

SIX NECESSARY QUALITIES OF SELF-LEARNING SYSTEMS

A Short Brainstorming

Gabriele Peters

University of Hagen, Human-Computer-Interaction, Universitätsstr. 1, 58097 Hagen, Germany

Keywords: Self-learning systems, Autonomous learning, Machine learning, Emergence, Self-organization.

Abstract: In this position paper the broad issue of learning and self-organisation is addressed. I deal with the question how biological and technological information processing systems can autonomously acquire cognitive capabilities only from data available in the environment. In the main part I claim six qualities that are, in my opinion, necessary qualities of self-learning systems. These qualities are (1) hierarchical processing, (2) emergence on all levels of hierarchy, (3) multi-directional information transfer between the levels of hierarchy, (4) generalization from few examples, (5) exploration, and (6) adaptivity. I try to support my considerations by theoretical reflections as well as by an informal introduction of a self-learning system that features these qualities and displays promising behavior in object recognition applications. Although this paper has more the character of a brainstorming the proposed qualities can be regarded as roadmap for problems to be addressed in future research in the field of autonomous learning.

1 INTRODUCTION

This article is about the old but still not answered question how information processing systems, be it biological systems or be it technological systems (see figure 1), can autonomously develop cognitive capabilities such as perception, recognition, reasoning, planning, decision making, and finally goal-oriented behavior. Given the inherent structure of the body respectively the hardware, the question is:

How can these cognitive capabilities be learned from nothing else but from data acquired from the environment?

Obviously the human brain is able to accomplish this task, and it performs better than any existing artificial system. Thus, to investigate this question a reasonable approach consists in the attempt to model artificial systems according to the human example. Answers to this question are beneficial not only for basic research in the field of cognitive sciences but also from an engineering point of view. Although I think that computer systems, which are engineered with the purpose to serve human needs, do not have to function necessarily and exclusively like the human brain, I am of the opinion that the investigation of the general possibility of emergent cognitive capabilities can bring out a large benefit for the development of technology.

This collection of thoughts is written from a basic researcher's point of view. Any self-learning system, either biological or technological, has to feature some necessary qualities to be able to display emergence of cognitive capabilities. In section 2 I specify six such qualities which in my opinion are necessary. Solving the problem of building an artificial system which displays the same learning capabilities as the human brain, means in my opinion that at least these qualities have to be modelled. As in existing machine learning approaches only some of the problems connected with these qualities can be regarded as being solved, section 2 can thus be read as a roadmap for problems to be addressed in future research in the field of autonomous learning.

The postulated qualities are supported in section 3 where I describe a self-learning system we have developed which features these qualities and shows promising behavior. The description is presented in a colloquial form. Technical details of this system are explained in the original research references given¹.

¹The list of necessary qualities may not be complete (i.e., these qualities may not be sufficient), but we obtained (at least weak) supporting results with a system that features these six qualities.



Figure 1: Left: Biological self-learning system. Right: Artificial self-learning system.

2 SIX QUALITIES

To accomplish tasks such as reasoning or decision making, a system needs to have knowledge about the world. One unsolved problem is the problem of getting knowledge into a system. I believe, that a model of the world can arise only in a system that operates in a hierarchical manner (Q1) - with the additional qualities Q2 to Q6 explained in the following.

2.1 Q1: Hierarchical Learning with Different Mechanisms on Different Levels

The need for a hierarchy is given, on the one hand, by the fact that representations of the world learned on a low-level are not sufficient if it comes to tasks like reasoning. On the other hand, high-level (i.e., more abstract) representations, that allow for higher cognitive skills, cannot be learned on a purely statistical basis. Thus, if learning takes place in a hierarchical manner, then probably different methods for knowledge acquisition are needed for different levels of learning. Whereas on a lower level a statistical approach for learning features has been proven to be sufficient, this does not necessarily hold true for higher levels as well².

