A Multi-Criteria Exemplar Model for Holistic Categorization in Autonomous Agents

Samer Schaat, Alexander Wendt, and Dietmar Bruckner
Institute of Computer Technology
Vienna University of Technology
Vienna, Austria
{schaat, wendt, bruckner}@ict.tuwien.ac.at

***Abstract*—When developing a model of perceptual categorization, a functional and holistic approach is required. Therefore, considering the influence of an agent's current needs, perceptual categorization is modeled as the valuation of a stimulus regarding its potential to satisfy an agent's needs. Based on an exemplar model, an activation-based model is developed that considers two kinds of categorization criteria, namely perceptual similarity and expectation, which represents bottom-up and top-down approaches of perception, respectively. The aggregation of these two criteria then corresponds to the integration of these approaches into a holistic model. A simplified simulation of the categorization model shows that the multi-criteria model leads to a more confident valuation of stimulus objects. In particular, the model is able to eliminate ambiguity in perceptual categorization. Although used equivalently, the simulation shows that the bottom-up similarity criterion is more significant than the top-down expectation-based criterion. In particular, the latter is only significant in case of ambiguity in an object's appearance.**

Keywords—Cognitive Architectures; Perception; Categorization; Top-down perception; Expectations; Agent-based Modeling; Artificial Intelligence; Cognitive Modeling

I. INTRODUCTION

The categorization of perceived objects is an important premise to cope with the world. In developing a human-inspired model of perceptual categorization for a cognitive architecture some key aspects have to be considered, which motivate this paper's work.

As in humans, a cognitive architecture should distinguish two qualities of information processing, which follow different rules and conditions. Based on the terms used in humans one can call them unconscious and conscious processes. The consideration of different methods of information processing is particularly relevant for perceptual categorization. As mentioned in [1] and elaborated in [2], a single categorization method is not sufficient to represent all facets of categorization. In conventional approaches of Artificial Intelligence (AI) the mechanisms of unconscious processing are often neglected although, since popularized by Freud [3], the significance of unconscious processes in human cognition and perception is emphasized repeatedly [4], [5]. Amongst other purposes, unconscious processes are preparatory work for conscious processes. This aspect is also indicated by Freud's theory of the human mind [3] by using the terms primary and secondary process for unconscious and conscious processes, respectively. An intensively discussed approach that considers the significance of unconscious processes in perception is top-down perception. This approach reflects the significance of subjective experience for perception, which is another key aspect that a model of perceptual categorization in a cognitive architecture has to consider.

Following the evolutionary principle, perceptual categorization is a mean to cope with the external world to satisfy an agent's needs. In this regard an agent does not need an objective image of the world, but only a subjective one that focuses on an agent's perspective, which is relative to its needs and experience and works to support the satisfaction of needs. Such a subjective approach treats perception as an active process, with the goal to satisfy an agent's needs.

With the agent's experience as a basis to cope with perceived objects and the fulfillment of its needs as the top concern, a holistic approach is needed to consider these aspects. Such an approach requires the integration of subjective and functional factors in perceptual categorization and in an agent's cognitive architecture. To reach this goal, objective and subjective influences have to be considered in perceptual categorization.

Hence, the problem at hand consists of finding a model of perceptual categorization for assessing perceived objects using associated memories and internal influences. In this paper the ARS (Artificial Recognition System) [6] model is used, particularly because it provides an appropriate cognitive architecture to consider above mentioned aspects. In ARS, categorization is considered as a multi-level process, in particular as an unconscious comparison and a conscious reasoning process. In this paper a model for comparison-based categorization that harnesses the agent's memory for perceptual categorization in multiple ways and that integrate influences by using expectations is presented.

II. STATE OF THE ART

Two topics are particularly relevant for this paper's work, namely perceptual categorization and the consideration of categorization influences.

A. Perceptual Categorization

In the scope of this discussion of the state of the art, perceptual categorization is the reduction of sensory data in a

symbolic form, i.e. the stimulus, to a higher-level label, i.e. its category. A model of categorization then is concerned with "the representations and transformations that link the input and response representations" [7].

