Chapter 12

Conceptual modeling and learning

Krista Lagus, Timo Honkela, Tiina Lindh-Knuutila, Mari-Sanna Paukkeri, Juha Raitio, Oskar Kohonen, Paul Wagner

12.1 Introduction

Conceptual modeling is a task which has traditionally been conducted manually. In artificial intelligence, knowledge engineers have written descriptions of various domains using formalisms based on predicate logic and other symbolic representations such as semantic networks and rule-based systems. The development of expert systems in 1980s was a notable example of such efforts. As a modern related attempt, the Semantic Web can be mentioned. It seems that the complexity and changing nature of most of the domains makes such formalisms problematic in many real-world applications.

A problem often neglected in symbolic knowledge representation tradition is subjectivity. For us, it seems evident that major portions of individual conceptual systems are learned. Due to the individual and cultural differences, it is not believable that concepts could be modeled with static structures without making use of adaptive processes.

Another challenging topic related to conceptual modeling is contextuality. Contextuality is illustrated in Fig. 12.1. Human activity takes normally place in rich contexts in which the relationship between prototypical meanings of expressions and the situation may be complex. Phenomena like subjectivity and contextuality serve as motivation for the research that is described in the following.

Figure 12.1: An illustration of contextual effects in the interpretation of linguistic expressions. There is a prototypical red but the redness of a shirt typically differs considerably from the redness of skin or wine. In an image, white snow may not altogether be very white.

The theories of knowledge have traditionally been based on predicate logic and related methodologies and frameworks. The basic ontological assumption is that the world consists of objects, events and relationships. The language and the conceptual structures are then supposed to reflect rather straightforwardly this structure. Learning has been seen as a means to memorize the mapping from the epistemological domain (to put it simply: words) into the ontological domain (objects, events and relationships). This view has been dominant at least partly because of the consistent formalization of the theory through the use of symbolic logic. Moreover, the use of the von Neumann computer as the model or metaphor of human learning and memory has had similar effects and has strengthened the idea of the memory as a storage of separate compartments which are accessed and processed separately and which are used in storing and retrieving information more or less as such. [4]

Realistic simulations of the socio-economical and cultural levels are seemingly difficult to build due to the complexity of the overall system. The richness of human culture makes it difficult as a phenomenon to model. Moreover, already the world knowledge of a single human being is so vast that it is difficult to approach it successfully. However, useful development may be possible by taking into account the aspects presented, e.g., in [9, 10, 1, 8, 3]. For instance, Vygotsky [10] has stated that "... the world of experience must be greatly simplified and generalized before it can be translated into symbols. Only in this way does communication become possible, for the individual's experience resides only in his own consciousness and is, strictly speaking, not communicable." Later, he continues: "The relation of thought to word is not a thing but a process, a continual movement back and forth from thought to word and from word to thought. In that process the relation of thought to word undergoes changes which themselves may be regarded as development in the functional sense." This means in practice that conceptualization is a complex process that takes place in a socio-cultural context, i.e., within a community of interacting individuals whose activities result into various kinds of cultural artifacts such as written texts.

The basic aim in our research group is to provide the means for a more or less automatic process of concept formation. This will facilitate both cost-effective development of knowledge-intensive systems as well as serve as a good basis for systems that can update themselves taking into account changes in the domain of interest.

Next we present three specific research areas within conceptual modeling with recent results. The *intersubjective communication model* aims at providing a general framework for explaining how communication between human or artificial agents that have different conceptual models can be successful and what kind of problems there also may be. We have also developed a *multiagent simulation model of conceptual development*. This model combines probabilistic modeling of concept naming with the self-organization of the underlying conceptual space in an agent population. In the third study, we have conducted an *analysis of philosophy students' conceptions*.

12.2 Intersubjective communication model

We have recently proposed a theoretical framework for modeling communication between two agents that have different conceptual models of their current context [5]. We have described how the emergence of subjective models of the world can be simulated and what the role of language and communication in that process is. We have considered the role of unsupervised learning in the formation of agents' conceptual models, the relative subjectivity of these models, and the communication and learning processes that lead into intersubjective sharing of concepts [5].

