

Semantics of 'deep' grammars - its applicability and evaluation

Lars Hellan
NTNU, Trondheim
lars.hellan@hf.ntnu.no

Dorothee Beermann
NTNU, Trondheim
dorothee.beermann@hf.ntnu.no

Abstract

Presenting advances in semantic specification using Minimal Recursion Semantics embedded in an HPSG grammar, the paper addresses the question how the semantic performance of such grammars can be evaluated. We suggest that if a grammar can contribute to correct steering of an artificial agent in a text-based instruction application, then this is as close as one can get to a correctness criterion. However, this is a type of situation we can still only imagine - in this paper, we show examples of semantic representations which stand a chance of providing the explicitness required in a steering scenario, and comment on the prospect of more systematically, and with more coverage, producing such representations using the type of grammar in question.

1 Introduction

An interesting task of natural language processing may be seen to reside in converting text to a format suitable for formal reasoning and instruction of artificial agents; the problem of managing that we may refer to as the 'conversion problem'.

A central task for 'deep' processing grammars is to arrive at measures for comparison and evaluation; the problem of defining such measures we may refer to as the 'evaluation problem'.

We here point to a respect in which these problems may interrelate.

Consider first possible measures for evaluating 'deep' computational grammars. Traditional evaluation procedures drawing on quantifiable

measures such as parse coverage clearly have limited (but not no) relevance. At this level, a parse accompanied by a mistaken semantics is of questionable value; however there is no mechanical procedure for determining correctness of semantic analysis. All one can do is consider proposed analyses in an academically defensible fashion.

Even if conversion correctness could be mechanically decided, a problem is for what it is needed. One possible 'client' could be communication with artificial agents. If the agents are steered satisfactorily using a language processing engine as one component, then that counts as an adequacy proof for that component. If the linguistic complexity of the instructions is so high that the processing could not be successful without a deep grammar conversion system, then once such a system works, that even more proves its adequacy. However, this is now a proof which pertains to actual interaction with agents, and not residing in the evaluation of semantic representations as such.

As consistency checkers for abstract grammars, computational grammars have one clear type of motivation. For non-grammarians, their potential usage as language processing tools in wider applications is perhaps their more interesting aspect. Moreover, when it comes to meaning representations as produced by grammars, it may well be claimed that their adequacy can only be tested in actual behavior, just as a language user's grasp of meaning can only be assessed through communicative consistency (to activate a Wittgensteinian point). These concerns point to the interfacing of grammars with actual agents as crucial to both purpose and correctness of the grammars.

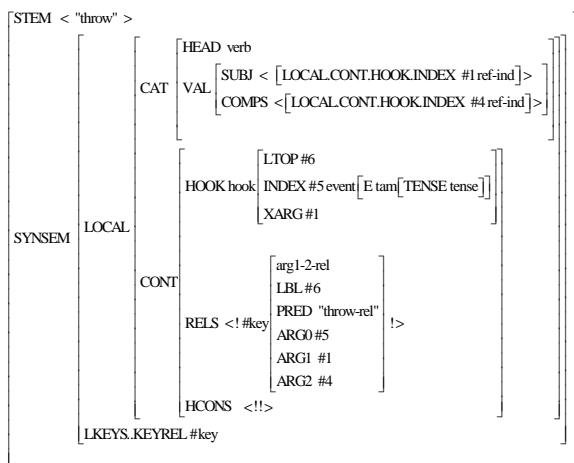
In the following, we will exemplify semantic representations as produced by a 'deep' grammar of choice (section 3). Since it is essential to the enterprise that the mechanisms producing such representations be hosted in linguistically reliable grammars, we will comment on the 'cost'

incurred by a grammar to produce the amount of semantic detail in question (section 4). We then (section 5) sketch the possibility of scenarios where the semantics demonstrated could be put into actual use, and comment on how this then could also be an 'evaluation' tool.

2 The Background – Multilingual Grammar Engineering using HPSG based Tools

A base-line requirement on semantic representation in grammar implementations within LFG and HPSG is that the composition of elementary predications and coreference be represented adequately. *Minimal Recursion Semantics* (MRS; cf. Copestake et al. 2005) is a notational format, corresponding roughly to predicate logic, where this requirement is fulfilled. HPSG grammars using the Linguistic Knowledge Builder (LKB; Copestake 2002) standardly have this format built in - cf. the DELPH-IN consortium of grammar development (<http://www.delph-in.net/>). The interaction between grammatical specification and MRS representations can be partly seen from (1) and (2) below, where the lexical specification of the verb *throw*, represented in (1), is reflected in the MRS produced for the sentence *The boy throws the ball* ((2)).