Categorization can be regarded as a reasoning, i.e. rule-based, or comparison, i.e. similarity-based, process. Rule-based models follow a classical definitional approach of categorization [1] that specifies sufficient and necessary properties for membership in a category [7]. In this case categories are mentally represented as definitions. An example of a rule-based model is shown in [8], where semantic nets are used to reason about features and relations of concepts. The comparison-based approach regards categorization as a function of similarity of a stimulus representation to category representations [9]. Popular examples of such approaches are exemplar [10], [1] and prototype [11], [1] models. The prototype approach introduced the idea of typicality to decide category membership, with a prototype being a summary representation of a category. Objects that are similar to the prototype are categorized as being typical members. Of course such prototype representations of a category are rarely real-world examples. Opposed to that, in exemplar models a category is represented by all concrete memorized exemplars of the category [7] and category membership is a function of similarity of an object to all known exemplars. The most influential [1] exemplar model is the General Context Model (GCM) [12]. Based on neurophysiological and empirical evidence there is an emerging consensus that exemplar-based models underlie important aspects of category representation [9] and are more powerful than prototype models [1], since they are able to represent categories comprehensively [7]. Recent hybrid models consider multiple category representations. An example is the SUSTAIN model [13], which considers both exemplar and prototype representations. Other hybrid models use exemplar- and rule representations [14]. Still, such hybrid models do not consider input additional to the stimulus. Furthermore they do not consider the parallel existence of multiple kinds of category representations and mechanisms and their interplay. As emphasized by [1] and [2] humans use different forms of category representations.

B. Top-down and Bottom-up Perception

When modeling perceptual processes two approaches may be used. In case of starting with sensory data as the primary source for the perceptual process, perception is regarded a bottom-up process. As opposed to that, in top-down perception the process is driven by prior knowledge as the primary source. These approaches implicitly raise the question how an interface between perception and cognition may look like, with different opinions emphasizing the primacy of bottom-up [15] or top-down processes [16], [17].

In both approaches perception can be interpreted as building a model of the world, with prior knowledge, i.e. memories, as the point of departure in case of a top-down process, or sensory data in case of a bottom-up process. In the former case the constructed model is used to drive perception via expectations [18], which are formed based on memory, and the sensory input is only used complementary. Hence, in a top-down approach memories trigger predictions by activating associated representations, which may be associated based on perceptual, conceptual or functional similarities. In particular, an activated representation would trigger the expectation of associated representations.

C. Priming

The concept of priming focuses on the process of influencing perception and therefore of perceptual categorization. It is defined as "… a nonconscious form of memory that involves a change in a person's ability to identify, produce or classify an item as a result of a previous encounter with that item or a related item" [19]. In the terms of priming, a target item is primed by a prime, i.e. a previous encounter with that item or a related item. Forms of priming include perceptual, conceptual and affective priming. In conceptual priming items are primed by primes with a similar meaning, e.g. objects from the same category. The influence of affective priming is shown in the faster categorization of a target if it is primed by an item of the same valence [20]. Another form of affective priming is shown by the influence of mood-congruent primes [21]. When considering a connection between an agent's bodily needs and its affective condition, the impact of drives on perception shows patterns of affective priming.

A model often used for the representation of priming is spreading activation, [22], where activation is spread from a prime to associated target items. Such activation of associated memory is assumed to occur unconsciously [21].

III. ARS MODEL

Conventional approaches in AI that use task models have not lead to artificial systems that are able to cope with complex tasks that are managed by humans on a daily base. The ARS approach develops a process model of the human cognitive architecture by using a top-down design approach. Due its consideration of holistic and functional aspects, the second topographical model of [3] is chosen as a framework to build a cognitive architecture, since it is the most appropriate unitary cognitive theory of the human mind that considers top-down and functional aspects [23]. Using a functional model, the ARS project follows a generative approach with the focus of describing functions that generate behavior instead of building a behavior model. This approach complies with the Artificial General Intelligence (AGI) approach to develop broad human-like intelligence that is able to cope with complex and dynamic problems rather than with narrowly and well-defined domains.

The ARS model's topmost level in the top-down design process is the second topographical model's three abstract functional units Id, Ego and Super-Ego. From there, together with psychoanalytical consultants, a finer grained, more detailed description of the functions in the ARS model is generated with each new level of description. The third level of the ARS functional model is shown in Fig. 1.

