VEDRAN GALETIĆ

BOLNIČKA CESTA 54

10090 ZAGREB

CROATIA

ESF

NetWordS

Prototypicality quantification by combining conceptual space theory and corpus analysis

Final report for short study visit

Zagreb, Croatia, 30 June 2013

Purpose of the visit

The main objective of the study visit to Professor Gärdenfors was to expand and fortify theoretical underpinnings to the semantic theory of conceptual spaces. It had been in the focus of my scientific attempts to:

1. Investigate the set of properties via which a conceptual space is spanned as well as the nature of these properties;
2. Determine the parameters that are relevant for situating concepts within geometrically organized conceptual space (e.g. property weights, distribution of typicality values over property values etc.);
3. Examine the means and prospects of (semi-)automatized extraction of these parameters' values via computational corpus analyses.

The design of such knowledge base of linguistic concepts is an extremely challenging task, yet such a model would without doubt be scientifically relevant. Namely, it would make an additional step towards bridging the gap between computational linguistics and conceptual spaces theory. Such combination would be welcome as distributional approaches, which are becoming increasingly popular in computational semantics, by and large do not exhibit cognitive plausibility. This problem would be addressed by endowing it with cognitively plausible theory such as conceptual spaces.

Having this in mind I had conceived a model of formal quantification of prototypicality of concepts (Galetić 2011). The situation of concept instances in a conceptual space is determined by two parameters: (1) representativeness of each of its property value with respect to the category and (2) property weights. This year I have been conducting an EEG experiment for verifying a hypothesis which, if confirmed, would allow for a certain method of (semi-)automatized extraction of parameters (primarily the property weight, potentially the representativeness function too) from the corpus. The attempt has been partially successful as some parameters, currently beyond my awareness, were apparently not taken into account. It was my intention to share the main underlying hypothesis with Professor and further refine the experiment's theoretical underpinnings.

Since there is no individual, let alone a team, at my PhD study whose work concentrates on conceptual spaces and related theories, I found it potentially highly beneficial to visit Professor Gärdenfors, both for my research as well as transferring novel knowledge and communicating current trends in pertaining area towards my home institution.

Description of work

On the very day of my arrival to Lund I met with Professor Gärdenfors and we discussed the focus of my work to be done during the visit. I had prepared an array of potential topics and upon discussion we agreed that the focus would be on property correlations. This phenomenon, manifesting in generalizing capability during concept learning and categorization, has been an intriguing topic for a long time. Professor acknowledged that he had not covered it to full extent in his work and proposed that I try to tackle the problem by studying extant work in the area (theories and empirical data) and proposing means of incorporating it within the attempts of formalizing and operationalizing conceptual spaces.

I met with Professor on six occasions during my stay. Each time we would discuss potential upgrades to the current conceptual space models and Professor would recommend a selection of materials for me to study. Also, I introduced Professor to some works that he had not come across and he found them very interesting and potentially useful. At each subsequent meeting I would share new observations that I obtained from the read materials as well as my drawn conclusions. Professor would comment on my ideas and we would consensually determine next steps of the research along with the pertaining materials.

A couple of times during my stay I had a Skype meeting with my mentor, Assistant Professor Jan Šnajder, at my home University of Zagreb and updated him with novelties regarding research. He would provide me with further materials from his expertise (computational linguistics, information extraction) that he found relevant for the current status of the research process. Since Professor Gärdenfors does not have as much expertise in the letter area, which is crucial for operationalization attempts, I believe such trilateral knowledge transfer is very beneficial for all involved parties (naturally, mostly for my own academic attempts).

Apart from meetings with Professor Gärdenfors and online meetings with my mentor, I got the opportunity to take part at a seminar about types of learning in artificial intelligence systems and their connections and parallelisms to empirical data on learning phenomena. This talk provided me with a valuable recapitulation of the theories and methodologies that I got introduced to during my pre-diploma study.

Finally, I got acquainted with an array of employees of the Institute of Philosophy, and we introduced each other to own work, exchanged contacts and shared our papers.

Obtained results

The outcomes of the study visit can be roughly summarized in the following manner:

- Expanded knowledge in the domain of conceptual semantics;
- Gained first-hand information on current trends in the domain's theoretical development;
- Collected and shared knowledge on prospects of combining conceptual spaces with distributional approaches of computational semantics, e.g. for (semi-)automatic extraction and typifying properties relevant for constructing conceptual spaces, for qualifying and quantifying property correlations etc.

The first observation I choose to emphasise, which determined the direction of most of work carried out during the stay, was empirical results indicating that humans are poor at learning concepts based on isolated correlations (Billman, Knutson 1996; Kornblith 1995) and pure perceptual similarity (Gelman, Markman 1986); rather, they rely upon bundles of correlated properties (Billman, Knutson 1996; Kornblith 1995; Jones, Smith 2002; Kloos, Sloutsky 2008; McClellald, Rogers 2003). These correlations need not only be numerous, but also appropriately structured (Billman, Knutson 1996) reflecting their ecological validity (Gärdenfors 2000:225). It is claimed that natural categories are compliant with these characteristics, which enables even very young children to generalize correctly without supervision (Billman, Knutson 1996).

It needs to be acknowledged that current interpretations of these phenomena do not mitigate their elusiveness from the perspective of computational approaches to (lexical) semantics. In line with this claim is the fact that Professor Gärdenfors, who has arguably developed the conceptual space theory to the largest extent yet, finds this problem one of the most crucial, especially for computational modelling attempts. As a result, I found this area worth dedicating the attention of my visit.