Low-level Learning. On a low-level plenty of features are available and they are simple. Reinforcement learning or probabilistic approaches such as statistical pattern classifiers are well suited to extract structure from the incoming data. Representations obtained by low-level learning are *numerical* descriptions on a *feature level*. What is learned can also be regarded as *implicit* knowledge (see section 3).

High-level Learning. There are some reasons for the necessity of high-level learning mechanisms. One

²I leave the question of the number of necessary levels in the hierarchy open. Of course, there is no reason why there should be two levels only. But for the sake of simplicity I will assume only two levels for the remains of these descriptions.

reason is that because of limited memory capacities and limited computational capacities of systems (i.e., brains as well as computers) harder problems such as the interpretation of a scene cannot be solved by statistical means only. Furthermore, if the complexity of a problem increases with the number of relevant dimensions, the experiences of examples are too rare for the system to be sufficient for a statistical approach. In a purely statistical approach a scene would for instance be represented by a single point in a very high-dimensional space. Third, if cognitive capabilities such as reasoning, planning, or anticipation should be displayed by a system, the numerical descriptions of the world have to be interpreted semantically. Thus, there is the need for more *symbolic* representations than those obtained by low-level learning. In addition, also the necessary quality Q4 of self-learning systems, i.e., the ability to generalize to more abstract knowledge, especially from few examples only, probably requires more advanced, e.g., rule-based learning mechanisms. In contrast to the above mentioned feature level, we can speak about the high level also as a *rule level*. What is learned on a higher level can be regarded as *explicit* knowledge (see section 3).

2.2 Q2: Learning Mechanisms have to Emerge for all Levels

Not only the mechanisms of low-level learning should emerge in an autonomous way. Here machine learning methods already exist which display this quality, e.g., reinforcement learning approaches. My intention is the explanation of the emergence of cognitive capabilities from data acquired from the environment only. Thus, also the mechanisms of high-level learning have to emerge. In my opinion it is reasonable to argue that symbolic representations of the world can be learned via the numeric representations acquired on a lower level. In section 3 I give an example how this can be achieved.

2.3 Q3: Multi-directional Transfer of Information between the Levels

Postulating a hierarchical approach to self-learning systems does not automatically determine the direction of information transfer within the hierarchy. In my opinion it is not sufficient for the acquisition of an appropriate world model to assume an incremental way only, i.e., a one-directional bottom-up process. Rather I believe that a constant exchange of information between low and high levels is necessary. So, I propose to endow self-learning systems with a

bottom-up emergent process as well as with a top-down process which will, after a while of runtime, be able to guide further knowledge acquisition via the low level.

Bottom-up Emergence. The need for a bottom-up emergent process should be obvious as also the high-level mechanisms are supposed to emerge. From, for instance, feature learning on the one end until scene understanding on the other end there has to be one point in the hierarchy where the acquired knowledge has to be transferred from a numerical, implicit form to a symbolic, explicit form which allows for reasoning. In section 3 I give an example for the general possibility that also such symbolic representations (especially in the form of rules) can be learned autonomously during runtime.

Top-down Guidance. Once one first (probably simple) symbolic representation exists it can be utilized to restrict or guide the low-level learning process. Top-down guidance of low-level processes probably plays a role in several perceptual phenomena in humans such as attention and selective perception, where (high-level) expectations affect the selection of perceived data³.

2.4 Q4: Generalization from Few Examples

Humans are very good at generalizing from few examples, generating abstractions, or making extrapolations from their experiences. Any self-learning system thus should allow for generalizations from its experiences to establish more general hypotheses on the world. Again, the fact of limited memory capacities is one argument for the necessity of this quality of self-learning systems. Without the ability to generalize a system would have to memorize all past experiences. And even this strategy would not necessarily yield usable knowledge as for harder problems such as scene interpretation with high-dimensional features the lack of enough observable examples inhibits or

