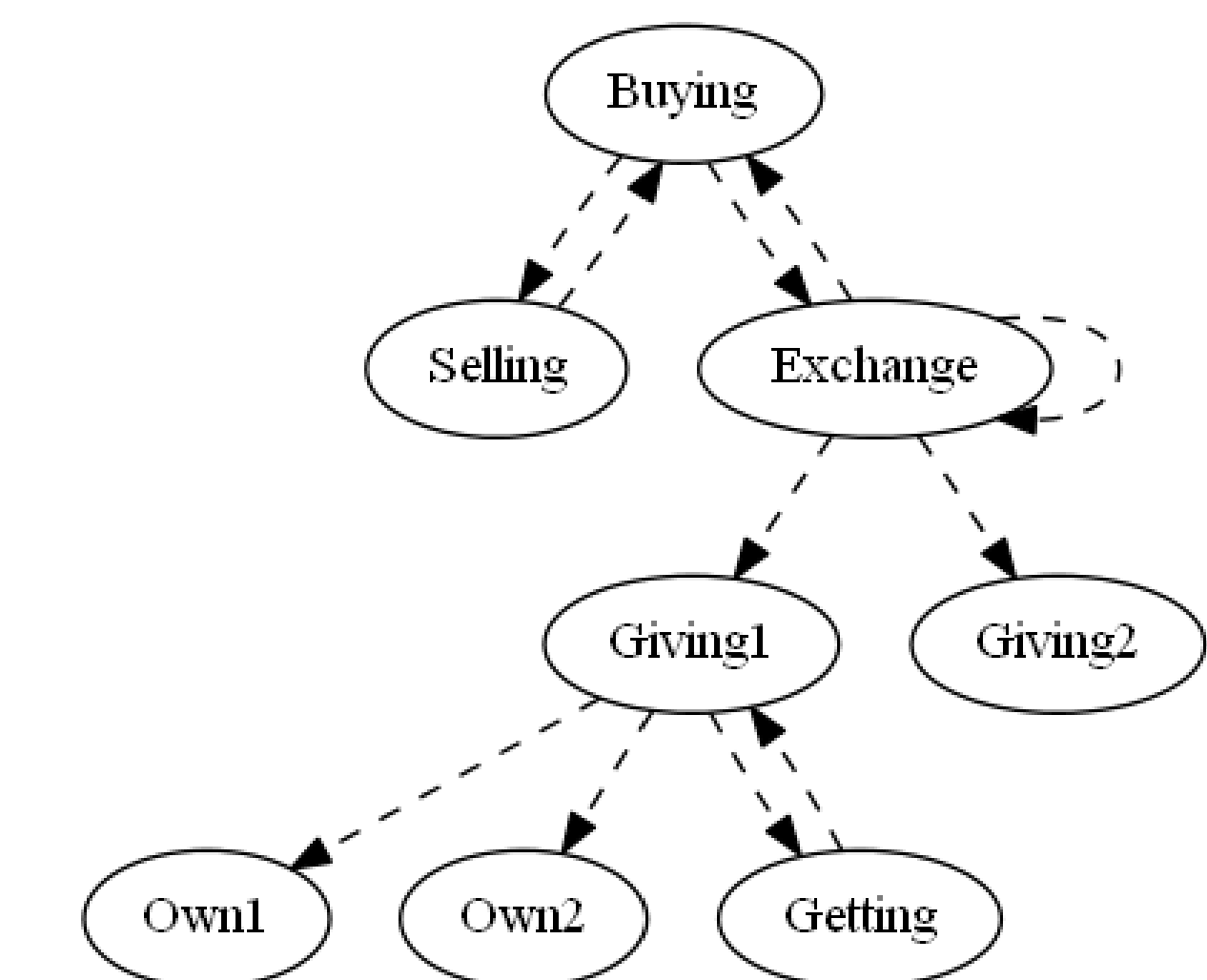
A solution is required to simplify the task.

Solution

How do we make things simpler?
> Easy – abstraction and generalization.