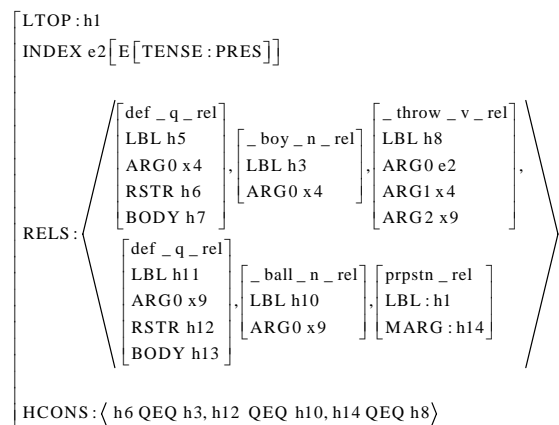
(1) Feature structure for the verb *throw*:



In accordance with a standard MRS set up as illustrated in (2), for any constituent *C* (of any rank), the RELS list in its CONT specification is a 'bag' of those elementary predications (EPs) that are expressed inside *C*. The sentence *The boy throws the ball* displays six EPs, of which one

reflects *throw*; cf. (2). The subject and object arguments of the verb are reflected by the coindexation of the ARG1 and ARG2 of the verb with the ARG0s of the determiner and the noun, corresponding to the 'bound variables'. The remaining EP represents the 'message type'. Scope properties are expressed in the HCONS list, 'x QEQ y' meaning essentially that x scopes over y. HCONS thus records the scopal tree of the constituent in question, as outlined in Copestake et al. (2005).

(2) MRS for the sentence *The boy throws the ball*:



3 Fine-grained conversion

To be presented here is a design that enriches MRS representations by lexical semantic information. By the demands of a conversion tool, such representations should also be 'calculation-sharp'; for example, when it comes to expressions of measurement, the representations delivered should provide a break-down of exactly what is being measured and which values are being assigned. For instance, for *This building is 5 meters higher than the church*, a calculation-sharp analysis should provide a representation corresponding to the quasi-paraphrase:

(3) "For a degree *d1* and a degree *d2* such that *d1* is the height of the building and *d2* is the height of the church, *d1* exceeds *d2* to an extent *d3*, and 5 meters measures out *d3*."

(3) stands in contrast to, for instance, a representation such as:
 "5-meters-taller-than (the building, the church)"
 where the naming of a function suggests a meaning rather than providing it.

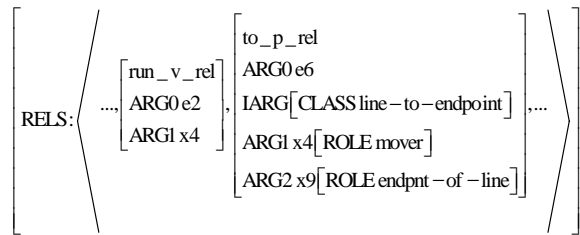
Similarly, a semantic representation should be able to distinguish directional from locative uses of prepositions, so as to capture the ambiguity of, e.g., *jump in the car*. Within the directional reading we need to be able to account for different thematic constraints on subjects of movement verbs, such as ‘mover’ and ‘line’ concepts to account for the fact that *lead* is a synonym of *go* in *The way leads along the ridge* but not in *The guide leads along the ridge*. Together with calculation sharpness, fine-grainedness will decide if one is able to represent, e.g., for a sentence such as *Walk along the ridge 5 kilometers towards the south*, a ‘mover’ who, in one and the same action, walks along a ridge *and* changes her position to a place whose degree of ‘toward-south’ exceeds the position held at the beginning of the action by 5 kilometers. Given that one is able to make these necessary distinctions, one is then also able to represent the distinct meaning of *Walk along the ridge and then 5 kilometers further south*, that is applying the above made descriptions to *time consecutive* actions.

In the following we are going to show how some of the above mentioned constraints can be stated within the formal frame of Minimal Recursion Semantics, as part of the HPSG grammar *NorSource* (<http://www.delph-in.net/>), which goes slightly further in fine-grainedness of semantic analysis than other grammars in this group. Since all specifications to be discussed in the following are clearly within the scope of a linguistic semantics (and not an aspect of physical or other analysis), they should be recognized as candidates for analysis by any linguistically realistic language processing tool, and in particular by a deep processing grammar.