³“Top-down guidance” in the context of cognitive sciences may correspond to “top-down control” in the context of engineering. As rule-based descriptions of the world are potentially more comprehensible by humans than numerical representations the learned explicit knowledge can be utilized by designers or operators of an autonomous system to observe, understand, and control its behavior. The higher levels of a self-learning system constitute potential gateways for human-computer interaction. Thus, this quality of self-learning systems probably complies with requirements of engineers who often urge for mechanisms to keep a system under control.

at least massively delays the emergence of an appropriate world model. Although mechanisms of generalization for different levels of learning presumably work in a different way (see Q1) learning from few examples should occur on all levels. Given symbolic representations, for instance, one new example may be able to prove or disprove rules (i.e., rules that comprise *exist* or *for-all* statements, respectively), or at least contribute to the belief or disbelief in such a rule. In section 3 I give an example for the ability of a learning system to alter its behavior in reaction to the experience of few examples.

2.5 Q5: Exploration

As another necessary quality I regard the capacity of a system to explore its environment, even if a sophisticated world model already has been acquired. “Self-learning” means learning during runtime, i.e., learning while the system interacts with the environment (in contrast to learning from preset, given datasets).

One reason for the necessity to explore the environment is the acquirement of knowledge from scratch. In the beginning, when high-level mechanisms are not yet sufficiently developed, hypotheses about the world will be rather unreliable. So, the system has to act and wait for the consequences its action brings to verify and modify the hypotheses. In the beginning of the learning process, i.e., as long as no adequate world model has been obtained, the explorative part of the behavior of a self-learning system should probably be larger than the part of the behavior which exploits (i.e., relies on) obtained knowledge. With increasing knowledge the exploration may be reduced. In section 3 I describe how a model can be learned from scratch by exploration.

A second reason for the necessity of exploration is the verification of an already obtained (and possibly sophisticated) world model. This reason is strongly connected to Q6 (adaptivity). To be able to keep up with changing environmental conditions it is necessary for a system to verify existing knowledge by exploration, probably both, on a regular basis (i.e., intermittently without external trigger) as well as on an experience-based scheme (i.e., with increasing exploration when hypotheses prove to be progressively invalid).

2.6 Q6: Adaptivity

Learned world models should not be allowed to be stationary, i.e., they should be subject to constant change for at least two reasons. The first reason consists in the possibility that the world model learned so



Figure 2: Left: A situation in which implicit learning takes place. Humans are not aware of the learning process, and they acquire procedural knowledge such as the sequence of actions during a dinner. Right: A typical situation in which explicit learning takes place. Here declarative knowledge is acquired. Humans are aware of the learning process and they can verbalize what they have learned, for instance grammar rules of a foreign language.

far may be insufficient, maybe because the system has not yet been exposed to appropriate examples. The second reason is the fact that the environment underlies constant change as well, both, because the learning system meshes with the environment via its actions, and for external reasons. In my opinion, the only way for a self-learning system to keep up with the environmental conditions is exploration, as mentioned above in an intermittent and in an experience-based way (when actions turn out to be disadvantageous).

A Final Remark: *Don't Start Simple in the First Place!*

For this remark I will shortly leave the basic researcher's point of view, take an engineering point of view, and briefly shift the attention from completely unsupervised learning to (at least in parts) supervised learning. In the introduction I formulated the question, how learning systems can acquire cognitive capabilities from nothing else but data perceived in the environment. Of course, even biological systems do not learn *exclusively* from scratch but *also* by supervision. In case a learning system should not be *exclusively self-learning*, supervision or training can certainly foster and accelerate the learning process. Often the learning of a foreign language is given as example here.

If the teaching of a system is intended, one approach consists in the consecutive, incremental exposure of the system to environments with increasing complexity. In such approaches learning (and teaching) starts with a selection of input data that allows for the acquisition of a very simple but appropriate model (not of the world but of the highly restricted world). One part of the supervision here consist in this selection.