Implicative behavior of the definition part depends on how this part is related to the concept being defined. This can be generalized into a few implicative relations.
They are handy:
- They have natural understandable names and meaning.
- They take care of context dependent inferences.
- Linguist just needs to connect two parts of the definition via one of these relations.
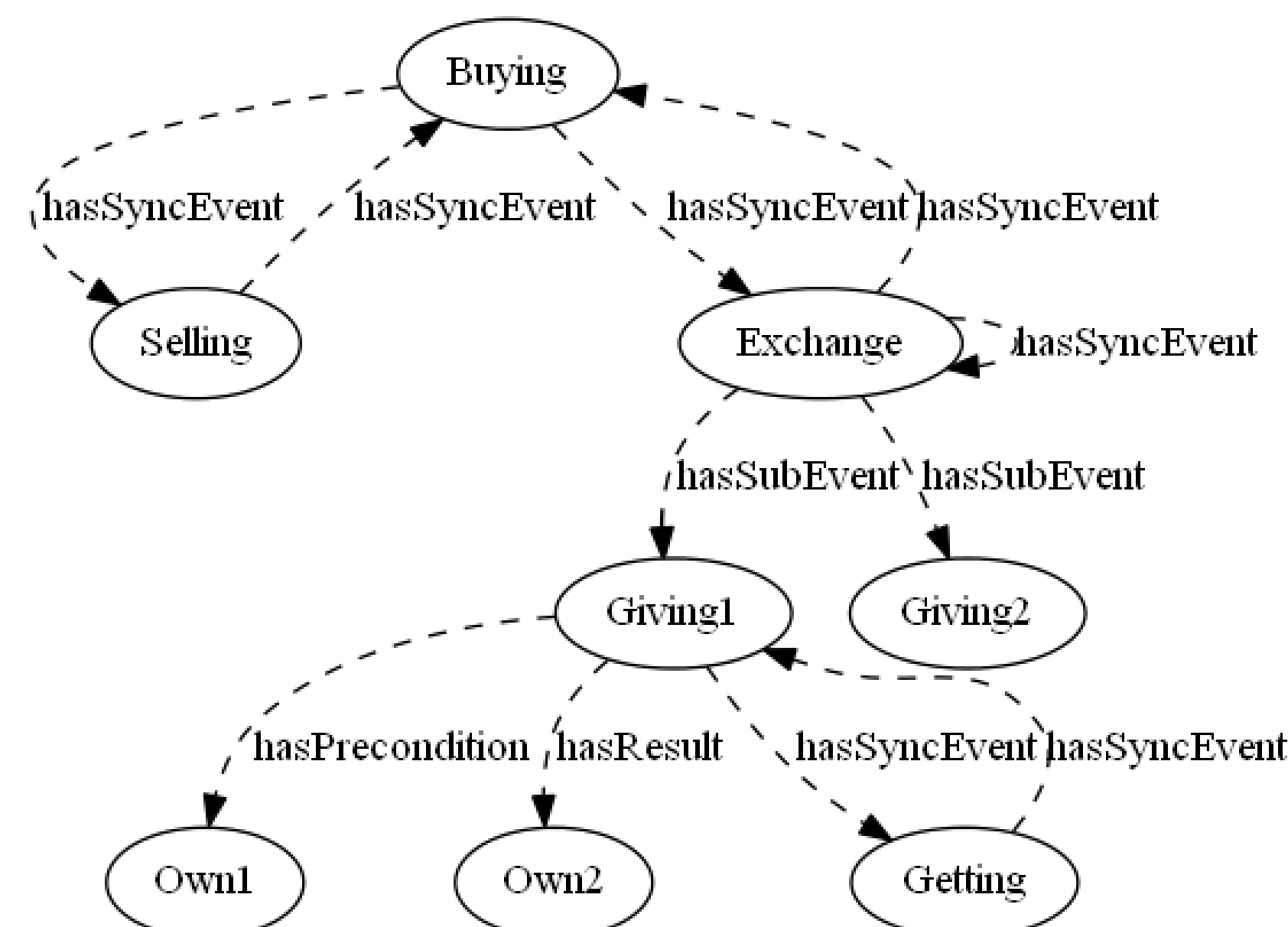
Implicative relations

Our current list of implicative relations are shown in the table below along with their facticity propagation logic in four different contexts:

Main event:	not started	not completed	started	completed
hasFact:	started	started	started	started
hasSyncEvent:	not started	not completed	started	completed
hasSubEvent:	not started			completed
hasPrecondition:			started	started
hasResult:				started
hasPreventedEvent:		completed	not started	not started

The list will likely grow in the future but not tremendously.

Conclusions

It was shown how facticity is handled in SemETAP semantic text analyzer, how it is modelled within the semantic structure and how it is transferred to the inferred propositions though the application of concept definition rules.

A notion of implicative relation was introduced, which reduces the complexity of the rules, simplifies the work of linguists and provides the natural and easy way of thinking in terms of relations instead of analyzing all possible contexts for all parts of the definition.

References

1. Boguslavsky I. et al. 2013. Semantic representation for NL understanding. In: *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue" (2013)*, p.132-144.
2. Boguslavsky I. et al. 2015. Semantic analysis and question answering: a system under development. In: *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue" (2015)*, p.62-79.
3. Boguslavsky I. 2017. Semantic descriptions for a text understanding system. In: *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue" (2017)*, p. 14-28.
4. Karttunen, L. 1971. Implicative verbs. *Language*, 47:340–358.
5. Nirenburg, S., and V. Raskin. 2004. Ontological semantics. The MIT Press. Cambridge, Massachusetts.
6. Rygaev I. 2016. Russian implicative verbs and their semantic analysis in ETAP-3 linguistic processor. *Paper for a student poster session at Dialog-2016 conference* (http://www.dialog-21.ru/media/3480/rygaev.pdf)
7. Rygaev I. 2017. Rule-based reasoning in semantic text analysis. *Proceedings of the Doctoral Consortium, Challenge, Industry Track, Tutorials and Posters @ RuleML+RR 2017* (http://ceur-ws.org/Vol-1875/paper11.pdf)

Further information

Any questions, claims or blames? Feel free to contact me at irygaev@gmail.com.