Chapter 9

Combining Discourse Representation Theory with FrameNet

Johan Bos, Malvina Nissim

1. Introduction

Thematic (or semantic) roles have been the topic of debate in lexical semantics since the 1960s, in the quest for finding semantic generalisations (see e.g. Fillmore 1968, Jackendoff 1972). In formal semantics, however, research mainly focuses on compositional rather than lexical semantics, and the issue of thematic roles has largely been neglected. But there are good reasons to combine a theory of thematic roles with a formal theory of natural language meaning, both from a practical and theoretical perspective:

- Thematic roles generalise and yield more abstract formal representations which will benefit the reasoning capabilities of any given semantic formalism;
- Mismatches that arise during the fusion can be helpful feedback to improve either theory;
- Performance of semantic role labelling algorithms can potentially be improved by using a semantically informed tagger.

Apart from these general arguments, it is also fair to say that a fusion between formal and lexical semantics is timelier than ever in the history of natural language processing. On the one hand, wide-coverage semantic parsers exist that produce reasonably adequate semantic representations for open-domain texts (Bos 2005). On the other hand, semantic role labelling has been recognised as a promising way to capture semantic generalisations in practical semantic applications.

In this article we investigate both theoretical and practical implications of integrating a theory of semantic roles, FrameNet (Baker, Fillmore and Lowe 1998), with a formal semantic theory, Discourse Representation Theory (Kamp and Reyle 1993). But why do we choose FrameNet, and why do we choose DRT?

We choose FrameNet because we like its bottom-up approach to building up an inventory of roles, thereby avoiding many practical problems in designing an adequate set of theoretically motivated roles. Besides, FrameNet is blessed with a large set of annotated examples, which we will be using to implement an automated role labelling tool.

We choose DRT because it has established itself as a well-documented formal theory of meaning, covering a number of semantic phenomena ranging from pronouns, abstract anaphora, presupposition, tense & aspect, propositional attitudes, to plurals (Kamp and Reyle 1993, Asher 1993, Van der Sandt 1992). Additionally, there is a strong practical argument for choosing DRT: we have at our disposal a robust parser that produces Discourse Representation Structures, the meaning representations proposed by DRT.

In this article, we aim to make a theoretical contribution by discussing the process of matching DRT and FrameNet. We will also discuss practical considerations by proposing a method for augmenting a DRT-based formalism with FrameNet roles. This implementation can both be used to produce richer semantic representations and for semantic role labelling.

2. Background

2.1. Event Semantics and Thematic Roles

Davidson (1967) suggested that sentences quantify over events and that adverbs and adjuncts essentially can be seen as modifiers of events, at least from a semantic perspective. This is attractive from a modelling perspective, for the meaning of sentences with events can be captured by a simple first-order language, rather than a more complicated language based on higher-order logic. In Davidson's view, "events introduce individuals", as is shown by the following example and its translation:

(1) Betty quietly left her house yesterday.
(1a) $\exists$e(leave(e,b,her-house) $\wedge$ quiet(e) $\wedge$ yesterday(e))

Parsons proposed a modification of Davidson's analysis (Parsons 1980). He stipulated that not only modifiers but also the arguments of verbs should be seen viewed as predicates of events. The resulting framework is what is usually referred to as the **neo-Davidsonian** system (Dowty 1989). It has two positive consequences: 1) we can uniformly introduce **thematic roles** (such as *agent*, *patient*, *theme*) as relations between events and individuals; 2) verbs with optional arguments can be covered with a simpler signature of predicates. Consider again example (1), but now with its translation in neo-Davidsonian style:

(1) Betty quietly left her house yesterday.
(1b) $\exists$e(leave(e) $\wedge$ agent(e,b) $\wedge$ theme(e,her-house) $\wedge$ quiet(e) $\wedge$
yesterday(e))

As Dowty notes, the neo-Davidsonian system has some attractive formal properties (Dowty 1989). There is only one way to state that an individual is participating in an event—namely by relating it to the event using a binary relation expressing some thematic role. Furthermore, the approach clearly distinguishes the participants of an event by the semantic roles they bear. Finally, it also allows us to characterize the meaning of thematic roles independently of the meaning of the verb that describes the event.

