EFGT net: A Resource Encoding Knowledge about Named Entities, Thematic Fields, Geographic Areas, and Temporal Periods

Levin Brunner and Klaus U. Schulz and Felix Weigel

CIS, University of Munich Lastname@cis.uni-muenchen.de

Named entities (e.g., "Ludwig van Beethoven", "Daimler-Chrysler", "Dresdner Bank", "Kofi Annan", "Second World War", "Coca-Cola") are ubiquitous in documents in the web and other document repositories. The information that a human user associates with named entities occurring in a document often suffices to derive a simplified picture, or a fingerprint, of its contents. It thus represents adequate meta-information for the document. In order to use this kind of information in automated document processing, resources are needed that make the information implicitly carried by named entities explicit, formalizing it in an appropriate way. The EFGT net [SW03], currently developed at CIS, University of Munich, represents such a resource, focussing on thematic fields (F), geographic locations (G) and temporal periods (T) associated with the entities (E). Entries of these four types are positioned in a well-founded (i.e., acyclic) navigation space. Large (small) entries w.r.t. the navigation order represent general (specific) topics and large (small) geographic or temporal areas. This acyclic organization of knowledge aims to support classification and indexing tasks. Each entry comes with a unique identifier that describes the role of the entry. A semi-formal semantics for the language of identifiers is given.

The current version of the EFGT net contains ca. 8.000 entries. An important feature of the system is the following: when adding a new entry, the position in the hierarchy is computed in a fully automated way. Relations such as thematic, geographic or temporal inclusion are taken into account. The same deduction mechanisms can be used for searching in documents and for semantic indexing.

References

[SW03] Klaus U. Schulz and Felix Weigel. Systematics and architecture for a resource representing knowledge about named entities. In Jan Maluszynski Francois Bry, Nicola Henze, editor, *Principles and Practice of Semantic Web Reasoning*, pages 189–208, Berlin, 2003. Springer-Verlag.