

Generating Lexical Data for Novel Analogies between Computer Graphics Publications using Dr Inventor

Donny Hurley, Diarmuid P. O'Donoghue, Yalemisew Abgaz (Maynooth University)
donny.hurley@nuim.ie

The "big data" era offers novel challenges for accelerating scientific progress and enabling new modes of discovery (Honavar 2014; Singh & Reddy 2014). We present some work on the "Dr Inventor" (O'Donoghue et al. 2014; O'Donoghue et al. 2015) creativity support tool that aims to uncover novel analogy-based comparisons (Gentner 1983; Fauconnier & Turner 1998) between academic publications.

Dr Inventor does not work directly on the publications, but instead generates Research Object Skeleton (ROS) graphs. Generation of a ROS graph starts when a pdf document enters the Text Mining Framework (Ronzano & Saggion 2015), addressing problems arising from the layout, text flow, images, equations etc. A parser generates the dependency tree for each sentence and Agarwal et al. (2015) we apply a set of rules to the dependency trees, generating connected triples of nouns and verbs forming the ROS graph. Crucially, multiple mentions of the same concept are uniquely represented within each ROS, using the co-reference resolution output from the dependency parser.

ROS graphs enable the application of Gentner's (1983) structure mapping theory to finding and evaluating analogies between ROSs - and thus between publications. This uses a combination of computational power and topologically driven analogical retrieval. Semantic web annotations (Ruiz-Iniesta & Corcho 2014) of sentences allow Dr Inventor to explore analogies between the "background" of one paper and the "approach" of another. Dr Inventor is being evaluated by experts in computer graphics for its ability to discover novel and useful analogies, inspiring its users and igniting their creativity.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme ([FP7/2007-2013]) under grant agreement no 611383.

References:

- Agarwal, B. et al. (2015). Concept-level sentiment analysis with dependency-based semantic parsing: A novel approach. *Cognitive Computation*, 1-13.
- Boden, M. (2004). "The creative mind: Myths and mechanisms" (2nd ed.). Routledge.
- Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive science*.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, Volume 7, 155-170.
- Honavar, V. G. (2014). The Promise and Potential of Big Data: A Case for Discovery Informatics. *Review of Policy Research*, 31(4), 326-330.
- O'Donoghue, D., et al. (2015) 'Stimulating and Simulating Creativity with Dr Inventor'. International Conference on Computational Creativity (ICCC), Park City, Utah.
- O'Donoghue, D. et al. (2014). Towards Dr Inventor: A Tool for Promoting Scientific Creativity. ICCC 2014 - Late Breaking Paper. Ljubljana, Slovenia.
- Ronzano, F., & Saggion, H. (2015). Dr. Inventor Framework: Extracting Structured Information from Scientific Publications. *Discovery Science*. Springer.
- Ruiz-Iniesta, A., & Corcho, O. (2014). A review of ontologies for describing scholarly and scientific documents. In *SePublica*.
- Singh, D., & Reddy, C. (2014). A survey on platforms for big data analytics. *J. of Big Data*.