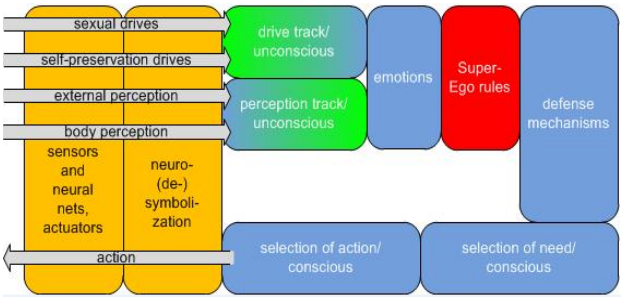

Fig. 1. Third level of the ARS functional model.

As indicated in Fig. 1 by the coloring, the functional model distinguishes models of the three functional units of the second topographical model (green: Id, red: Super-Ego, blue: Ego) and the separation of the primary (upper part) and secondary process (lower part), which refer to unconscious and conscious processes, respectively. Additionally, it considers three layers of processing, inspired by neuro psychologists, namely neural processing (left), neuro-symbolic processing (half-left), and the mental layer (the Freudian model). Psychoanalysis describes the primary process as structureless, which in this regard means that, opposed to the secondary process, it does not consider logical principles or order, e.g. temporal and spatial relations. The central principle of the primary process is the pleasure principle. In psychoanalytic terms this represents the dynamics of drive wishes: psychic activity aims for maximal satisfaction of a drive wish, which is the representation of an agent´s bodily need.

For this paper's topic the motivational and perceptual system of the ARS agent are particularly relevant and are described next.

A. *Drives*

The concept of drives is used in ARS as the central part of the agent's motivational system. Drives are the psychic representation of a bodily need. Through the impact of the drives on the agent's goals and behavior, the impact of the body, i.e. embodiment, and the agent's autonomy are considered.

A drive consists of a drive source, a drive object and a drive aim. The drive source is represented by the organ which signals the bodily need, the drive aim is the satisfaction of the drive by an act and the drive object is the object used in this act. Every drive source triggers two kinds of drives, named aggressive and libidinous drives. After the triggering of a drive by a bodily need, it is rated by a quota of affect, which reflects the organ's tension, i.e. the tension of the drive source. The higher the tension, the higher the bodily need, and subsequently the higher the drive's quota of affect. After the quantification of the bodily need, possible actions and objects are remembered satisfy the drive according to the agent's experience, and are subsequently used as potential drive aims and objects (see Fig. 2). In further steps the agent searches for possibilities to satisfy the drive in the external world, based on the remembered drive objects and aims. An example of a libidinous drive, which can be identified as "hunger" in the secondary process, is given in Fig. 2.

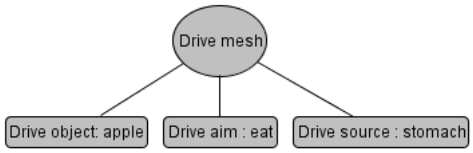

Fig. 2. Drive mesh structure.

B. *Perception and Recognition*

Perception in the ARS agent is a multi-step process that starts with the conversion of sensorial data to symbolic representations, which are called neuro-symbols in the ARS agent [24], since they reflect features of neurons and symbols and bridge the gap between activated artificial neurons and internal representations. The next perceptual step in the ARS agent is the recognition and categorization of the perceived object. Following a similarity-based approach incoming perceptions are matched against memory traces and activate them in case of a match. The concept of a memory trace in ARS states that memory content is stored in an associated manner and that memories are activated by associations of co-occurrence, similarity and accessibility [25].

The result of object recognition in ARS is a representation given by a thing presentation mesh [25], which is the associated representation of an object's physical attributes.

IV. Holistic Perceptual Categorization

When considering a functional and subjective approach and the principles of the primary process, exemplar models, priming and top-down perception are appropriate concepts that have to be integrated in a holistic model of perceptual categorization. The avoidance of abstraction and conceptualization in the primary process and the significance of drives and concrete memories favor an exemplar-based categorization model in the primary process. Therefore an exemplar categorization model is used to value perceived objects. Categorization influences, especially the impact of an agent's current drives, are integrated into the exemplar model by using the concept of top-down perception and priming (in the form of spreading activation). Following the pleasure principle, the purpose of perceptual categorization in the primary process is to determine the drive object categories of perceived objects, i.e. which drives may be satisfied by the perceived objects.