In this section, we introduce the basic definitions and notation used in our communication model for two agents. The key concept is the agent's internal view of its context, the *concept space*. The concept space is spanned by a number of features. We can use the terminology coined by Gärdenfors [2] calling each feature (f_i) a quality dimension. Dimensionalities of the concept spaces can be different for each agent . The concept space of agent 1 is N-dimensional metric space C^1, and for agent 2, C^2.

This work has several theoretical and practical implications including the possibility of approaching interoperability of information systems from a novel point of view. Some of these implications are discussed next (see the original article [5] for additional details and references).

The traditional notion of uncertainty in decision making does not cover the uncertainties caused by differences in conceptual systems of individual agents within a community. We claim that in all transactions including symbolic/linguistic communication the differences in the underlying conceptual systems play an important role. For instance, serious efforts have been made to harmonize or to standardize the classification systems used by business agents, e.g., using Semantic Web technologies. However, even if the standardization is conducted, there can not be any true guarantee that all the participating agents would share the meaning of all the expressions used in the business transactions in various contexts.

One implication is that in business transactions there should be means for checking what is a meant by some expressions by an access to a broader context (cf. symbol grounding). Moreover, rather than relying solely on a standardized conceptual system, one could introduce mechanisms of meaning negotiation. Before two business agents get into negotiation about, for instance, the price of some commodity, they should first check if they agree on what they refer to by the expressions that are used in the negotiation. This concern is valid both for human and computerized systems, even though humans are usually capable to conduct meaning negotiations even when they are not aware of it [5].

The harmonization of conceptual systems, such as the creation of ontologies for business transactions, has obvious benefits when, for instance, the interoperability of related information systems is considered. It appears ideal that all systems within some domain would use similar terminologies and shared ontologies. However, this approach can be claimed to be idealistic because the continuous change through innovations and other activities and the underlying learning processes within the human community lead into the situation carefully considered in the earlier chapters of this paper. All the agents have a conceptual system of their own, at least to some degree. Therefore, the harmonization of the conceptual systems should be considered only as a relative goal. One may aim for a larger degree of sharing of the conceptual system as before. A central theme is then to assess the associated benefits and costs. Here we do not try to provide any means to estimate the benefit of well working harmonized conceptual system implemented, e.g., as an ontology within the Semantic Web framework [5].

The costs stem from two main sources: the development of a shared conceptual systems