3.1 Spatial semantics

The two representations given below in (4) and (5) show partial semantic descriptions for the directional construction *He runs to the forest* and the comparative expression *The building is 5 meters higher than the church*, respectively (generated for the Norwegian counterparts, but for convenience represented with English). We first comment on (4):

- (4) Partial MRS representation of *He runs to the forest*



Illustrated here is that spatial relations are typed according to the conceptual type of the arguments they select. Let us start with the specifications that are embedded under the attribute IARG (for ‘internal argument’). *To* is specified as a preposition of type *line-to-endpoint* where *line* and *endpoint* stand for concepts imposed on the two arguments of the preposition. Specifications such as *mover* and *endpoint* live in a type system where the top level distinction is that between *line* and object of any dimension, called *xdim*. The *line* concept is imposed on the first argument of the spatial relation, in (4), via the motion verb, on the subject playing the role as projecting a line through movement. The *endpoint* concept is satisfied by the second argument, that is the object of the preposition. (Cf. Hellan and Beermann 2005, forthcoming) for presentations of the system.

(4) illustrates how a standard MRS semantic output can be enriched by a system of spatio-temporal types, providing a system that distinguishes, e.g., lines and spans from objects. For applications that use MRS representations as input (e.g., in information extraction), the design offers potential advantages in two respects. For one thing, to the extent that semantic information is channelled through a finite set of defined features, search can be principally restricted relative to this set, rather than oriented towards the potentially more unrestricted array of predicates like *run_v1_rel* and *to_p2_rel*. A test demo along these lines is described in Beermann et al. (2004), where mountain hiking routes are rendered in pictographic form via RMRS grammar outputs and XML conversion.

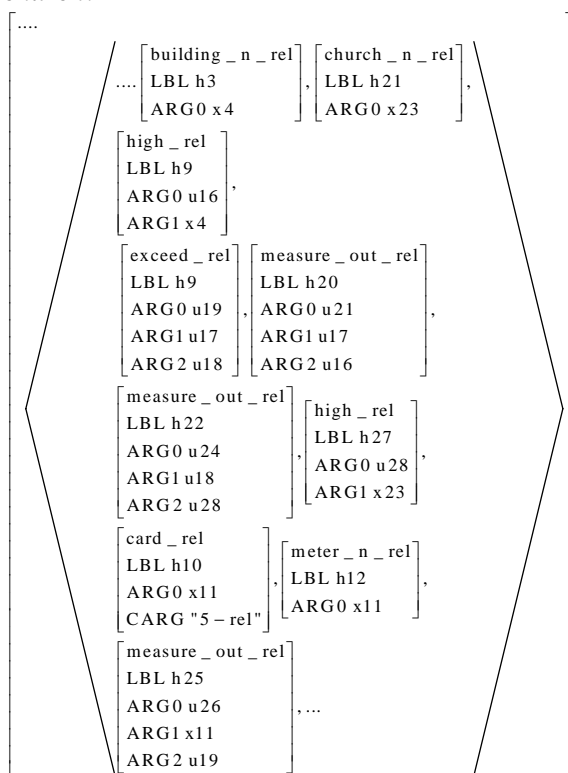
Secondly, to the extent that the system underlying the annotations exemplified in (4) is cross-linguistically applicable, it opens the possibility for grammars of different languages to produce MRSEs with identical annotations for sentences with essentially the same content in relevant respects. This has potential benefits for machine translation, and also for information extraction: Once a given extraction algorithm has been defined for MRSEs of this kind for one language, the same algorithm may then in principle be ap-

plicable for the corresponding MRSes for a different languages.

3.2 'Calculation sharp' representations

In (5) below is a representation of comparison where the dimension of comparison is height, viz. the height of the building vs. the height of the church, explicating the paraphrase suggested in (1) for the sentence *This building is 5 meters higher than the church* (as noted, we here use English counterparts of the structures created by the Norwegian grammar):

(5) (Part of) MRS representation for the sentence *This building is 5 meters higher than the church*:



The first two EPs here introduce $x4$ and $x23$ as variables representing the building and the church, respectively. Next, $u16$ is introduced as a variable representing $x4$'s height, and further down, $u28$ as representing $x23$'s height. The EPs with predicate value *measure_out_rel* assign the measure values $u17$ and $u18$ to these respective heights, and the EP *exceed_rel* states that $u17$ exceeds $u18$, i.e., that the building is higher than the church. The ARG0 of this EP - $u19$ - in turn represents the extent to which this difference holds, and this value is associated with the measure $x11$ (last EP), with the EPs preceding stating that this is meter in a number of five. Of these

EPs, those for *measure_out_rel* and *exceed_rel* clearly do not stand in a one-to-one relation to occurring words or morphemes: the last *measure_out_rel* reflects a syntactic specifier-head constellation between *five metres* and *higher*, and the other two plus the *exceed_rel* EP emerge from the comparative morpheme *-er*. The analysis mirrors proposals made in formal semantics, and indicates down to which detail insights from this field can also be implemented in a computational tool.

