

Ontology Learning and Beyond

Vít Nováček^{1,2}

¹**Digital Enterprise Research Institute (DERI)**
National University of Ireland, Galway
email: {vit.novacek}@deri.org

²**NLP Lab, Faculty of Informatics**
Masaryk University, Brno, Czech Republic

Svratka, Czech Republic — December 1, 2008



Outline

- 1 Introduction
 - Vit's Involvement
 - Overview of Vit's Work
- 2 Results
 - Ontology Learning and Merging
 - Ontology Dynamics
 - Empirical KR&R
- 3 Dissemination Summary
- 4 Final Words



Vit's Involvement in the SeWe Project

- **half-time** employed as a **specialist** from July 2004 till June 2005
 - working on the **ontology learning** topic
- **full-time** employed as a **specialist** from July 2005 till March 2007
 - working on the **ontology learning and merging, uncertainty** topics
- **part-time** employed as a **consultant** from April 2007 till December 2008 (after starting a full-time PhD with DERI, Galway)
 - working on the **ontology dynamics, empirical KR&R** topics



Overview of Vit's Work Relevant to the Project

- ontology **learning** and **merging** framework with a native support for **uncertainty** (master thesis)
- dynamic ontology **integration** (co-leading the Ontology Dynamics research work package of the Knowledge Web EU NoE)
- light-weight **representation** of and **reasoning** with **emergent knowledge** (current PhD topic)



Automated Ontology Learning and Merging

- **learning** of ontologies from (English) free text:
 - lexico-semantic **pattern matching**
 - support for **import** of existing and **custom design** of new patterns
 - unsupervised **conceptual clustering** and WordNet-based cluster **annotation** with common concept names
- **merging** of learned ontologies:
 - assignment of **certainty measures** to learned statements based on the **frequency** of respective relations among data
 - heuristic **merging** of learned statements from small **document-bound** ontologies into larger ones covering the **whole domain**
 - automated empirical **refinement** and **reinforcement** of **irrelevant** and **relevant** knowledge, respectively
- illustrative **web interface** to the respective **software** used to be at <http://nlp.fi.muni.cz/projects/ole/>



Ontology Dynamics (KWeb EU NoE)

- work elaborated within **coordination** of the Knowledge Web EU NoE work package on **Ontology Dynamics**
- primary topic: methodology and implementation of **dynamic ontology life-cycle** with support for **integration** of a **automatically extracted** knowledge and **manually designed** master models
 - a proof-of-concept Java **application** (support discontinued for now, though) available at
<http://140.203.154.209/~vit/tools/dino/>
- secondary topic: preliminary **analysis** of possibilities of **multi-version** ontology **reasoning** built on the top of **linear temporal logics** and/or **contextualised** Description Logics **inference**



Towards Managing the Emergent Knowledge

- **proposal** and **implementation** of light-weight non-logical **knowledge representation** with the following features:
 - 1 **robustness** regarding noisy and imprecise emergent data (e.g., learned ontologies, data mining or Semantic Web crawling results)
 - 2 **scalability** and **dynamics** matching the web growth and fluidity
 - 3 **support** for a light-weight **similarity-based inference** superseding logical reasoning methods that are inappropriate in these settings
 - 4 **accessibility** regarding common users (even your mother should grasp it and be able to play with it)
- **prototype** of a **knowledge-based** publication **search engine** using the **empirical KR&R** technology currently submitted to the Elsevier Grand Challenge semi-final round
- to be **linked** from `http://smile.deri.ie` soon



Overview of Vit's Dissemination Activities

- book chapters: 1 (Springer, LNBI series, 2008)
- journal articles: 1 (Elsevier, Journal of Biomedical Informatics, 2008)
- conference papers: 6 (LREC'06, SOFSEM'06, ESWC'06, ECCB'06, FLAIRS'07, ICEIS'07)
- workshop papers: 8 (various, 2005–2007)
- theses: 1 (MSc. thesis, 2006)
- tutorials: 1 (EuroMasters Summer School, 2006)



Final Words

- sorry for not being present in person
- sorry for the English (looks a bit better than “cestina”, though ;))
- thanks to the actual presenter for tackling this on my behalf
- thanks everybody for the attention

