HOW MUCH DO FORMAL NARRATIVE ANNOTATIONS DIFFER?

A Proppian case study

RENS BOD
Institute for Logic, Language and Computation
Universiteit van Amsterdam

AND

BENEDIKT LÖWE
Institute for Logic, Language and Computation
Universiteit van Amsterdam

AND

SANCHIT SARAF
Institute for Logic, Language and Computation
Universiteit van Amsterdam

Abstract. The formal study of narratives goes back to the Russian structuralist school, paradigmatically represented by the 1928 study *Morphology of the Folktale* by Vladimir Propp. Researchers in the field of computational narratology have developed the general Proppian methodology into various formal and computational frameworks for the analysis, automated understanding and generation of narratives. Methodological issues in this research field give rise to concrete research questions such as "How much does the representation of a narrative in a given formal framework depend on subjective decisions of the formalizer?"' touching philosophy of computing and philosophy of information. In order to approach the mentioned question, we consider the process of formal representation of a narrative as a natural analogue of the task of annotation in computational linguistics and corpus linguistics. We use the Russian folktales formalized by Propp and let them be formalized by annotators according to Propp's system, evaluating these results according to the standards of inter-annotator agreement.

The formal study of narratives goes back to the Russian structuralist school, paradigmatically represented by the 1928 study *Morphology of the Folktale* by Vladimir Propp (1928) in which he identifies seven *dramatis personae* and 31 functions that allow him to formally analyse a corpus of Russian folktales.

Researchers in the field of computational narratology (or "computational models of narrative") have developed the general Proppian methodology into various formal and computational frameworks for the analysis, automated understanding and generation of narratives. Examples for this are Lehnert (1981)'s *Plot Units*, Rumelhart (1980)'s *Story Grammars*, Schank (1982)'s *Thematic Organization Points* (TOPs), Dyer (1983)'s *Thematic Abstraction Units* (TAUs), or Turner (1994)'s *Planning Advice Themes* (PATs). Over the last decades, the main interest of this research community lay in the technical challenges that the computational treatment of narratives brings, but recently, there is again increased interest in the methodological and conceptual issues involved, linking this research closely to questions of the philosophy of information (cf. the paper (Löwe to appear) presented at the *3rd Workshop for the Philosophy of Information*). This interest is witnessed by workshops such as the recent AAAI workshop on *Computational Models of Narrative* that brought researchers from this field together with philosophers, narratologists and professional story tellers. The methodological issues involved give rise to concrete research questions such as

1. How do you compare formal frameworks of narrative? (Cf. Löwe 2010 and Löwe to appear.)
2. How do you assess the quality of a formal framework of narrative?
3. How much does the representation of a narrative in a given formal framework depend on subjective decisions of the formalizer?

Question 1. is a genuinely philosophical question, but also the more technical questions 2. and 3. are very relevant for gaining philosophical insight into what constitutes the formal core of the concept of narrative. In this paper, we approach question 3. of the above list. To this end, we think of the process of formal representation of a narrative in a formal system as a natural analogue of the task of annotation in corpus linguistics and computational linguistics. Whereas typical annotation tasks involve annotation of sentences or discourses (cf., e.g., Marcus et al. 1993, Brants 2000, Passonneau et al. 2006), the formalization or annotation of a narrative is at the next level of complexity, involving sequences or systems of discourses, connected to a narrative. First studies suggest that question 3. is not easy to tackle for the following reasons: First, ambiguity which in typical linguistic annotation is a rather confined phenomenon becomes ubiquitous at the level of narratives: the natural answer to a formalization task is not one annotation, but a family of consistent annotations (cf. Löwe 2010, §2). Secondly, even allowing for multiple annotations, it is not clear whether consensus about whether a given annotation is a valid representation of a narrative is easy to achieve.

Of course, these questions naturally reflect a well-known discussion from computational linguistics: in sentence- or discourse-level annotation, the quality of annotation is typically studied as *inter-annotator agreement* (Carletta et al. 1997, Marcu et al. 1999). For the annotation or formalization of narratives, no such analysis has ever been done, not even with the oldest and best-known formal approach to narrative structure, the Proppian narratemes.

In this study, we use English translations of the Afanas'ev tales formalized by Propp (Afanas'ev 1973), train a group of annotators in the use of Propp's system, and then let them formalize a selection of tales in that formal framework. We evaluate these results according to the standards of inter-annotator agreement from computational and corpus linguistics (Carletta et al. 1997).

References

Afanas'ev, A (1973). *Russian fairy tales*. Pantheon. Translation by Norbert Guterman from the collections of Aleksandr Afanasev. Folkloristic commentary by Roman Jakobson.

Brants, T. (2000). Inter-annotator agreement for a German newspaper corpus. In: *Proceedings Second International Conference on Language Resources and Evaluation LREC-2000*.

Carletta, J.C., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G. & Anderson, A (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13-31.

Dyer, M.G. (1983). *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Articial Intelligence Series. MIT Press.

Lehnert, W.G. (1981). Plot units and narrative summarization. *Cognitive Science*, 4:293-331.

Löwe, B. (2010). Comparing formal frameworks of narrative structures. In M. Finlayson (Ed), *Computational Models of Narrative. Papers from the 2010 AAAI Fall Symposium*, (pp. 45-46). Volume FS-10-04 of AAAI Technical Reports.

Löwe, B. (to appear). Methodological issues in comparing formal frameworks for narratives. In P. Allo & G. Primiero (Eds), *3rd Workshop on the Philosophy of Information*. Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.

Marcu, D., Romera, M. & Amorrortu, E.A. (1999). Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In: *Workshop on Levels of Representation in Discourse*, (pp. 71-78).

Marcus, M.P., Santorini, B. & Marcinkiewicz, M.A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19:302-330.

Passonneau, R., Habash, N. & Ramnow, O. (2006). Inter-annotator agreement on a multilingual semantic annotation task. In: *Proceedings LREC-2006.*

Propp, V. (1928). *Morfologiya skazki*. Leningrad: Akademija.

Rumelhart, D.E. (1980). On evaluating story grammars. *Cognitive Science*, 4:313-316.

Schank, R.C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press.

Turner, S. (1994). *The creative process. A computer model of storytelling*. Lawrence Erlbaum Associates.