In my opinion, however, one part of the learning problem humans are able to solve, is the *autonomous* selection of those data from the environment, that

are simple enough to allow for the construction of a first, simple world model - in the beginning and on a low learning level. This is implicit learning (see subsection 3.1). In foreign language learning, for instance, the supervision takes place already on a higher learning level. That is explicit learning (see subsection 3.1). From my point of view, "starting simple" makes sense only for explicit learning, not in the first place for implicit learning. As explicit knowledge probably always emerges only after first implicit representations have been acquired, my proposition at this point thus is the following:

For pure *self-learning* systems anyway no preselection of the data should occur. But also if supervision to a certain extend should be permitted, I believe that a learning system should nevertheless be thrown into the full complexity of the world and aim at the acquisition of a first (simple) world model from the structure of the full available data, not just from a selection of them. I guess, that the training of a learning system is most fruitful if it takes place at the earliest on a higher hierarchical level of learning. In section 3 I give an example for a learning system that in principle allows for supervision on a high learning level by utilizing the symbolic representation, which takes the form of rules.

3 SUPPORTING SIMULATIONS

In subsection 3.2 I roughly describe our self-learning system in a colloquial way and I point out how the postulated qualities are implemented, after I briefly illustrate one main inspiration for our approach in subsection 3.1. This section concludes in subsection 3.3 with some results obtained with the proposed system in a concrete computer vision application.

3.1 Inspiration

In our approach we implemented two levels of learning. This is inspired by psychological findings which support a two-level learning model for human learning (Sun et al., 2005). In psychological terminology, on the lower level, humans learn *implicitly* and acquire *procedural* knowledge. They are *not aware* of the relations they have learned and can hardly put them into words. On the higher level humans learn *explicitly* and acquire *declarative* knowledge. They are *aware* of the relations they have learned and can express them, e.g., in form of *if-then* rules. These two levels do not work separately. Depending on what is learned, humans learn top-down or bottom-up (Sun et al., 2007). In completely unfamiliar situa-

tions mainly implicit learning takes place and procedural knowledge is acquired. The declarative knowledge is formed afterwards. This indicates that the bottom-up direction plays an important role. It is also advantageous to continually verbalize to a certain extent what one has just learned and so *speed up* the acquisition of declarative knowledge and thereby the whole learning process (see figure 2).

3.2 A Self-learning System

The system I will now briefly introduce is intended for learning adequate behavior based on simple features it perceives in the environment. We combine two very different approaches from opposite ends of the scale of machine learning techniques. Low-level learning is realized by reinforcement learning (RL), more specifically $Q(\lambda)$ -learning (Sutton and Barto, 1998), high-level learning is realized by techniques of belief revision (BR) (Spohn, 2009). In figure 3 the systems' functionality is illustrated. Technical details are given in (Leopold et al., 2008a). By the combination of RL and BR techniques the system is able to adjust much faster and more thoroughly to the environment and to improve its learning capabilities considerably as compared to a pure RL approach. In the following I will address the realization of the single postulated qualities in our system.

Hierarchy (Q1) is obviously implemented as explained before.

Emergence on All Levels (Q2) is given as well. For the lower level it is given inherently by employing RL. But also on the higher level a world model (in the form of *if-then* rules) *emerges* as the generation of rules is driven by the numerical representation (in the form of values of state-action pairs) that arises on the lower level. Also from the BR point of view alone, this construction is interesting. One drawback of BR techniques consists in the fact that it is often difficult to decide *which parts* of existing rules (i.e., which parts of logical conjunctions (see Q4)) should be given up when a new belief comes in, in such a way that no inherent contradictions are introduced. In this context the rewards obtained in the RL context can be regarded as measures for the correctness of parts of the rules learned so far.