Potential advantages of such a degree of preciseness lie not only in the sheer adequacy of the representations, but also in the way such MRSes, suitably rewritten via XML conversion or other means, may be exchanged into calculable descriptions or instructions to natural or artificial agents. For such an endeavor to be feasible, the domain will probably have to be quite restricted, and with a highly predictable vocabulary.

4 Hosting fine-grained conversion in a medium-to-large scale grammar

The semantics illustrated above is computed by a core grammar of not inconsiderable size (appx. 500 phrasal types, 950 word and lexeme types), without compromising general efficiency. In order to fully link the prepositional typology to information encoded by verbs and nouns, and in the latter case real-world knowledge concerning their inherent properties, one most likely must resort to additional independent sources of information, such as ontologies.

A grammar producing fine-grained and precise semantics of the kind mentioned requires a high degree of what one may call *construction control*: MRSes should occur exclusively with those constructions whose meaning they reflect, and parsing ambiguity should be kept to a minimum to not 'conceal' correct MRSes among a multitude of generated inappropriate MRSes. One thus needs an architecture allowing a close form - meaning correspondence. In this respect, the over-all architecture of HPSG grammars is well suited, in its parallel levels of representation; for more detailed expositions of the mechanisms relevant, cf. Beermann and Hellan (forthcoming), and Hellan (forthcoming).

5 Test scenarios?

Imagine a robot, for instance under-water, which one wants to direct using linguistic input (be it

speech or text). Given an in-built mechanism of measuring distance travelled, and the ability to carry out 'exceed' predications as summation of entities specified, it ought to be possible to enable such a robot to execute instructions like *move twenty metres more forward*, based on a transmission channel including a conversion mechanism of the type of granularity as was exemplified above. Directions like 'forward', 'to the right' etc. are agent relative, and thus definable by an internal mechanism. Through training into recognition of salient objects, such as pipes and valves, expressions for end-of-path would also seem possible, as in *go forward to the pipe*.

Handling of objects, as in "turn the wheel to the right", would, apart from the further physical components involved, also involve understanding of concepts like 'grasp' and 'hold', and dynamic manipulation of whatever is held; for the conversion mechanism at hand, sufficient richness of expression will again seem available.

As mentioned, in the grammar given, the specifications which would be crucial to the enabling of the actions mentioned are stated in language independent ('interlingua') terms. Moreover, the modules of the grammar producing the specifications are designed as independent possible add-on's to grammars of other languages as well. This will indeed seem a prerequisite for the viability of such a design, and within consortia like the DELPH-IN consortium mentioned, with its high degree of shared design between grammars, it is a practically feasible design.

Once situation types like the one indicated can be made operational, it would seem an interesting project to design a standard repertoire of scenario-types as test-scenarios across grammars, and also across types of semantic representation frameworks. In that way, a test bed of analytic adequacy across semantic formalisms might become available.

6 Conclusion

This paper has given some indications how one can provide fine-grained and 'calculation-sharp' semantic representations produced by a computational grammar using the MRS formalism. If we succeed in this, we may be able to make such grammars useful in applications directed towards instructing robots or otherwise manipulate scenarios in real or virtual scenarios. If we manage that, we may also have an interesting evaluation environment for deep processing grammars.

References

- Beermann, D., J.A. Gulla, L. Hellan and A. Prange (2004) Extraction of spatial Information using a Deep Processing Grammar. Paper presented at CLIN 2004, Leiden; to appear in proceedings.
- Beermann, D. and L. Hellan (forthcoming) The preposition *til* ('to') in Norwegian: accounting for its full range of uses in an implemented HPSG grammar.
- Copestake, A. (2002) *Implementing Typed Feature Structure Grammars*. Stanford: CSLI Publications.
- Copestake, A., D. Flickinger, I.A. Sag and C. Pollard. (2005) Minimal Recursion Semantics: an Introduction. *Journal of Research on Language and Computation*. 281-332.
- Hellan, L. (forthcoming) Syntax and Semantics of Comparatives in an Implemented HPSG grammar.
- Hellan, L. and D. Beermann (2005). Classification of Prepositional Senses for Deep Grammar Applications. In Kordoni, V. and A. Villavicencio (eds) *Proceedings of the Second ACL-SIGSEM Workshop on The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*. University of Essex.
- Hellan, L. and D. Beermann (forthcoming). Prepositions in Minimal Recursion Semantics. In Lønning, J.T., S. Oepen, H. Dyvik and D. Flickinger (eds) *The LOGON Machine Translation Effort*.