A. *Drive Object Categorization*

Perceptual categorization in the ARS agent's primary process is modeled as *integrated drive object categorization*, which is the valuation of perceived objects as drive objects. It is implemented by determining the *graded category membership* of a perceived object, where multiple category memberships are possible. In the ARS agent an exemplar is represented by a memorized drive object. Following a memory-based categorization approach, the primary goal of drive object categorization is the reduction of an agent's uncertainty in searching and selecting *appropriate exemplars*, whose categories are fully or partly used as categories of the stimulus. Therefore the exemplar's similarity to the stimulus and internal influences are used. To find appropriate exemplars

an *activation-based* approach, which is supported by the associative composition of the ARS data structures, is used. An exemplar‘s appropriateness for the stimulus’ categorization is dependent on its degree of activation. Two sources of activation are considered (see Fig. 3). On the one hand, an exemplar can be activated by the *similarity* to the stimulus. The more similar an exemplar is, the stronger it gets activated. On the other hand, an exemplar can be activated by the agent’s *expectations*. The stronger an exemplar is expected, the stronger it will be activated. Thus, internal influences are treated as expectations and used to activate appropriate exemplars. An overview of the overall process is given in Fig. 3. A particular relevant example of an expectation-based activation is the activation of *expected drive objects*, i.e. exemplars, by an agent’s current drives, in the scope of psychoanalysis called cathexis [3]. For instance, if the agent is hungry, stored drive objects, i.e. exemplars, are activated that are expected to satisfy hunger. The better these exemplars have satisfied hunger according to the agent’s memories, the more activation they get.

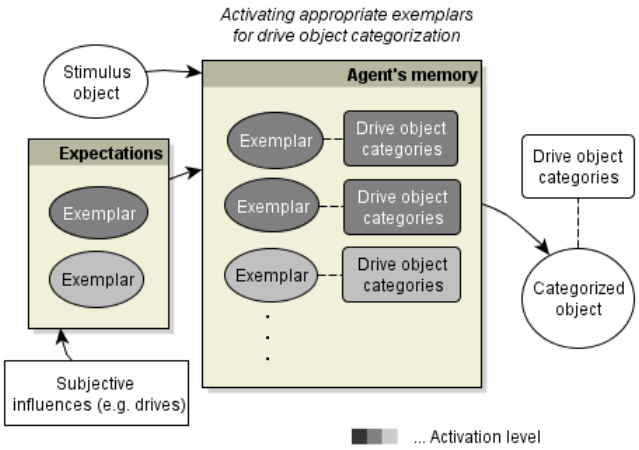

Fig. 3. Overview of integrated drive object categorization.

B. Integrated Categorization

The integration of categorization influences by using the concept of expectation is compliant with the principle of top-down perception and priming and is appropriate for a bionic and holistic model of perceptual categorization in a cognitive architecture. By using expectations, which are represented as *activation levels*, categorization influences are treated in a *generic* way. To integrate these categorization influences in the categorization process they are transformed to categorization criteria. For that reason and for the sake of a generic processing of all criteria, *activation sources* and *activation functions* are defined for each criterion. The activation function of a criterion activates exemplars that are associated with the sources, taking into account how much each criterion is fulfilled. Therefore *associative activation* is considered in the search algorithm that is used to find appropriate exemplars (see Fig 4 and 5). The better an exemplar fulfills the criteria, the stronger it gets activated. An example of an objective criterion is the *similarity criterion*, i.e. the similarity of an exemplar to the stimulus object. In this case the object’s properties are the activation sources (see Fig. 5) and a similarity function serves as activation function. An exemplar with the same properties as the stimulus object would get the maximal activation value. This case is called *appearance recognition*.

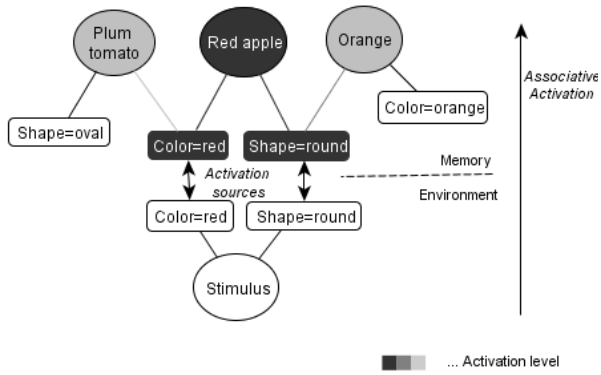

Fig. 4. Similarity activation.

