From Reports to Maps

Ulrich Schade, Miłosław L. Frey, Sebastian Becker
FGAN / FKIE
Neuenahrer Straße 20
D-53343 Wachtberg, Germany
e-mail: schade;m.frey;becker@fgan.de

In this paper, we will sketch a project in progress. The project aims at an application of a command and control system. The application is meant to process military reports written in natural language. It exploits computer linguistic techniques, especially information extraction and ontological augmentation. A prototype has already be completed.

A real world application of report processing has to go beyond pure syntactic parsing. Semantic analysis is needed and the meaning of the report has to be constructed. Even more, the meaning has to be represented in a format such that it can be visualized within the so called "common operational picture" (COP). The COP is an interactive map displaying information. COP standards are provided by NATO. Since military operations of our days – war operations as well as peace-keeping and nation building ones – involve forces of many nations, the COP serves as main tool for synchronizing actions and plans. The paper at hand will provide some insights what kind of problems come along if language processing has to result in map visualization. It also will describe some solutions to overcome these problems.

1 The application's blueprint

The application operates as follows. First, a report is divided into sentence-like units. Second, these units are transformed into a formal representation by Information Extraction. Third, this representation is augmented by ontological processes. Finally, the augmented representation is visualized within the COP. In the following, we will focus on the interplay of ontological augmentation and visualization, but start with some remarks on Information Extraction.

2 Information Extraction

The task of Information Extraction is to transform the report's content into a formal representation. The IE is based on the SMES system (Neumann, 2003). It is adapted to the military domain by exploiting corpora of military reports, cf. (Hecking, 2003) for more details about the adaptation. The resulting representation is a feature structure, the standard format used in computational linguistics (Shieber, 1986; Bresnan, 2001).

3 Ontological Processing

Feature structures are underspecified. This is imperative for unification processes, but it also helps with respect to augmentation. Ontological augmentation is meant to explicit the report's meaning. In addition, the requirements for map visualization have to be met. To show a report's content within the COP, coordinates have to be provided which explicitly signify where to put the symbols representing units and equipment mentioned in the report. Sometimes, coordinates can be checked up. This is especially true, if own troops have to be visualized which are equipped with GPS. But more often, coordinates have to be calculated by ontological means. The application's ontology provides knowledge about specific military objects, e.g., control features, the front line, the structure of convoys, and more. It also provides knowledge about spatial relations and reference. Let us assume a unit reconnoiters enemy movements and reports "Zwei Bradyland Raketenwerfer hinter Vinstedt" (Two Bradyland rocket launchers behind Vinstedt). In this case, it is safe to assume a secondary deitic reference: The weapons are positioned "behind" the village with respect to the reporting unit. Thus, an axis is calculated from the reporting unit through Vinstedt, and the rocket launchers' symbol is placed on this axis next to the village. A similar case is given, if the unit reports "Feindlicher Konvoi in Zufahrt" (Hostile convoy approaching). Here, either the forefront of the reporting unit (primary deictic reference) or if it is not known, the frontline (intrinsic reference) determines the calculation. Intrinsic reference also is used whenever a relatum has a prominent intrinsic orientation. E.g., "Demonstration vor Kloster Ebstorf" (Demonstration in front of Ebstorf monastery) is analyzed under the assumption that "vor" means "outside at the entrance".

4 Visualization

Visualization of objects within the COP requires coordinates on the one hand and symbols on the other hand. Symbols are composed according to given military standards (MAS, 1997). Basically, these standards define

taxonomical hierarchies which provide an object's blueprint symbol. For example, a wheel loader can be found under "ground track – equipment – ground vehicle – engineer vehicle – construction vehicle." Each step down the hierarchy adds to the complexity of the blueprints. To complete the symbol, additional information is needed. E.g., the color of a symbol is determined by its objects affiliation attribute (friend = blue, hostile = red, neutral = green, unknown = yellow). In some cases, missing attributes might be inferred by ontological means. Since inferring normally adds uncertainty, some inferring processes are made facultative. Facultative processes can be activated or shut-down as required. For example, there is a hostility determination process which proposes an affiliation value for units of unknown affiliation. This process takes equipment reported into account and exploits the military knowledge about standard equipment. Another example is the unit determination process. This process suggests a type of unit if only equipment is mentioned. As a result, the most probable type of unit for operating the equipment is visualized within the COP.

5 Survey

The visualization of reports within a map enforces specific semantic processes. Coordinates as well as symbols representing the objects and actions mentioned in the report have to be determined. The application we are developing uses ontological means in order to fulfill the demands. More details about the project are given in (Schade, 2004) and (Schade and Frey, 2004).

References

J. Bresnan. *Lexical–Functional Syntax.* Blackwell, Oxford, UK, 2001.

M. Hecking. Information extraction from battlefield reports. In *Proceedings of the 8th ICCRTS*, Washington, DC, 2003. CCRP.

MAS. Military symbols for land-based systems (APP–6A). Military Agency for Standardisation, 1997.

G. Neumann. http://www.dfki.de/ neumann/pd-smes/pd-smes.html, 2003.

U. Schade. Automatic report processing. In *Proceedings of the 9th ICCRTS*, Washington, DC, 2004. CCRP.

U. Schade and M. Frey. Beyond information extraction: The role of ontology in military report processing. In E. Buchberger, editor, *Proceedings of KONVENS 2004*, Wien, 2004.

S. M. Shieber. *An Introduction to Unification–Based Approaches to Grammar.* CSLI, Stanford, CA, 1986.