Table 9.1: Comparing Davidsonian with neo-Davidsonian style semantic representations

	Davidsonian	**neo-Davidsonian**
Betty left.	$\exists e(\text{leave}_1(e,b))$	$\exists e(\text{leave}(e) \wedge \text{agent}(e,b))$
Betty left her house.	$\exists e(\text{leave}_2(e,b,\text{her-house}))$	$\exists e(\text{leave}(e) \wedge \text{agent}(e,b) \wedge \text{theme}(e,\text{her-house}))$

2.2. Linking Theory

Even with this formal apparatus at our disposal, it is not obvious how grammatical objects are linked to semantic roles. Unfortunately, there is no straightforward one-to-one correspondence between the grammatical function of linguistic units and the thematic roles in which they participate. The standard example illustrating this problem is the verb "to open" used in various contexts:

(2) Betty opened the door. (subject $\mapsto$ agent; object $\mapsto$ theme)
(3) The key opened the door. (subject $\mapsto$ instrument; object $\mapsto$ theme)
(4) The door opened. (subject $\mapsto$ theme)

In (2) the subject of "opened" corresponds to the semantic role of agent, whereas in (3) and (4) it corresponds to instrument and theme, respectively. In (2) and (3) the semantic role of theme is introduced by the direct object of "opened", but in (4) it is introduced by the subject. Examples like these make clear that any framework of semantic roles has to say something about matching grammatical function to semantic roles, a component usually referred to as **linking theory**. Various systems have been proposed for an adequate linking theory, and the reader is kindly referred to Dowty's work for an overview (Dowty 1989). Most of these early attemps were based on rule-based systems. In the context of this article, we will use a memory-based learning system to fill the gap of linking theory.

2.3. Frame Semantics

But what is the inventory of roles we are going to use? Ideally, one would like to have a small inventory of abstract roles covering all possible situations at one's disposal. However, given the immense variety of possible situations,

Frame A: $_{Speaker}$[Ben] $_{\textbf{Statement}}$[had mentioned] $_{Message}$[his withdrawal].
Frame B: $_{Theme}$[Ben] had mentioned $_{\textbf{Departing}}$[his withdrawal].

Figure 9.1: Annotated FrameNet example for the frames Statement (A) and Departing (B).

this has simply proven impossible, and many early attempts at formalizing semantic roles shipwrecked on exactly this issue. FrameNet, the system we are going to use in the context of this article, overcomes this problem by following a bottom-up approach.

Fillmore introduced the notion of **frames** as an abstract, conceptual model of the world, where frames describe types of situations, events, and actions, and their participants. The participants (the **frame elements**) play the various **roles** specifically associated with a certain frame: in FrameNet, each frame comes with its own set of frame elements.

Frames are invoked by lexical items (typically verbs, but also by other parts of speech such as nouns) which are referred to as **lexical units** or **targets** in FrameNet semantics. Ambiguous words can be lexical units of different frames, where the chosen frame disambiguates the sense of the word. (An example of an annotated sentence taken from the FrameNet corpus is shown in Figure 9.1).

Well-known examples of instances of roles found in the literature are agent, patient, theme, recipient, instrument, and so on, as demonstrated in the examples above. In FrameNet, some of the thematic roles are general, occurring in various frames, whereas others are specific to certain frames. As a consequence, there is no practical limit to the number of different roles used in FrameNet's inventory, as it is unclear how many different frames one should consider.

The current version of FrameNet contains annotated data of more than 800 different frames, exemplified in more than 100,000 example sentences. Semantic roles correspond to **frame elements** or **frame roles** in FrameNet's terminology, and words or phrases that evoke frames are referred to as **lexical units** or **targets**. There is a distinction between core and non-core frame elements, where the former are normally part of a frame, and the latter are optional.

2.4. Discourse Representation Theory

DRT is a formal semantic theory originally designed by Kamp to cope with anaphoric pronouns and temporal relations (Kamp 1981). DRT uses an explicit intermediate semantic representation, called DRS (Discourse Representation Structure), for dealing with anaphoric or other contextually sensitive linguistic phenomena such as ellipsis and presupposition. DRSs are structures comprising two parts: 1) a set of discourse referents; and 2) a set of conditions constraining the interpretation of the discourse referents. Conditions can be simple properties of discourse referents, express relations between them, or be complex, introducing (recursively) subordinated DRSs. Anaphoric links are represented by equality conditions.

The standard version of DRT formulated in Kamp and Reyle incorporates a Davidsonian event semantics (Kamp and Reyle 1993), where discourse referents can also stand for events and referred to by anaphoric expressions or constrained by temporal relations. An example of a Kamp and Reyle-style DRS is shown on the left-hand side in Figure 9.2.

3. Combining DRT with FrameNet

3.1. Basic Idea

First of all we will deviate from Kamp and Reyle and employ a DRS language that adopts a neo-Davidsonian view. This will make it easier to incorporate FrameNet roles as two-place relations between discourse relations, since semantically, thematic roles denote relations between entities.