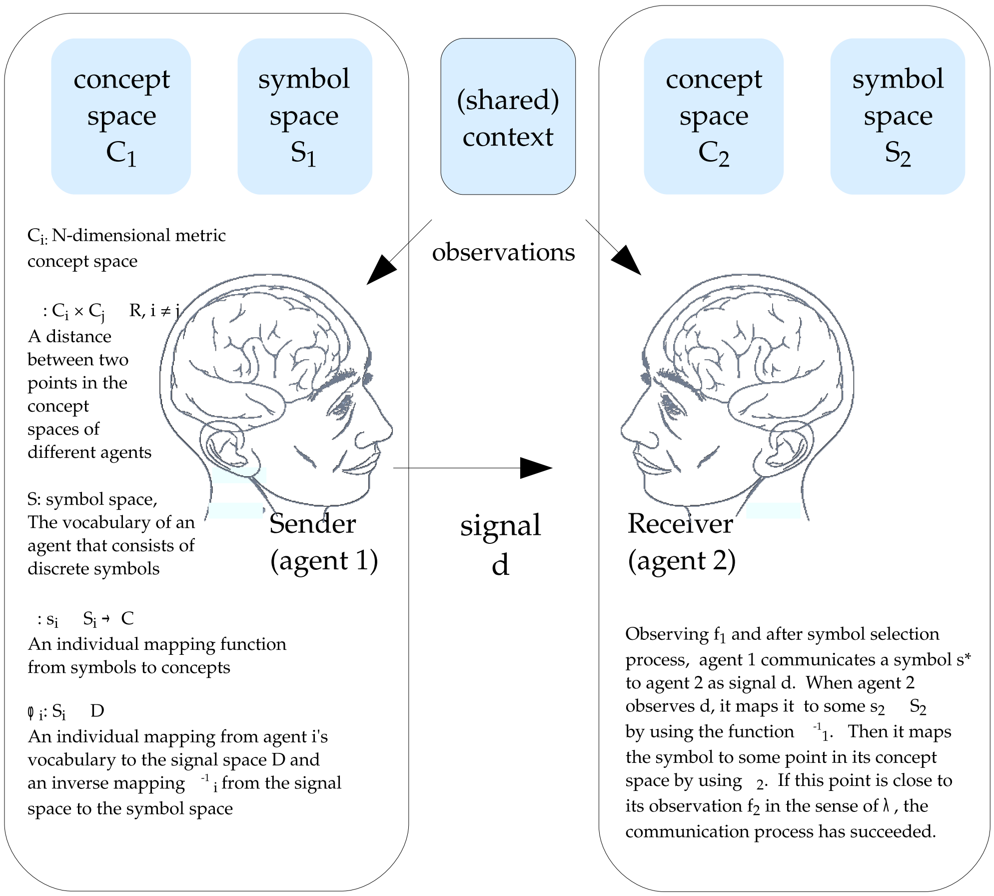

Figure 12.2: In the intersubjective communication framework, there are two distance measures ω and λ. ω gives a distance between two points inside the concept space of the agent, i.e. $\omega : C^i \times C^i \to \mathbb{R}, i = 1, 2,$. λ gives a distance between two points in the concept spaces of the different agents, i.e. $\lambda : C^i \times C^j \to \mathbb{R}, i \neq j$. The symbol space S^1 of the agent 1 is its vocabulary that consists of discrete symbols. Similarly, the vocabulary of agent 2 consists of symbols S^2. An agent i has an individual mapping function ξ^i that maps the symbol $s^i \in S^i$ to C^i. An agent i expresses each symbol $s^i \in S^i$ as a signal d in the signal space D. The signal space D is multidimensional, continuous and shared between the agents. Each agent i has an individual mapping function ϕ^i from its vocabulary to the signal space, i.e. $\phi^i : S^i \to D$ and an inverse mapping ϕ^{-i} from the signal space to the symbol space.

and the use of it. The development of an ontology typically consists of defining the concepts and the relationships between the concepts. The typical stages of an ontology building process are the following: (1) domain analysis resulting into the requirements specification, (2) conceptualization resulting into the conceptual model, (3) implementation that leads into the specification of the conceptual model in the selected representation language, and (4) the ontology population i.e. the generation of instances and their alignment to the model that results into the instantiated ontology [5].

The estimation of costs related to the use of ontologies is rather difficult. There are many kinds of uses of ontologies that require higher or lower degree of familiarity of conceptual structures of the domain. A widely cited claims from expertise research is the 10-year rule, first proposed in relation to expertise development among chess players, and later generalized to other domains. The essential content of the rule is that anyone seeking to perform at world-class level in any significant domain must engage in sustained, deliberate practice in the activity for a period of at least ten years. This figure serves only as a upper bound of an estimate for a person to learn to master the conceptual content of a complex domain [5].

12.3 Multiagent simulation model of conceptual development

In [6], we present a model that combines probabilistic modeling of concept naming with the self-organization of the underlying conceptual space in an agent population. In this multi-agent simulation framework, we study emergence of a common vocabulary. The self-organizing map is used for the purpose of transferring sensory perceps into a conceptual level representations.

In the community of agents, we assume that each agent has its own representation. While the representations are alike due to similar training data, the representations are not exactly the same. On top of the concept emergence, we studied shared vocabulary emergence using a naming game paradigm , in which two agents share a common perceived context, and they attempt to find a name to match their observation. Each agent matches the observation to their concept map by finding the best-matching unit for that data point in the self-organizing map. For that given map unit, each agent then selects the term to denote that observation based on the maximum likelihood, $\max(P(C|T))$, which is estimated as the number of successful uses of the term for a given map node, proportional to all of the successful uses of all the terms in that node. The likelihood is estimated for all the terms associated with the BMU and for those nodes adjacent to it, and the term with the highest likelihood is selected and uttered. If no term is found to be associated with the color or its neighborhood in the self-organizing map, a new term is invented. The hearer estimates the likelihood $P(C|T)$ in similar fashion. When a number of games is played, a common vocabulary emerges in the population.

