Machine Learning for Interpretation of Spatial Natural Language in terms of QSR

Parisa Kordjamshidi[1], Joana Hois[2],
Martijn van Otterlo[1], and Marie-Francine Moens[1]

[1] Katholieke Universiteit Leuven, Departement Computerwetenschappen,
{parisa.kordjamshidi,martijn.vanotterlo,sien.moens}@cs.kuleuven.be
[2] University of Bremen, Faculty of Computer Science,
joana@informatik.uni-bremen.de

Abstract. Computational approaches in spatial language understanding distinguish and use different aspects of spatial and contextual information. These aspects comprise linguistic grammatical features, qualitative formal representations, and situational context-aware data. We apply formal models and machine learning techniques to map spatial semantics in natural language to qualitative spatial representations. In particular, we investigate whether and how well linguistic features can be classified, automatically extracted, and mapped to region-based qualitative relations using corpus-based learning. We structure the problem of spatial language understanding into two parts: i) extracting parts of linguistic utterances carrying spatial information, and ii) mapping the results of the first task to formal spatial calculi. In this paper we focus on the second step. The results show that region-based spatial relations can be learned to a high degree and are distinguishable on the basis of different linguistic features.

1 Introduction

Spatial language interpretation (SLI) is essential in areas such as *artificial intelligence*, *human-computer interaction*, and *geographic information systems*, where different formalisms have been developed that cope with the problem of interpreting natural language expressions in terms of their contextualized meaning. Here, we focus on interpreting (English) spatial language by automatic mapping its situational meaning to *qualitative spatial representations* (QSR). Although the mapping from (unstricted) spatial language to QSR covers only parts of the entire SLI process, it is one of SLI's central tasks as spatial language describes spatial information primarily in qualitative terms. Moreover, the mapping from spatial language to formal models allows spatial reasoning, whereas this is hardly feasible with spatial language itself [2].
To map language to QSR, we distinguish two levels [1]:

- **(I) the linguistic level.** Natural language is analyzed and parts of a sentence, i.e., *linguistic features* are syntactically and semantically categorized as having different spatial roles that convey certain spatial information
- **(II) the spatial level.** The linguistic features are mapped to specific QSR, i.e., *spatial calculi*

Table 1. Number of annotations in CLEF.

tr	lm	sp	trPh	lmPh	spPh	mo	pa	for	dy	gum	Total
354	183	43	258	142	56	12	4	2	2	39	1095

For example, in the sentence *"The book is on the table"*, the first level is the identification of spatial linguistic features, e.g., a spatial relation *"on"* (the spatial indicator with the semantics of *"support"*) that holds between the *"book"* (the trajector) and the *"table"* (the landmark), and the second level is the mapping to an adequate spatial calculus relation, such as *externally connected* (EC) or *disconnected* (DC) in RCC-8 [3].

2 From Natural Language to Formal Spatial Relations

The aim is to investigate whether a mapping from complex spatial utterances to formal spatial relations can be learned from examples. Complying with the two levels, we first perform what we call *spatial role labeling* (**SpRL**) to extract and identify spatial linguistic features selected on the basis of holistic spatial semantics [8, 7] and ontological information (GUM [2]). In a second step, called *spatial qualitative labeling* (**SpQL**), we learn to map these features to qualitative spatial relations [6].

Linguistic features used in the second level are primarily trajector (tr), landmark (lm), and spatial indicator (sp) together with more extensive features, namely trajector phrase (trPh), landmark phrase (lmPh), spatial indicator phrase (spPh), motion (mo), path (pa), frame of reference (for), dynamicity (dy), and the spatial modality according to GUM. The following examples shows how these features are annotated:

> **The book is on the table.**
> trajector: the book; landmark: the table; spatial indicator: on; general-type: region; specific-type: RCC-8; spatial-value: EC/DC; DY: static; path: none; frame of reference: none; gum-spatial-modality: Support)

Corpus Data The annotated data consists of textual descriptions of 613 images taken from the IAPR TC-12 Image data set [4] (CLEF). It contains 1213 English sentences and 1716 extracted spatial relations. The data contains 1040 annotated *topological* relations. An overview of the input features are shown in table 1.

2.1 (I) Linguistic Level: Extracting Spatial Relations from Language

The first step in our overall approach is spatial role labeling, which is extensively described in recent work [5]. There we have employed machine learning techniques (conditional random fields; CRF) to tag sentences with the roles *trajector*, *landmark*, and *spatial indicator*, with the aim of extracting spatial relations, e.g., triples (spatial indicator, trajector, landmark) such as (*on*, *book*, *table*). The resulting performance with linear chain CRF's for joint learning was an F-measure of 0.72 for the ImageCLEF benchmark and an F-measure of 0.89 for a subset of the MapTask corpus.

2.2 (II) Spatial Level: Mapping Linguistic Features to Qualitative Spatial Relations

We designed a set of experiments to map from spatial roles to the following sets, and the influences of individual spatial roles are analyzed by adding them gradually. Two different sets of spatial relations are learned (RCC-8, and RCC-mod), and spatial modalities in GUM-space are used as features.

RCC-8. *RCC-8={EC, DC, TPP, TPPI, NTPP, NTPPI, PO, EQ, NONE}*

RCC-mod. *RCC-mod={EC, DC, PP, PO, EQ, NONE}*, i.e., subsuming {*TPP, NTPP, TPPI, NTPPI*} under {*PP*}

GUM-space. Spatial modalities of GUM are used in two ways: either as an input feature together with the spatial roles or as an output feature of the mapping process.