In view of the FrameNet terminology, the expressive power of the DRS language faces the following requirements:

- FrameNet targets (i.e., lexical units) correspond in DRT to entities that introduce fresh discourse referents;
- FrameNet roles (i.e., frame elements) correspond in DRT to two-place relations, of which one denotes the entity introduced by the target, and the other the entity introduced by a frame element.

In FrameNet, targets are expressed by verbs, nouns, or adverbs. In DRT, only finite verbs and noun phrases introduce discourse referents. Adverbs, however, do not introduce discourse referents. As a consequence, frames or roles anchored by adverbs cannot be represented in the DRS language. This affects however the minority of cases and is not a major concern here. Most FrameNet roles correspond to verbal arguments, prepositions, noun-noun compound relations, or possessives. In a neo-Davidsonian version of DRT, these are all expressed by two-place relations. These preliminary observations are summed up in Figure 9.2, exemplifying how to move from a Davidsonian DRS à la Kamp and Reyle to a neo-Davidsonian DRS with FrameNet roles.

e x y z
named(x,ben)
mention(e,x,y)
male(z)
z=x
withdrawal(y)
of(y,z)

e x y z
named(x,ben)
Statement(e)
Speaker(e,x)
Message(e,y)
male(z)
z=x
Departing(y)
Theme(y,z)

Figure 9.2: Example DRSs for the sentence "Ben had mentioned his withdrawal". The standard DRS following Kamp and Reyle (left), and the DRS augmented with FrameNet roles (right). Differences between the two DRSs are marked in bold-face.

3.2. Theoretical Discrepancies

Several issues came to the surface when trying to combine FrameNet with DRT. Recall that the main assumption we made is that roles are modelled as two-place relations between entities. This presupposes that whatever is viewed as a target in FrameNet should introduce a discourse referent in the corresponding DRS. In DRT, rougly speaking, only noun phrases and finite verbs introduce discourse referents. In FrameNet however, targets can als be introduced by adverbs. This forms a first theoretical clash between FrameNet and DRT, and it has consequences both for targets and roles triggered by adverbs.

We will illustrate this mismatch with the example "Bill drove quickly". In FrameNet, this sentence selects the frame **Operate_vehicle**, with two roles, a core role **Driver** and a non-core role **Speed.** In DRT, only the first role can be expressed in a DRS (namely as a relation between Bill and the driving event). The second role cannot be expressed simply because the adverb quickly does not introduce a discourse referent. In other words it is the limited expressive power of standard DRT that stops us from representing the speed role.

It is beyond the scope of this paper to extend the DRS language in order to cope with this mismatch between DRT and FrameNet. In Figure 9.3 is a sketch of a possible solution. This solution requires extensive lexical knowledge of adverbs.

e x
named(x,bill) drive(e,x) quick(e)

e x
named(x,bill) **Operate_vehicle**(e) **Driver**(e,x) quick(e)

e x y
named(x,bill) **Operate_vehicle**(e) **Driver**(e,x) quick(y) **Speed**(e,y)

Figure 9.3: Example DRSs for the sentence "Bill drove quickly", in a Davidsonian notation (left), one in neo-Davidsonian with one FrameNet role (middle), and one with a possible extension in lexical representation express both FrameNet roles (right).

Secondly, mismatches occur in cases were FrameNet assumes a role which is expressed on the DRT only indirectly via two two-place relations. The main case here is introduced by the possessive construction, which normally, for three entities x, y and z, yields the two relations R1(x,y) and R2(y,z) in a DRS, whereas the FrameNet role is expressed as a relation between x and z. Although we classify this discrepancy between FrameNet and DRT as another mismatch, it is one that is relatively easy to deal with, by simply adding rather than substituting a two-place relation.

4. Automatically Learning FrameNet Roles

4.1. Method and Data

The method that we employ to enrich DRSs with FrameNet roles is fully automatic and implemented as part of an extension of a wide-coverage parser for English. This is a robust state-of-the-art parser based on CCG, Combinatory Categorial Grammar (Clark and Curran 2004), and can be used in a pipeline with Boxer to produce DRSs. Boxer already generates DRSs in a neo-Davidsonian style—however it is restricted to a limited set of roles, which we will refer to as **proto roles** below.

As starting point we took the lexical units and the example sentences that come with FrameNet and are marked up in XML. In our study we only considered those targets annotated with parts of speech corresponding to verbs. We translated the XML files into a format accepted by the CCG parser, preserving the annotation of roles and target. Then we ran Boxer to generate DRSs for the example sentences, adding the FrameNet target and role information to the DRS.