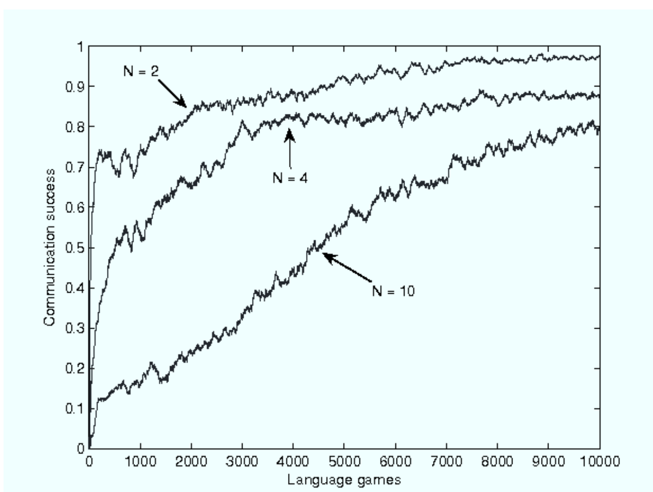

Figure 12.3: Communication success for $N = 2$, $N = 4$ and $N = 10$ agents in the population.

Figure 12.3 shows the communication success for two, four and ten agents, each averaged over 10 simulation runs. In the two-agent case, the communication success, the fraction of successful games of the previous hundred games played, rises rapidly to $CS = 0.8$ and then steadily up to $CS = 0.95$ during the $10,000$ simulated games. The communication success for four agents grows slower than in the previous experiment, but still increases up to $CS = 0.86$, where it seems to settle. The bigger population size, in the ten-agent case yields into considerably slower convergence, reaching approximately $CS = 0.8$

in 10,000 games.

All the language games are played pair-wise, i.e. only two agents of the whole population participate in each game, and other agents have access to the words only through subsequent language games with the same topic. This means that when the population size grows, the convergence to common vocabulary is considerably slower. More competing words for a given topic emerge, and it simply takes longer for each agent to see a representative subset of the topics.

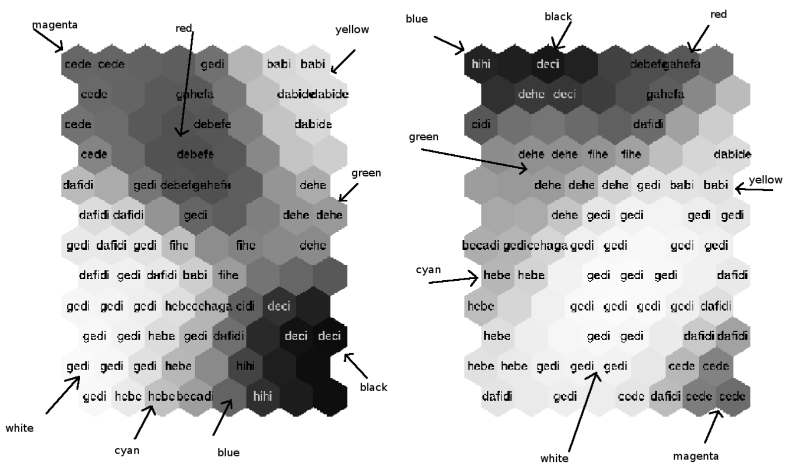

Figure 12.4: The conceptual memories of the agents in the two-agent simulation. Only the most probable label for each node is shown.

Figure 12.4 shows the conceptual maps of the two agents in the first experiment. The colors denote the converged RGB values of the prototype vectors of the map. The map has organized well and transformations from one color to other are smooth. The eight prototypical colors used are more prominent, since they are represented more in the data than the intermediate colors that have resulted from added noise.