3 Experimental Results

We use support vector machines (SVM) of the well-known WEKA toolbox to learn the mapping of the second level. Evaluation is performed with 10-fold cross-validation. The experimental results for RCC-8 and RCC-mod are reported in terms of F-measure in the tables of Figure 1. As expected, the simplification by RCC-mod improves the overall performance.

RCC-8	**#Instances**	**F-measure**
DC	147	0.859
TPP	369	0.847
NTPP	12	0.571
EC	458	0.850
EQ	5	0.8
NTPPI	1	0
TPPI	25	0.412
PO	15	0.583
NONE	684	0.962
Weighted Avg.	1716	0.883

RCC-mod	**#Instances**	**F-measure**
DC	147	0.86
EC	458	0.857
EQ	5	0.667
PP	407	0.862
PO	15	0.583
NONE	684	0.964
Weighted Avg.	1716	0.898

Fig. 1. RCC-8 (left) and RCC-mod (right) classification results for CLEF.

The reported results were obtained using all linguistic features introduced above for learning the mapping. However, to analyze the influence of individual linguistic features, the graphs in Figure 2 show the learning performance for gradually adding the different features.

4 Conclusions

Our experimental results show, most importantly, that indeed both the extraction of spatial features and the mapping to qualitative relations can be learned,

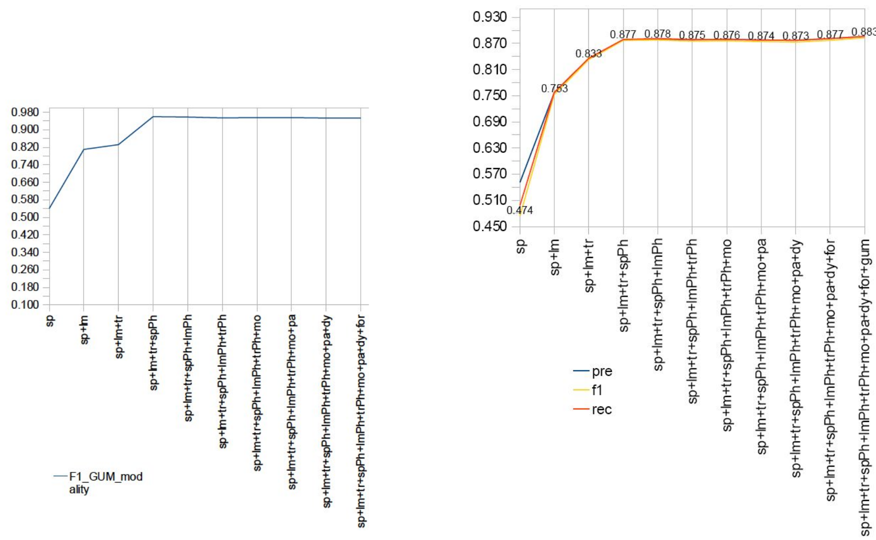

Fig. 2. Learning curve for GUM spatial modality (left) and RCC-8 (right) by gradually adding linguistic features.

and that the approach is computationally feasible. Feature analysis indicates that the most influential features are trajector, landmark, and spatial indicator. Although phrase information seems to have no effect on the learning, it may become more important when the phrase contains functional information. As we mapped primarily static and region-based spatial descriptions in our experiments, the spatial features path, motion, and frame of reference had only a negligible effect on the results. The linguistic features showed high performance for predicting GUM's spatial modalities on our corpus. In the future we consider mapping natural language to spatial ontologies considering hierarchical information and also mapping to multiple spatial calculi.

References

1. Bateman, J.A.: Language and space: a two-level semantic approach based on principles of ontological engineering. Int. J. of Speech Technology 13(1), 29–48 (2010)
2. Bateman, J.A., Hois, J., Ross, R., Tenbrink, T.: A linguistic ontology of space for natural language processing. Artificial Intelligence 174(14), 1027–1071 (2010)
3. Cohn, A.G., Bennett, B., Gooday, J., Gotts, N.M.: Representing and reasoning with qualitative spatial relations. In: Stock, O. (ed.) Spatial and Temporal Reasoning, pp. 97–132. Springer (1997)
4. Grubinger, M., Clough, P., Müller, H., Deselaers, T.: The IAPR benchmark: A new evaluation resource for visual information systems. In: Int. Conf. on Language Resources and Evaluation, LREC'06 (2006)
5. Kordjamshidi, P., van Otterlo, M., Moens, M.F.: Spatial role labeling: Towards extraction of spatial relations from natural language. ACM Transactions on Speech and Language Processing, under revision

6. Kordjamshidi, P., van Otterlo, M., Moens, M.F.: From language towards formal spatial calculi. In: Workshop on Computational Models of Spatial Language Interpretation, CoSLI'10 (2010)
7. Kordjamshidi, P., van Otterlo, M., Moens, M.F.: Spatial role labeling: Task definition and annotation scheme. In: 7th Conf. on Int. Language Resources and Evaluation, LREC'10 (2010)
8. Zlatev, J.: Spatial semantics. In: Geeraerts, D., Cuyckens, H. (eds.) The Oxford Handbook of Cognitive Linguistics, pp. 318–350. Oxford Univ. Press (2007)