The result of this process was a collection of DRSs with (partially) labelled FrameNet roles and targets. This data was used as training and testing material for building and evaluating a classifier that assigns FrameNet roles. We split the automatically annotated DRSs into three sets (Table 9.2: 10% was used for developing and tune our system (development data); 80% to build the model (training data), and 10% to evaluate the final system (test data).

4.2. Class Labels and Features

The idea behind statistical classification can be summarised as follows: given a set of training objects (a list of feature values) associated to the desired output (a class label), a model is created which can predict the class label of a new unseen object (test instance). This training process is referred to as "supervised learning" since the model is learnt over manually annotated data. For the task

Table 9.2: Distribution of training, development and test data

Set	Instances	Different frames	Different roles
Training	76,312	294	481
Development	9,534	279	362
Test	10,683	280	373
Total	96,529	294	487

we have in mind, we would need two classifiers: one that assigns frames and one that, given a frame, labels frame elements. The system we describe here is a classifier of the second type, and assumes that the frame is already known. For our experiments we use TimBL, a memory-based learner (Daelemans and Van den Bosch 2005).

There are two obvious ways to build a model for assigning frame elements: (a) training per frame, that is having separate training sets (and thus models) for each frame, or (b) training on the entire data set using the frame just as a feature. The first approach has the advantage of explicitly limiting the set of class labels to the set of roles that belong to the frame, but the amount of training data for many frames is too little to obtain reliable results. The second approach incurs the risk of labelling entities with roles that are not associated with the assigned frame, because the class labels comprise all frame elements occurring in the training data. However, since many general roles (such as "agent" and "theme") occur across frames, this method does not suffer as much from the sparseness problem. These considerations naturally led us to opt for the second approach.

Selecting appropriate features is an important ingredient for a successful classifier. We selected seven features to build our model, as Table 9.3 shows. Feature 1 is the frame of which we need to label its frame elements; Feature 2 is the morphological root of the lexical unit that evokes the frame; Feature 3 is the proto role assigned in the DRS as output by Boxer; Feature 4 and 5 are boolean values denoting whether the frame element has adjectival or cardinal modification; Feature 6 is the WordNet top hypernym (out of 25) of a frame element; Feature 7 is the named entity type assigned by the parser.

Table 9.3: Overview of features used for the semantic role classifier

No	Feature	Value
1	frame name	{. . . }
2	lexical unit	{. . . }
3	proto role	{agent, patient, theme, . . . }
4	adjectival modification	{yes, no}
5	cardinal modification	{yes, no}
6	wordnet hypernym	{. . . }
7	named entity type	{per, loc, tim, org, nil}

4.3. Results and Discussion

The classifier achieved an accuracy of 0.839 on the development data, assigning 8000 out of 9534 roles correctly, and an accuracy of 0.921 on the test data, with 9844 correct assignments out of 10683. We think this is a promising result compared to other role-labelling algorithms (Gildea and Jurafsky 2002). To interpret these accuracy results we also analysed the influence of the amount of training instances and the contribution of each feature.

In order to assess the influence of the amount of training data on the accuracy or role labelling, we ran the classifier with increasing portions of training data. Each run was done five times with five different randomly chosen portions of data. Results on the development data are averaged over the five runs and visualised in Figure 9.4. This graph predicts that accuracy can still improve when given more training material. This is not surprising. There are nearly five hundred class labels (see Table 9.2) mainly because the majority of roles are frame-specific. Therefore, only little training data is available for many of these class labels.

The contribution of each of the seven features was computed by running seven leave-one-out classifiers. The obtained results on the development data are shown in Table 9.4. The feature with the most impact is the frame: without it, performance drops dramatically from 84% to 57%. The second most important feature is the proto role as produced by Boxer. If not included