The most significant subjective criterion in drive object categorization, which is derived from above described concept of cathexis, is called *embodiment criterion*. In this case the current drives serve as activation sources. The activation function considers an exemplar’s potential to satisfy the agent’s current drives. An exemplar’s fulfillment of this criterion is determined by comparing the agent’s current drives, particularly their quota of affect, with an exemplar’s associated drives, i.e. the drives the exemplar has satisfied according to the agent’s memories (see Fig. 5). An exemplar that would satisfy all current drives best would get the maximal activation.

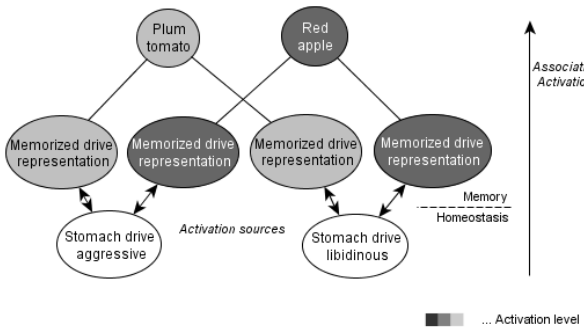

Fig. 5. Embodiment activation.

Hence, perceptual categorization in the ARS agent’s primary process is a *multi-criteria activation* process. To determine the overall activation of activated exemplars their criteria activation values have to be *aggregated* and hence *integrated* in the exemplar model. In the course of this process the criteria are weighted to consider their influence. A criterion activation value c is determined by

$$c = \frac{\sum_{i=1}^{n} s_i}{c_{max}}.$$

s_i…*i-th* source activation value
n … number of activation sources
c_{max}…criterion’s maximal activation value

The aggregated activation value a is given by

$$a = \frac{\sum_{i=1}^{n} w_i * c_i}{\sum_{i=1}^{n} w_i}.$$

c_i… *i-th* criterion activation value
n … number of criteria
w_i…criterion's weight

In the case of the similarity criterion the source activation values are given by the object properties' similarity to an exemplar's properties (see Fig. 4). In case of the embodiment criterion it is given by comparing the current drives with an exemplar's associated drives (see Fig. 5).

The calculation of a criterion's weight is depending on the criterion. In case of the similarity criterion it is given implicitly, i.e. the higher the similarity, the higher the impact of the criterion. In case of the embodiment criterion the criterion's weight is given by the aggregation of an agent's current drives. Due to properties of the pleasure principle and the need of normalization, the weight w_{i+1} is given by

$$w_{i+1} = w_i + (1 - w_i) * q_{i+1}$$

w_i… weighting factor after the consideration of i drives
q_{i+1}… $(i+1)$-*th* drive's quota of affect

Using this formula on the one hand normalization is considered in accumulating all drives' quota of affects, on the other hand low quota of affects do not distort the weighting factor (as e.g. when using a weighted mean value). After ranking the exemplars with respect to their overall activation value, the most appropriate exemplars are chosen for the categorization of the stimulus object. The agent's confidence in choosing appropriate exemplars is dependent on the activation level and its unambiguity. This leads to identified drive object categorization, i.e. the most appropriate exemplar's categories are used, or to generalized drive object categorization, i.e. generalization over the most appropriate exemplars (see Fig. 6). Both cases correspond to graded category membership, where multiple category memberships are considered.

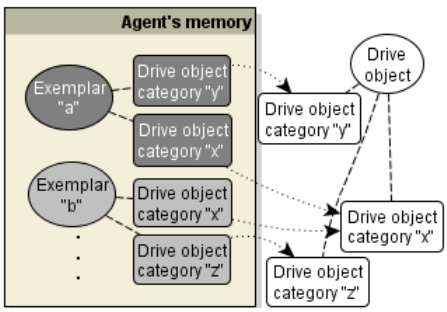

Fig. 6. Generalized drive object categorization.