Multi-directional Transfer (Q3) is given by stages 3 and 7 of a learning step (see figure 3). At stage 3 (*top-down guidance*) the system uses current beliefs to restrict the search space of actions for the low-level process. At stage 7 (*bottom-up emergence*) feedback to an action in the form of a reward is used to acquire specific knowledge from the most recent experience by which the current symbolic knowledge

is revised. While the implementation of the top-down guidance in our system is straightforward the probably more important bottom-up process is the most delicate part of our architecture. The ultimate revision of the ranking function by new information is indeed realized using standard techniques of BR. The challenge however consist in the formalization of the new information (here: which are the best actions in a given state from a RL point of view) in such a way that it can be utilized by BR techniques.

As revisions of the symbolic knowledge have a strong influence on the choice of future actions they have to be handled carefully, i.e., the system should be quite sure about the correctness of a new rule before adding it to its belief. For this reason we chose a probabilistic approach to assess the plausibility of a new rule. We use several counters counting, for instance, how often an action has been a best action in a specific state, before the symbolic representation is adapted. Thus, stage 7 is not necessarily carried out in each step of the learning process but only after enough evidence for a revision has been obtained from the lower learning level.

Generalization from Few Examples (Q4) is facilitated in our system by the introduction of the BR component. In general, the possibilities to generalize from learned knowledge to unfamiliar situations are more diverse with BR than with RL techniques. In our approach rules take the form of conjunctions of several, multivalued literals. For example, a rule such as "If $A=x$ and $B=y$ then perform action z " would be represented by the conjunction $(A = x) \wedge (B = y) \wedge (Action = z)$. This allows the definition of similarities between conjunctions, e.g., simply by counting how many literals with the same values they share. Revisions of existing rules can then be based on the similarities between conjunctions. Thus generalization can easily occur by revising similar rules in one single learning step. In Q4 I claimed that generalizations should be possible from few examples only. In principle, it is possible in our approach that a single experience (in form of a reward given for an action) could cause the introduction of a new rule into the symbolic world model. In practice, the necessary number of supporting examples can be adjusted by tuning the relevant parameters of stage 7.

We address the topic of generalizations in the context of BR in more detail in (Häming and Peters, 2010) and (Häming and Peters, 2011b). In (Häming and Peters, 2011c) we propose a method to exploit similarities in symbolic descriptions especially in the case of high-dimensional spaces.

The quality of **Exploration (Q5)** in our approach is implemented by the RL component, with a larger