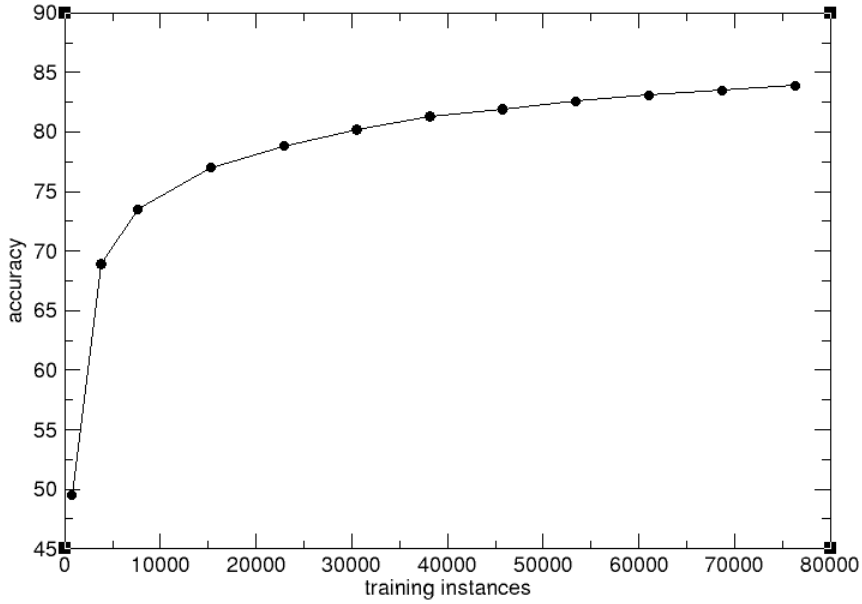

Figure 9.4: Influence of training size with respect to accuracy of role labelling

in the feature set, it causes a loss of 7 percentage points of accuracy. It is interesting to observe that leaving out Feature 2 (the lexical unit) gives a small boost in performance. All other features do not seem to make a significant contribution.

Table 9.4: Contribution of features, ordered by impact on accuracy

Feature left out	Correct/Total	Accuracy
3	5450/9534	0.571638
1	7310/9534	0.766730
6	7902/9534	0.828823
7	7925/9534	0.831236
4	7955/9534	0.834382
5	8005/9534	0.839627
2	8087/9534	0.848227

5. Conclusion

We have argued that a neo-Davidsonian approach to representing events fits best for combining FrameNet roles in a DRT-based semantic formalism. We also demonstrated that the majority of roles, as annotated in the FrameNet corpus, can be included in Discourse Representation Structures. However, due to the way adverbial modification is handled in DRT, FrameNet roles that are realised via adverbs pose a representational problem.

The experiments that we have described show that automatically learning FrameNet roles using semantic features is promising, achieving an accuracy of 92.1%. This is an upper bound, since we assume that the frame is already known. The automatic identification of frames will necessarily yield a drop in performance, as the lack of frame information causes a drop in a accuracy of seven percentage points, as shown by the experiments on the development data. The most important information for assigning frame elements comes from the proto roles, thus implying that semantic features have added value.

Finally, because many of the roles are frame specific, we can expect accuracy of role labelling to rise if more training data becomes available.

As far as we are aware this is a first attempt of automatically combining deep semantic representations with semantic role information. Semantic role labelling is, of course, not an aim in itself, and a next step would be to examine how DRSs with FrameNet roles would affect performance of automatic question answering or recognising textual entailment.

References

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer Academic Publishers.

Baker, C. F., Fillmore, C. J. and Lowe, J. B. (1998). "The Berkeley FrameNet project". In *Proc. of 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, Montreal, Canada.

Bos, J. (2005). "Towards wide-coverage semantic interpretation". In *Proc. of Sixth International Workshop on Computational Semantics IWCS-6*, Tilburg, Netherlands, pp. 42–53.

Clark, S. and Curran, J. R. (2004). "Parsing the WSJ using CCG and Log-Linear Models". In *Proc. of 42nd Annual Meeting of the Association for Computational Linguistics (ACL '04)*, Barcelona, Spain.

Daelemans, W. and Van den Bosch, A. (2005). *Memory-Based Language Processing*. Cambridge University Press.

Dowty, D. (1989). "On the semantic content of the notion thematic role". In *Properties, Types, and Meanings*, Kluwer.

Fillmore, C. J. (1968). "The Case for Case". In Harm, E. B. R. T. (Ed.), *Universals of Linguistic Theory*, New York: Holt, Rinehart, and Winston, pp. 1-90.

Gildea, D. and Jurafsky, D. (2002). "Automatic labeling of semantic roles". *Computational Linguistics*, 28 (3), pp. 245–288.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Kamp, H. (1981). "A Theory of Truth and Semantic Representation". In Groenendijk, J. and Janssen, T. M. V. and Stokhof, M. (Eds.), *Formal Methods in the Study of Language*, Amsterdam: Mathematical Centre, pp. 277–322.

Kamp, H and Reyle, U. (1993). *From Discourse to Logic; An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and DRT*. Dordrecht: Kluwer.

Parsons, T. (1980). "Modifiers and quantifiers in natural language". *Canadian Journal of Philosophy*, 6, pp. 29–60.

Van der Sandt, R. A. (1992). "Presupposition Projection as Anaphora Resolution". *Journal of Semantics*, 9, pp. 333–377.