V. Simulation

For the simulation and evaluation of the model a simplified use cases is defined to show the principle of drive object categorization. The goal is the drive object categorization of an apple. Different memories and internal states (i.e. current drives) lead to different scenarios and results. With the premise of appearance recognition following scenarios arise (see Fig. 7). Scenario I represents the default scenario, where the drive object categories are identified. Scenario II and III, where a memorized apple and tomato have the same object properties, show that in the case of appearance recognition ambiguity is still possible. In Scenario II the agent categorizes the stimulus object according to the categories of an apple, despite ambiguity in appearance.

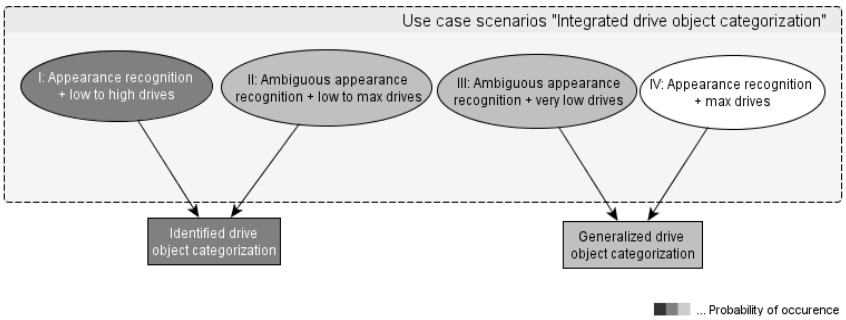

Fig. 7. Use case scenarios.

As shown in TABLE I. , the stimulus' similarity to the tomato exemplar is the same as the apple's, but because the apple would satisfy the current drives better, it is expected more than an tomato, and hence has a higher embodiment activation. The embodiment impact is sufficient to eliminate ambiguity and to identify the categories.

TABLE I. Categorization values scenario II

Categorization variables	*Apple exemplar*	*Tomato exemplar*
Category appropriatness	0.91	0.76
Similarity activation	1	1
Embodiment activation	0.76	0.34
Embodiment impact	0.58	0.58

In case of low current drives this is not the case, as scenario III reveals. Finally, scenario IV shows the distortion of drive object categorization by subjective influences, i.e. in case of maximal impact of the embodiment criterion. This is only the case when all drives are, which is exceptional; it would correspond to a form of hallucination, e.g. due to starvation, i.e. before the agent dies in the Artificial-Life-Simulation.

VI. Conclusion

The presented model shows how a multi-criteria exemplar model is used to value perceived objects and support the agent's needs. The simulation of the model show that the usage of multiple categorization criteria leads to the reduction of an agent's uncertainty in selecting appropriate exemplars, whose drive object categories can be used as the stimulus' categories. Hence, a holistic approach supports deciding the categorization process. The simulation also shows that the expectation-based

criterion is particularly significant in case of ambiguity in an object's appearance. Using the agent's current drives as trigger for expectations, the simulation shows the dynamic character of the expectation-based criterion. In particular, the impact of the criterion is dependent on the agent's dynamic drives. Generally, one can observe that the similarity criterion is more reliable to reduce uncertainty and that additional categorization criteria are significant only if appearance is weak or ambiguous.

One can observe that the consideration of similarity and expectation-based criteria and their integration in a model of perceptual categorization complies with aspects of top-down and bottom-up perception and their integration in a holistic approach.

The activation of appropriate exemplars by the agent's needs also emphasizes the directed and dynamic nature of drive object categorization. The usage of activation sources as point of departure in the search for appropriate exemplars enables a directed associative memory retrieval. This eliminates a common disadvantage of conventional exemplar-based models in reducing the search space.

VII. Future work

Possible future work includes the investigation of multiple expectation-based criteria. It is particularly interesting, if they are able to reach the similarity criterion's significance of reducing the uncertainty in drive object categorization. Examples of additional subjective influences that can be used as expectation-based categorization criteria are expectations that are triggered from planning and associative memory formation. The latter reflects contextual considerations in perceptual categorization.

Additional future work includes the interaction of the comparison-based exemplar model with a rule-based categorization model in the ARS agent's secondary process.

References

[1] G. L. Murphy, *The Big Book of Concepts.* Cambridge: MIT Press, 2002.

[2] R. L. Goldstone, "The role of similarity in categorization: providing a groundwork," *Cognition,* vol. 52, no. 2, pp. 125–157, 1994.