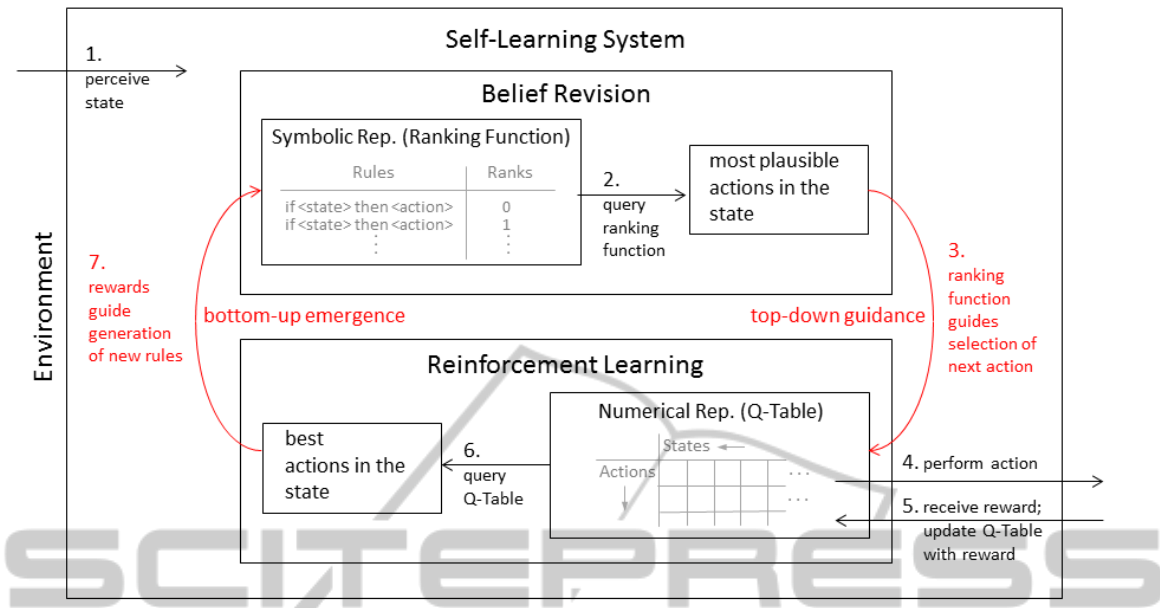


Figure 3: One single learning step of the proposed self-learning system. Displayed are 7 stages of one learning step. Before the execution of the first step of a whole learning process neither any symbolic knowledge (in the form of rules) nor any numeric knowledge (in the form of values of state-action pairs in the Q-table) has been acquired. 1. Signals perceived in the environment (i.e., the current state of the environment) have a dual representation: a symbolic one and a numerical one. BR perceives the symbolic representation, RL the numerical one. 2. Query the ranking function of BR about the most plausible action in the given state. (The ranking function assigns a degree of plausibility or belief to each rule, that has emerged so far.) 3. (Top-Down Guidance) Here the ranking function acts as a filter for the Q-table of RL for the selection of the next action: Look up actions in the Q-table with best values for the state, but search only among those actions that have been ranked plausible in the preceding stage. 4. Chose a random action among those actions that have been left over by the third stage and perform this action. 5. (Pure RL) Receive a reward from the environment and update the Q-table with this reward. 6. Query the (updated) Q-table about the best actions in the given state. 7. (Bottom-Up Emergence of High-Level Mechanisms) The ranking function of BR is revised (i.e., the plausibility of existing rules is adapted or new rules are generated) with the aim to make those actions most plausible in a given state that have the greatest values in the Q-table, i.e., that have been rated advantageous on the lower learning level via the received rewards.

part of exploration in the beginning of the learning process then with advancing progress. Though, also with an already established world model the environment is consistently checked by random explorations. Exploration in the proposed system is responsible for the acquirement of knowledge from scratch in the numerical form, as well as in the symbolic form. We have not implemented yet an increasing exploration in reaction to (repeatedly disadvantageous) experiences.

Adaptivity (Q6) is given inherently in our system by the explorative character of the RL component. Thus, the learned numerical as well as symbolic representations underlie a constant change. As will be described in the subsection below, even rules that have been formulated as conditionals by an operator and that have been used for a revision of the ranking function before the learning process started, can be discarded by the system after it made some inconsistent experiences during exploration.

3.3 Example Application

We have investigated the capabilities of our learning system in applications from the domain of computer vision (Leopold et al., 2008b; Häming and Peters, 2011a). Given a variety of unfamiliar objects, the system should learn to rotate an unfamiliar object until it can (visually) recognize it. The left part of figure 1 shows a biological system, which is able to accomplish this task. Instead of rotating an object in front of a fixed camera we modelled a camera moving around an object, as shown in the right part of figure 1.

The sensory input of the system comprises simple visual features, e.g., form and texture attributes of the objects. In more detail, the system is able to identify the values of attributes such as the shape of the front view, the size of the side view, or the complexity of the texture.

The behavioral output of the system is a sequence of actions, e.g., rotations, that allow for an efficient

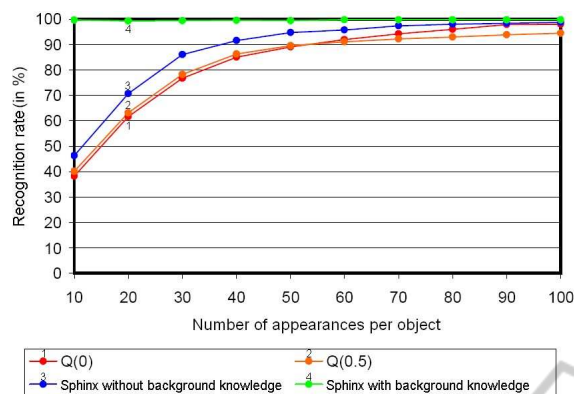


Figure 4: Recognition rates in the computer vision application depending on the number of appearances per object. The red curves are the rates for two different, pure RL learning systems, the blue curve shows the recognition rates for the system illustrated in figure 3, and the green curve displays the rates for the same system but now endowed with a priori knowledge (see end of subsection 3.3).