When comparing the figures, it is evident that for most prototypical colors there are one or two words that are preferred: *deci* is preferred for black or dark, *hihi* for blue, *fehe* for green, *hebe* for cyan, *defebe* and *gahefa* for red, *cede* for magenta, and *babi* and *dabide* for yellow. For white, the most common word used is *gedi*, but there are also competing labels for bluish white, pinkish white and so on because white covers a larger area in the space. The conceptual memories support the conclusion already visible in the communication success ratio – that a common vocabulary for the agents has emerged.

12.4 Analysis of philosophy students' conceptions

In collaboration with researchers from Helsinki University, Anna-Mari Rusanen, Otto Lappi and Mikael Nederström, we have used the self-organizing map algorithm to analyze and visualize the initial conceptions of philosophy students [7].

The general theoretical approach of this study was based on the conceptual change paradigm. There is a large body of research which shows that novices conceptions do differ from those of experts, but researchers still remain divided not only about the nature of those differences, and also the status of novices' belief systems. Some researchers claim that novices belief systems are weakly organized systems that are internally inconsistent, piecemeal and incoherent. Other researchers argue that novice belief systems are not only internally quite coherent but they may also share the essential properties of scientific theories.[7]

To obtain information on the students' conceptions, we used a multiple choice-questionnaire. The questionnaire included 63 thematically selected items. Three thematic sets of questions probed (1) the subjects' ontological commitments with regard to the mind and the body, (2) hypothetical questions that relate to the possible spatial and temporal attributes of bodyless minds (3) hypothetical questions that relate to the possible perceptual and cognitive attributes of bodyless minds. Each of these conceptual subdomains was probed with multiple questions, and the students' responses were examined, coded in the binary format and used to train a self-organizing map for visualization.[7]

To summarize the results, the overall structure of the map suggests that the students do not share a clear and coherent set of beliefs on the spatiotemporal attributes of an immaterial mind, whereas in the case of sensory and cognitive capacities they are quite consistent. The question of internal coherence of this recurring set of belief remains an open question, however. The SOM map cannot address this question directly. However, it does show that if the students are incoherent, they are consistently incoherent in the same way.[7]

References

[1] M. Cole and Y. Engeström. *Distributed Cognition*, chapter A Cultural-Historical Approach to Distributed Cognition, pages 1–47. Cambrigde University Press, 1991.

[2] P. Gärdenfors. *Conceptual Spaces.* MIT Press, 2000.

[3] K. Hakkarainen, T. Palonen, S. Paavola, and E. Lehtinen. *Communities of networked expertise: Professional and educational perspectives.* Elsevier, 2004.

[4] Timo Honkela, Teemu Leinonen, Kirsti Lonka, and Antti Raike. Self-organizing maps and constructive learning. In *Proceedings of ICEUT'2000, International Conference on Educational Uses of Communication and Information Technologies*, pages 339–343. IFIP, 2000.

[5] Timo Honkela, Ville Könönen, Tiina Lindh-Knuutila, and Mari-Sanna Paukkeri. Simulating processes of concept formation and communication. *Journal of Economic Methodology*, 15(3):245-259, 2008.

[6] Tiina Lindh-Knuutila, Juha Raitio, and Timo Honkela. Combining self-organized and Bayesian models of concept formation. In J. Mayor, N. Ruh, and K. Plunkett, editors,*Connectionist Models of Behaviour and Cognition II Proceedings of the Eleventh*

Neural Computation and Psychology Workshop, number 18 in Progress in Neural Processing, pages 193-204. World Scientific, April 2009.

[7] Anna-Mari Rusanen, Otto Lappi, Timo Honkela, and Mikael Nederström. Conceptual coherence in philosophy education - visualizing initial conceptions of philosophy students with self-organizing maps. In B. C. Love, K. McRae, and V. M. (Eds.) Sloutsky, editors, Proceedings of the 30th Annual Conference of the Cognitive Science Society, pages 64-70, Austin, Texas, 2008.

[8] J. Lindblom and T. Ziemke. Social situatedness of natural and artificial intelligence: Vygotsky and beyond. *Adaptive Behavior*, 11(2):79–96, 2003.

[9] L. Vygotsky. *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, Massachusetts, 1978. (Originally published in 1934).

[10] L. Vygotsky. *Thought and Language*. MIT Press, Cambridge, Massachusetts, 1986. (Originally published in 1934).