[3] S. Freud, "The ego and the id," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 19, London: Hogarth Press, 1923.

[4] J. A. Bargh, and T. L. Chartrand, "The unbearable automaticity of being," *American Psychologist,* vol. 54, pp. 462-479, 1999.

[5] A. Dijksterhuis, and L. F. Nordgren, „A theory of unconscious thought," *Perspective on Psychological Science,* vol. 1, no. 2, pp. 95-109, 2006.

[6] D. Dietrich, G. Fodor, G. Zucker, and D. Bruckner, *Simulating the Mind - A Technical Neuropsychoanalytical Approach*. Vienna: Springer, 2009.

[7] J. K. Kruschke, "Models of categorization, " in *The Cambridge Handbook of Computational Psychology,* R. Sun, Ed., New York: Cambridge University Press, 2008.

[8] J. L. McClelland, and T. T. Rogers, "The parallel distributed processing approach to semantic cognition," *Nature Reviews Neuroscience*, vol. 4, pp. 310–322, 2003.

[9] T. J. Palmeri, and I. Gauthier, "Visual object understanding," *Nature Reviews Neuroscience*, vol. 5, pp. 291–304, 2004.

[10] D. L. Medin, and M. M. Schaffer, „Context theory of classification learning," *Psychological Review*, vol. 85, no. 3, pp. 207–238, 1978.

[11] E. Rosch, "Cognitive representations of semantic categories," *Journal of Experimental Psychology*, vol. 104, pp. 192–233, 1975.

[12] R. M. Nosofsky, "Attention, similarity, and the identification-categorization relationship," *Journal of Experimental Psychology:General*, vol. 115, pp. 39–57, 1986.

[13] B. C. Love, D. L. Medin, and T. M. Gureckis, "Sustain: A network model of category learning," *Psychological Review,* vol. 111, pp. 309–332, 2004.

[14] J. R. Anderson, and J. Betz, "A hybrid model of categorization," *Psychonomic Bulletin & Review*, vol. 8, no. 4, pp. 629–647, 2001.

[15] Z. Pylyshyn, "Is vision continuous with cognition? The case for cognitive impenetrability of visual perception," *Behavioral and Brain Sciences,* vol. 22, no. 3, pp. 341-367, 1999.

[16] G. W. Humphreys, M. J. Riddoch, and C. J. Price, "Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy," *Phil. Trans. R. Soc. Lond. B,* vol. 352, no. 1358, pp. 1275–1282, 1997.

[17] B. H. Barlow, "The knowledge used in vision and where it comes from," *Phil. Trans. R. Soc. Lond. B,* vol. 352, no. 1358, pp. 1141–1147, 1997.

[18] M. Bar, "The proactive brain: memory for predictions," *Phil. Trans. R. Soc. Lond. B,* vol. 364, no. 1521, pp. 1181–1182, 2009.

[19] D. L. Schacter, I. G. Dobbins, and D. M. Schnyer, "Specificity of priming: a cognitive neuroscience perspective," *Nature Reviews Neuroscience*, vol. 5, pp. 853–862, 2004.

[20] J. D. Houwer, D. Hermans, K. Rothermund, and D. Wentura, "Affective priming of semantic categorisation responses," *Cognition & Emotion,* vol. 16, no. 5, pp. 643-666, 2002.

[21] J. Storbeck, and G. L. Clore, "The affective regulation of cognitive priming," *Emotion,* vol. 8, no. 2, pp. 208–215, 2008.

[22] A. M. Collins, and E. F. Loftus, "A spreading-activation theory of semantic processing," *Psychological Review,* vol. 82, no. 6, pp. 407–428, 1975.

[23] T. Deutsch, "Human Bionically Inspired Autonomous Agents - The Framework Implementation ARSi11 of the Psychoanalytical Entity Id Applied to Embodied Agents," PhD thesis, Institute of Computer Technology, Vienna University of Technology, 2011.

[24] R. Velik and D. Bruckner, "Neuro-symbolic networks: introduction to a new information processing principle" *Proceedings of the 6th IEEE International Conference on Industrial Informatics,* pp. 1042–1047, 2008.

[25] H. Zeilinger, "Bionically Inspired Information Representation for Embodied Software Agents," PhD thesis, Institute of Computer Technology, Vienna University of Technology, 2010.