By dwelling into problems of interpreting, theoretically describing, formalizing and utilizing the empirical results on detecting property correlations, it became obvious to me that so-called "theories" play a crucial role in undisputed effects of observed property correlations on concept generalization. According to Murphy's and Medin's (1985) paper, which I take as one of the referent works in this area, theories reflect "a complex set of relations between concepts, usually with a causal basis" (pp. 290). There is a strong connection to the notion of psychological essentialism (account in Kornblith 1995) emphasizing the causality between objects' essential properties and their superficial, perceptual properties. To my knowledge, there does not exist a conceptual space model that includes theories or formalizes their influence.

Additionally, one of the greatest and elusive problems in formalizing semantic representations based on conceptual spaces is the question of context. It is common

acknowledgement (e.g. Frassinelli, Lenci 2012; Gärdenfors 2000; Song, Bruza 2001) that context plays a crucial role in determining property salience. Although formalizing context was not in focus during my visit, Professor Gärdenfors stated, and I accepted, that any future cognitively plausible semantic model needs to take context into account which further determines my future research activities.

In the context of my academic attempts, I find the visit very useful for narrowing the scope of my research and thereby making the topic of my PhD thesis more focused. Namely, throughout the stay, during interaction with Professor and studying materials, an idea for the following future research steps has begun taking form:

1. Further theoretical development with an emphasis on the "underlying", causally connected properties, by consolidating available theoretical accounts and empirical data.
2. Attempt to extract correlated features of concepts by using a (semi-)automatic method of corpus analysis.
3. Eliminate all properties that are correlated arbitrarily, i.e. "by chance" of their utilization value (cf. Gibson's affordances), which is true for most (possibly all) artefacts. It is necessary to focus only on truly causally correlated properties present within natural categories in order to investigate the nature of the underlying properties.
4. Modify (or upgrade) a current conceptual space model (Gärdenfors' or mine (Galetić 2011)) with respect to these underlying properties.
5. Examine the influence of introducing a specific subcategory to the category (e.g. learning that an ostrich is a type of bird) on the weight of properties for that category (for proposal of formal treatment of influence of property weights on categorization see Galetić 2011).

Concerning the second and third point, it is envisioned to combine conceptual space theory with corpus analysis methods (for a review of such an attempt see Andrews, Vigliocco, Vinson 2009). A necessary presumption is the ability to categorize different property types (e.g. Baroni, Lenci 2008; Frassinelli, Lenci 2012).

Concerning the last point, I should emphasize that I presented my hypotheses on relation between property weights and property value distribution to Professor and he found three of four hypotheses interesting and worth further investigation. If any of these hypotheses (or, more likely, their combination) can be empirically validated (e.g. via a behavioural or an EEG experiment, such as the one I mentioned at the beginning of the document), this would provide justification of employing corpus analysis methods for (semi-)automatic extraction of conceptual space model's parameters.

Next steps

The main framework for future research steps is outlined at the end of the previous section. Moreover, Professor Gärdenfors and I agreed to stay in contact. Meanwhile, I will continue transferring the newly acquired knowledge to my mentor Professor Šnajder. I will take into account the obtained theoretical observations and empirical data and, with the aid of his guidance, work on incorporating them into the attempts of computational modelling of semantic representations.

While there have not been any concrete agreements with Professor Gärdenfors on producing a collaborative scientific article, the visit has left me with a firm impression that such an option is open. The administrative potential of formal collaboration between the University of Zagreb and Lund University is beyond the scope of my knowledge and influence; however, Professor Gärdenfors' participation at 2012 NetWordS summer school in Dubrovnik may hopefully provide a positive step towards this direction.

Finally, the work and conclusions of my study visit will be presented at the Third NetWordS workshop in Dubrovnik, 19-20 September 2013, where I expect to gain novel contacts, hopefully further knowledge exchange and, possibly, an opportunity to establish future collaborative work.

References

Andrews, Mark, Gabriella Vigliocco, and David Vinson. "Integrating experiential and distributional data to learn semantic representations." *Psychological review* 116.3 (2009): 463-498.

Baroni, Marco, and Alessandro Lenci. "Concepts and properties in word spaces." *Italian Journal of Linguistics* 20.1 (2008): 55-88.

Billman, Dorrit, and James Knutson. "Unsupervised concept learning and value systematicity: A complex whole aids learning the parts." *Journal of Experimental Psychology-Learning Memory and Cognition* 22.2 (1996): 458-475.

Frassinelli, Diego, and Alessandro Lenci. "Concepts in context: Evidence from a feature-norming study." *Proceedings of CogSci 2012* (2012).

Galetić, Vedran. "An Aggressive Robin in the Backyard: Formal Quantification of Prototypicality Level within the Frame of the Prototype Semantic Theory of Cognitive Linguistics." *Suvremena lingvistika* 37.71 (2011): 0-0.

Gelman, Susan A., and Ellen M. Markman. "Categories and induction in young children." *Cognition* 23.3 (1986): 183-209.

Gärdenfors, Peter. *Conceptual spaces: The geometry of thought*. MIT press, 2004.

Jones, Susan S., and Linda B. Smith. "How children know the relevant properties for generalizing object names." *Developmental Science* 5.2 (2002): 219-232.

Kloos, Heidi, and Vladimir M. Sloutsky. "What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories." *Journal of experimental psychology. General* 137.1 (2008): 52.

Kornblith, Hilary. *Inductive Inference and Is Natural Ground: An Essay in Naturalistic Philosophy*. The MIT Press, 1995.

McClelland, James L., and Timothy T. Rogers. "The parallel distributed processing approach to semantic cognition." *Nature Reviews Neuroscience* 4.4 (2003): 310-322.

Murphy, Gregory L., and Douglas L. Medin. "The role of theories in conceptual coherence." *Psychological review* 92.3 (1985): 289-316.

Song, Dawei, and Peter Bruza. "Discovering information flow suing high dimensional conceptual space." *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2001.