recognition of the objects. *Efficient* means that the system should perform the least number of rotations per object to successfully recognize it. In more detail, the actions can be a) rotations to three views per object: the front, the side, and an intermediate view and b) recognize actions. The recognize actions are either the recognition of the different objects, such as *bottle* or *tree*, or the refusal to recognize anything.

On the higher learning level the system starts with no rule at hand. On the lower learning level the system is rewarded with a value of -1 for each rotation action, a value of 10 for each correct recognition action, and a value of -10 for each false recognition action. It starts with arbitrary rotations (Q5), it can perform several sequences of actions with each object, and a sequence ends after a recognition action or after 10 rotation actions. Three of the rules learned in this way are the following:

- If the shape of the front view is an upright triangle and the size of the side view is large then recognize a bottle.
- If the shape of the front view is a circle and the shape of the side view is unknown and the texture is simple then rotate the object to the left.
- If the texture is complex then do not recognize a bottle.

In one of the simulations we endowed the system with a priori knowledge before the learning process started, i.e., we revised the ranking function with the rule, to look at all three views before a recognition action. At the end of the learning process the system has learned that this rule is disadvantageous and that

it can achieve even higher recognition rates when it looks at only two views (Q6) (see figure 4).

The technical details of this application are described in (Leopold et al., 2008b). In another application we showed the general applicability of our system to learn a scanning strategy for accessing such views of a 3D object that allow for a discrimination of the object from a very similar but different object (Häming and Peters, 2011a).

ACKNOWLEDGEMENTS

This research was funded by the German Research Association (DFG) under Grant PE 887/3-3.

REFERENCES

- Häming, K. and Peters, G. (2010). An Alternative Approach to the Revision of Ordinal Conditional Functions in the Context of Multi-Valued Logic. In *20th International Conference on Artificial Neural Networks (ICANN 2010)*, pages 200–203. Springer-Verlag.
- Häming, K. and Peters, G. (2011a). A Hybrid Learning System for Object Recognition. In *8th International Conference on Informatics in Control, Automation, and Robotics (ICINCO 2011)*.
- Häming, K. and Peters, G. (2011b). Improved Revision of Ranking Functions for the Generalization of Belief in the Context of Unobserved Variables. In *International Conference on Neural Computation Theory and Applications (NCTA 2011)*.
- Häming, K. and Peters, G. (2011c). Ranking Functions in Large State Spaces. In *7th International Conference on Artificial Intelligence Applications and Innovations (IAI 2011)*.
- Leopold, T., Kern-Isberner, G., and Peters, G. (2008a). Belief Revision with Reinforcement Learning for Interactive Object Recognition. In *18th European Conference on Artificial Intelligence (ECAI 2008)*, pages 65–69.
- Leopold, T., Kern-Isberner, G., and Peters, G. (2008b). Combining Reinforcement Learning and Belief Revision - A Learning System for Active Vision. In *19th British Machine Vision Conference (BMVC 2008)*, pages 473–482.
- Spohn, W. (2009). A survey of ranking theory. In *Degrees of Belief*. Springer.
- Sun, R., Slusarz, P., and Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review*, 112(1):159–192.
- Sun, R., Zhang, X., Slusarz, P., and Mathews, R. (2007). The interaction of implicit learning, explicit hypothesis testing learning and implicit-to-explicit knowledge extraction. *Neural Networks*, 20(1):34–47.
- Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge.