

# Neuro-Symbolic Networks: Introduction to a New Information Processing Principle

Rosemarie Velik  
Vienna University of Technology, ICT  
velik@ict.tuwien.ac.at

Dietmar Bruckner  
Vienna University of Technology, ICT  
bruckner@ict.tuwien.ac.at

**Abstract-** Neural networks and symbolic systems are two different approaches to model cognitive functions. Both methods have certain strengths but also suffer from a number of weak points, which are however disparate. By combining characteristics of neural and symbolic processing, these weaknesses could be overcome. This paper introduces a new information processing principle based on so-called neuro-symbolic networks, which incorporates these two approaches. The utility of the suggested principle is outlined for applications in machine perception.

## I. INTRODUCTION

Both neural networks – also called connectionist systems – and symbolic systems have a broad field of applications. According to [7], neural networks and symbolic systems are two disparate approaches to model cognitive processes and engineer intelligent systems. Primary efforts have generally focused always only on one of these two disparate approaches.

Both approaches have their specific strengths and weak points: *Connectionist systems* are robust and can learn from examples. They are fault tolerant, can handle incomplete information, and are able to generalize to similar input. As they are parallel distributed systems, they also potentially provide increased speed of processing [14]. The drawbacks of neural networks lie in their incapacity to provide an explanation for the underlying reasoning mechanisms, wherefore they are considered as black box models [7], [13]. *Symbolic systems* can explain their inference process and use powerful declaration languages for knowledge representation. They allow explicit control, fast initial coding, dynamic variable binding, and knowledge abstraction [14]. Problems of symbolic systems are their lack of robustness as well as the inability to handle incomplete information and to generalize. They generally fail to learn new associations between symbols and to do things on their own [7], [13]. One basic problem with symbolic systems is the question how symbols get their meanings, because symbols and concepts need to be grounded somehow in reality [3].

A comparison of the characteristics of neural networks and symbolic systems shows that symbolic systems have certain problems that connectionist systems seem to solve and vice versa. However, although it seems quite obvious

that the weaknesses of connectionist and symbolic systems could potentially be overcome through a judicious integration of techniques and tools of both approaches, there has been only little cooperation between these two disciplines until now. By taking up the challenge to combine connectionist and symbolic approaches, a new method could be developed that shows the advantages of both without suffering from their weaknesses [7].

In this paper, a new method for cognitive information processing is introduced using so-called *neuro-symbolic networks*. The utility of neuro-symbolic networks will be outlined for the application domain of machine perception. The structural organization and information processing principles of neuro-symbolic networks are derived from neuroscientific and neuro-psychologic research findings about the human brain [8], [9]. The utility of the suggested principles is outlined for applications in machine perception. Therefore, the structural organization of the perceptual system of the brain is of particular interest [1], [2].

## II. NEURO-SYMBOLIC NETWORKS

### A. Neuro-symbols as Basic Information Processing Units

The basic information processing units of neuro-symbolic networks are so-called neuro-symbols. These neuro-symbols combine characteristics of neural as well as symbolic information processing.

The structure of a neuro-symbol has its paragon in the structure of biological neurons (see figure 1). A neuro-symbol has several inputs and one output. A neuro-symbol can receive input information from several other neuro-symbols which corresponds to the function of dendrites of neurons. The input information contains – among others – the activation grade of the symbol it originates from. The activation grades from all incoming neuro-symbols are summed up likewise in the cell body of a nerve cell. If this sum exceeds a certain threshold, the neuro-symbol is

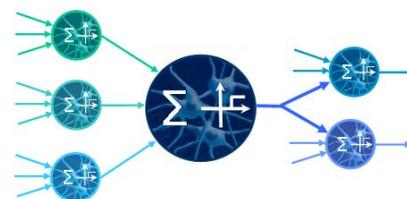


Figure 1: Structure of a neuro-symbol

activated. The information about the activation is passed to other neuro-symbols it is connected to in analogy to the axon of a neuron. If necessary, the input information representing the activation grades from connected neuro-symbols can be weighted. This corresponds to the synaptic connections between neurons. Similar as in the brain where many neurons are active at the same time, different neuro-symbols can process information in parallel.

To activate a neuro-symbol, information at the inputs does not necessarily have to arrive concurrently. There also exists the possibility to handle data arriving asynchronously or to consider certain sequences of data if they are of importance.

Like in symbolic systems, each neuro-symbol represents a certain symbolic meaning. For the application of machine perception, each neuro-symbol stands for a *perceptual image* representing e.g., a line, a person, a sound, a voice, a certain smell, etc. A neuro-symbol is activated, if the perceptual image it stands for is perceived in the environment.

Neuro-symbols can also have so-called properties that specify the perceptual image they represent in more detail. For example, a neuro-symbol “person” could have properties like “size”, “sex”, “age”, etc. By such properties, the number of different neuro-symbols being necessary for an application can be reduced. For perceptive tasks, a very important property of neuro-symbols is their “location” property, which specifies where in the surrounding a certain perceptual image was perceived.

### B. Combining Neuro-symbols to Neuro-symbolic Networks

#### Neuro-symbolic Network Architecture

To generate complex behavior, neuro-symbols have to be associated with other neuro-symbols. That way, neuro-symbolic networks emerge. However, a random interconnection of neuro-symbols does not lead to “intelligent” behavior. Their structural organization is of crucial importance. For perceptive tasks, which are the application envisioned in this paper, it is suggested to take the structural organization of the perceptual system of the brain as archetype.

Figure 2 illustrates the cerebral organization of the perceptual system of the human brain as described by



Figure 2: Cerebral organization of the perceptual system of the brain

Luria [8]. Perception starts with information coming from sensory receptors. This information is then processed in three levels, which are referred to as primary cortex, secondary cortex, and tertiary cortex. Each sensory modality of human perception has its own primary and secondary cortex. This means that in the first two levels, information of different sensory modalities is processed separately and in parallel. In the primary cortices, simple features like lines, color, movement, sounds of certain frequency, etc. are extracted from sensor data. In the secondary cortices, the features of each modality are combined and result in a unified unimodal perception of each modality. In the tertiary cortex, information coming from all sensory modalities is merged. These results in a unified multimodal perception. As can be seen from this description, the perceptual system follows a “simple to complex hierarchy” starting with the processing of simple unimodal perceptual features and ending with unified multimodal perceptual representations.

Following this principle of modular hierarchical information processing, neuro-symbols are arranged to neuro-symbolic networks in different modalities and different hierarchical levels (see figure 3). Information coming from different sensory modalities is first processed separately and in parallel before being merged. In a first processing step, so-called feature symbols are extracted from sensory raw data. Information processing in this level correlates with information processing performed in the primary cortex of the brain. This feature symbol layer can in fact consist of a number of sub-layers.

In the next two steps, feature symbols are combined to sub-unimodal and unimodal symbols. These two levels correspond to the function of the secondary cortex of the brain. The reason for using two layers is inspired from the fact that in the brain, sensory modalities – e.g., the

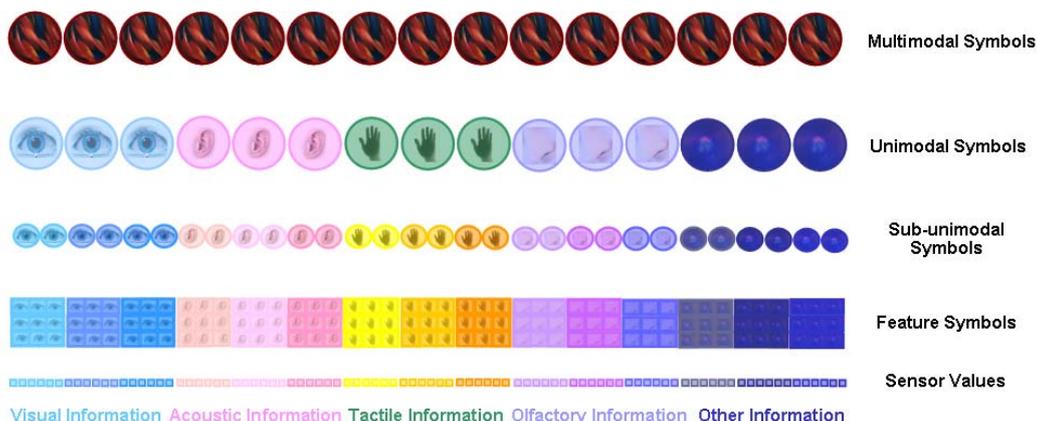


Figure 3: Architecture of neuro-symbolic networks for perceptive tasks

somatosensory system – can consist of further sub-modalities. As will be shown later, the sub-unimodal processing stage can be skipped if one modality only has to process information coming from one sensor type.

On the multimodal level, which corresponds to functions performed in the tertiary cortex of the brain, information from unimodal symbols of all modalities is merged to multimodal symbols.

### Neuro-symbolic Flow of Information

In the human brain, perceptual processes can be considered as the result of *bottom-up* processing and *top-down* processing working together. Bottom-up processing is based on incoming data from the receptors of our sense organs. Top-down processing is based on knowledge. Additionally, focus of attention influences perception top-down. Besides this, in the human brain, there exist feedbacks between neural levels.

Accordingly, in the neuro-symbolic network, there exists an information flow, which is directed bottom-up, there exist feedbacks within layers, and there exists an information flow, which is directed top-down.

Bottom-up information flow starts from sensor data and goes from lower levels to higher levels (see figure 4a).

By feedback connections, information from a higher processing stage can serve as input for a lower processing stage (see figure 4b).

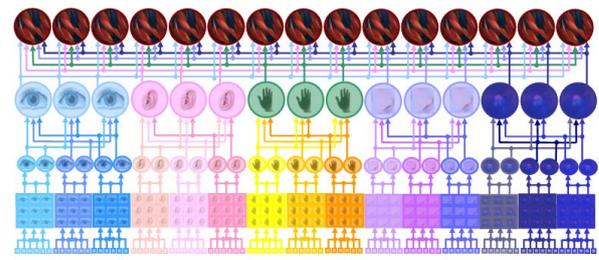
Besides these two sensor data-driven processes, neuro-symbol activations can also be influenced by top-down processes, which are referred to as cognitive information in figure 4c. For perceptual tasks, such top-down processes can result from knowledge about the environment and the objects and events being likely to occur or memory about objects and events that occurred in the past. Additionally, a mechanism called focus of attention can influence information processing top-down by concentrating processing power on currently relevant data and inhibiting the processing of data coming from other sources.

### C. Event-based Neuro-symbolic Information Processing

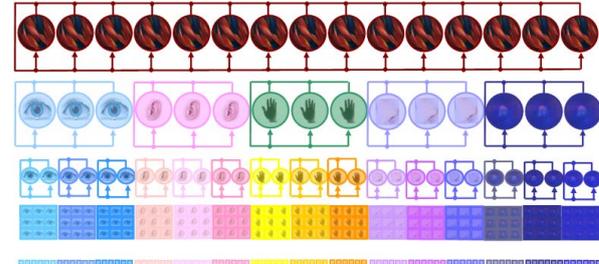
Information processing in neuro-symbolic networks is event-based, which means that information is not processed and transmitted to other neuro-symbols continuously but only if there occur changes in input data, which are sensor data for the case of perceptive tasks. Every time that at least one input value of a neuro-symbol changes, its activation grade and property values are recalculated and the results are sent to neuro-symbols connected to the output. By using event-based information exchange, computational power can be pared down, because a transmission and recalculation only needs to be performed if changes in the environment are perceived. This is especially important if the neuro-symbolic network is not implemented into a chip, which can perform real parallel processing, but is simulated on a computer where parallel processing is only “virtual”.

### D. Neuro-symbolic Learning

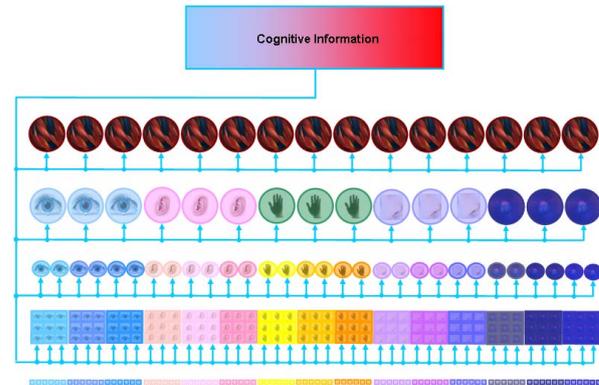
To allow flexibility, it is desirable not to have fixed, predefined correlations between neuro-symbols of different levels, but to allow learning of correlations. According to descriptions from neuro-science, learning of



(a) Bottom-up information flow



(b) Information feedback



(c) Top-down information flow

Figure 4: Information flow

adequate connections in different hierarchical layers takes place successively. Higher levels can only evolve if lower levels are already developed [8]. However it is agreed that “not everything can be learned from nothing”. Therefore, at the lowest levels, there already have to exist certain connections, which are innate and specified by genes. In accordance to this description, correlations between sensor data and the feature symbol level, which is the lowest hierarchical level, have to be predefined by the system engineer. All higher hierarchical levels can be learned. Learning is performed step by step. First, there are learned forward connections between feature symbols and sub-unimodal symbols. Next, there are determined needed feedback connections between sub-unimodal symbols. Afterwards, forward connections from the sub-unimodal level to the unimodal level can emerge followed by feedback connections between unimodal symbols and so on. The learning process shows certain analogies to supervised learning in neural networks as there are also used examples in form of input-target-data pairs to train neuro-symbolic networks. When using neuro-symbolic networks for perceptive tasks, the target data are the object, events, and situations that need to be perceived, and the input data are the sensor data that are triggered

when they occur in the environment. The presented examples need to cover all different objects, events, and situations that the system shall perceive. For each object, event, and situation, a number of examples are necessary, because there can occur deviations in how they are represented by sensory data.

Figure 5 illustrates the learning process for determining forward connections between the sub-unimodal and the unimodal level and feedbacks within the unimodal level of the visual perceptual system. For other perceptual modalities and other hierarchical levels, the same principle can be applied.

Similar to supervised learning in neural networks, there have to be presented input data and target data to the system. As already outlined, for the concrete application, input data are sensor data from sensors or arrays of sensors that are triggered when certain objects, events, or situations occurs. As the lower neuro-symbolic levels are already pre-connected when the current learning phase takes place, based on these sensor data, certain lower level symbols are activated, which then serve as actual input data for the learning algorithm. The target data are the particular higher-level symbols that are assigned to the current situations occurring. During the learning phase, the system memorizes the examples. Based on these examples, it determines and extracts correlations that exist between lower-level and higher level symbols and sets connections between them accordingly.

Besides forward connections from lower levels to higher levels, there can also exist feedback connections from higher levels downwards. These feedback connections can be determined and set by presenting the same examples as used for setting the forward connections to the system a second time. Feedback connections are necessary to inhibit the undesired concurrent activation of different neuro-symbols. For a more detailed explanation see [10] and [11].

In the learning phases, not only connections between neuro-symbols are set. There are also learned other correlations like values of properties and timing behavior of incoming data. Additionally, there can be identified and eliminated redundant neuro-symbols and it can be detected if additional neuro-symbols need to be added to the neuro-symbolic structure. The adding of neuro-symbols is performed in the same phases within which the forward connections between different levels are set. The elimination of neuro-symbols takes place in the same phases within which feedback connections are determined. Besides bottom-up connections and feedback connections, there also exist top-down connections from higher cortical processes. However, in the current implementation of the model, these top-down connections still have to be determined by the system engineer in form of explicit rules.

### E. Concept Clarification

In order to clarify the concept of neuro-symbolic information processing, it shall be applied to a concrete example, which is intentionally kept simple. The example is taken from the field of machine perception and aims to detect different activities of persons in a room. It shall be detected if a person stands in the room, walks around in

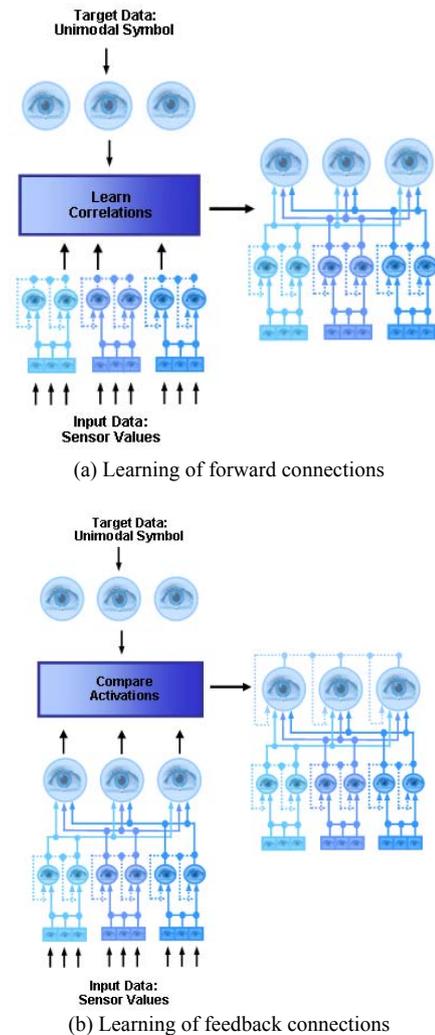


Figure 5: Neuro-symbolic learning

the room, enters the room, or leaves the room. For this purpose, the room is equipped with a number of diverse sensors: a motion detector, two tactile floor sensors, two light barriers, a video camera, and a microphone (see figure 6). The sensors have the property to have partly overlapping sensory fields of perception and to provide partly redundant information. The motion detector, tactile sensors, and light barriers provide a binary sensor signal (zero or one) and are referred to as tactile modalities. The motion detector detects movement in the room. The tactile floor sensors detect if an object is present in the room and whether it is present in the left or the right half of the room. The two light barriers detect whether an object passed the door and – by combining the information of both – in which direction it passed. The video camera evaluates whether a person is present in the room. The microphone detects the noise that occurs when someone walks.

To perceive what is going on in the environment, information coming from these sensory receptors is processed by a hierarchically structured arrangement of neuro-symbols. In figure 6, the symbol hierarchy and their interconnections being determined during the learning phases are illustrated.

In a first processing stage, feature symbols are extracted from the sensory raw data. In case of the binary sensors,

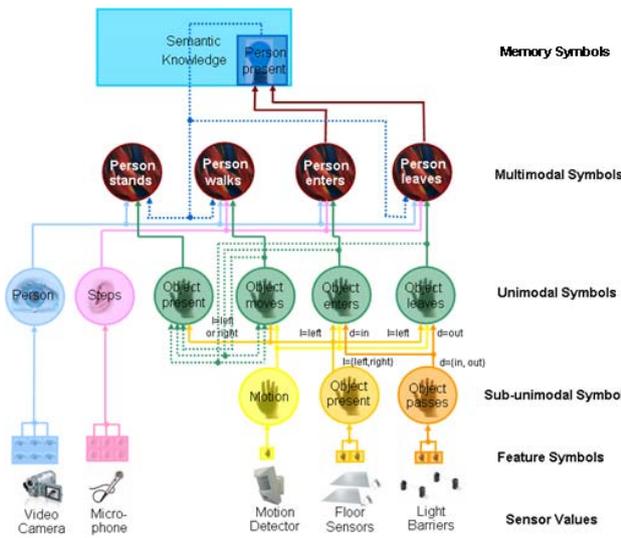


Figure 6: Neuro-symbol hierarchy for test configuration

the associations between the sensor values and the feature symbols are very simple due their binary sensor output values and the fact that the amount of used sensors is limited. Conceivable feature symbols in the visual information processing flow are lines, edges, curves, colors, moving forms, etc. Feature symbols for auditory processing are the frequency components or the loudness of a sound signal. Out of the feature symbols, sub-unimodal symbols in case of the binary sensors and unimodal symbols in case of the visual and auditory senses are formed. From the visual feature symbols it is detected whether a person is present in the room. From the auditory feature symbols it is extracted whether the characteristic noise of steps is perceived.

The sub-unimodal and unimodal symbols of the tactile perceptive system describe the states and activities of objects. They are not directly associated with states and activities of a person, because the sensors need not necessarily be triggered by persons. Some of the sub-unimodal symbols contain properties, which specify the neuro-symbols in more detail. The neuro-symbol “person present” has the property “location l”, which indicates at what position (left or right half of the room) an object is present. The symbol “object passes” has the property “direction d”, which comprises the information whether an object enters or leaves the room. The sub-unimodal tactile symbols are combined to unified unimodal tactile symbols.

The neuro-symbols “object enters” and “object leaves” are both connected to the same lower-level neuro-symbols. The decision which of the two neuro-symbols is activated depends on the property “direction d” of the symbol “object passes”. Some of the unimodal tactile neuro-symbols also depend on the value of the property “location l” of the symbol “person present”.

Using this bottom-up data processing, a critical point that has to be considered is the fact that whenever one of the neuro-symbols “object enters”, “object leaves”, or “object moves” is generated, there is also activated the symbol “object stands”, because it results from a combination of a subset of the same sub-unimodal tactile symbols. Additionally, each activation of the symbol “object enters”

or “object leaves” also triggers the symbol “object moves”. To overcome such an undesired concurrent activation of more than one symbol at a certain moment, inhibitory feedbacks are inducted between these neuro-symbols. In figure 6, they are depicted as dotted lines.

From the unimodal symbols of the visual, auditory, and tactile system, one of the four multimodal symbols is activated. In the example, the activation of multimodal symbols is also influenced top-down by knowledge. In the test set, the system “knows” that a person can only walk around in the room, stand in the room, or leave the room if he entered before. This means that the symbols “person walks”, “person stands”, and “person leaves” can only be activated if the symbol “person enters” has been activated before. Otherwise, their activation is inhibited. To perform this task, it has to be memorized that a person entered the room. This task is performed by so-called memory symbols – in the particular case the memory symbol “person present”. This memory symbol is activated when a person enters the room and deactivated when the person leaves the room again.

### 3. IMPLEMENTATION AND RESULTS

To test and evaluate the function and performance of neuro-symbolic networks, they were simulated in a virtual environment. For their implementation, the modeling language AnyLogic was used, which offers the possibility of modular hierarchical design and simulated parallel processing. AnyLogic provides certain design elements (active objects, ports, connectors, messages, etc.), which allow a fast and efficient implementation of neuro-symbols as well as their interconnection to neuro-symbolic networks [12].

The neuro-symbolic model was tested for perceptive tasks where different activities of persons in a building had to be perceived. For the provision of these sensor data, a simulator was used, which allowed it to generate sensor values based on activities going on in a virtual office environment [4], [5]. After completion of the learning phase, the neuro-symbolic network, which was trained based on a range of examples provided by the simulator, proved to be suitable for detecting learned activities of persons. It also showed the ability to generalize over unseen examples.

To get an impression how neuro-symbolic information processing is performed within such a network a concrete example shall be given. The sensor configuration and neuro-symbol hierarchy for this example is depicted in figure 6. Figure 7 shows what neuro-symbols from the sub-unimodal level upwards are activated when a person enters the room. The first five signals of the diagram show the activation over time of the tactile sub-unimodal and the visual and auditory unimodal neuro-symbols. The next six signals show what higher-level neuro-symbol activations result from these activations. The last row shows when the memory symbol “person present” is set, which happens when the system detects that a person enters the room.

The numbers above the time signals indicate the activation value that is sent from the corresponding neuro-symbol to other connected neuro-symbols. For symbol activations

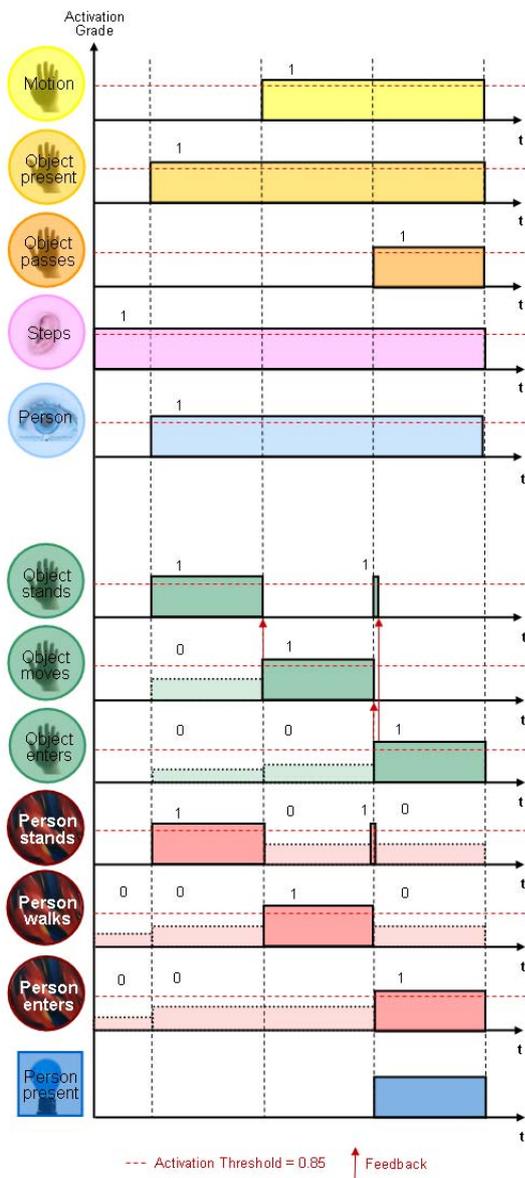


Figure 7: Activated neuro-symbols when a person enters a room

having a value below the threshold, the value zero is transmitted to other symbols via the output. In figure 7, these cases are depicted with dotted lines and in lighter shades.

Points in time where a feedback message is sent to another symbol are marked with an arrow. As long as a feedback is effective, the activation grade of the symbol that receives this information is set to zero. Figure 7 shows that for the symbol “object stands”, there can occur a short activation peak when the symbol “object enters” is activated. This is due to the fact that the symbol “object moves” has an inhibitory influence on the symbol “object stands”. On the other hand, the symbol “object enters” has an inhibitory influence on the symbol “object moves” and on the symbol “object stands”. If now the inhibitory feedback message of the symbol “object enters” reaches the symbol “object moves” before the symbol “object stands”, the symbol “object moves” is deactivated and its inhibitory influence on the symbol “object stands” is rescinded before the feedback of the symbol “object enters” can get effective. This results in an activation of the symbol “object stands” for a very short time. In the simulation, where the time for transmitting and processing can be set to zero, the length of such peaks can be assumed to be zero. However, when implementing the

model into hardware, signal run times and time for data processing have to be considered, which will result in such peaks of a certain length. It has to be assured that these peaks do not influence succeeding information processing stages by declaring signals only as valid if they are active for a certain minimum time.

#### 4. CONCLUSION

In this paper, a new information processing principle called neuro-symbolic information processing has been introduced, which aims to combine characteristics of neural and symbolic information processing to get the advantages from both methods and eliminate their weak points. The basic information processing units for this principle are neuro-symbols, which are structured to neuro-symbolic networks. Within these networks, there exist different types of connections. It has been outlined how to emulate the structural organization of the perceptual system of the brain by neuro-symbolic networks in order to use them for perceptive tasks. The gained results are very promising. The application of the concept to other application fields will be subject to further research.

#### REFERENCES

- [1] E. B. Goldstein, “*Wahrnehmungspsychologie*”, Spektrum Akademischer Verlag, 2002.
- [2] E. B. Goldstein, “*Sensation and Perception*”, Thomson Wadsworth, 2007.
- [3] S. Harnad, “The Symbol Grounding Problem”, *Physica D*, Vol. 42, 1990, pp. 335–346.
- [4] H. Hareter and G. Pratl and D. Bruckner, “A Simulation and Visualization System for Sensor and Actuator Data Generation”, *Proc. 6th IFAC International Conference on Fieldbus Systems and their Applications*, 2005, pp. 56–63.
- [5] H. Hareter, “*Worst Case Szenarien Simulator für die Gebäudeautomation*”, PhD Thesis, Vienna University of Technology, to be published 2008.
- [6] V. Honavar and L. Uhr, “*Symbolic Artificial Intelligence, Connectionist Networks, and Beyond*”, Iowa State University of Science and Technology, Department of Computer Science, 1994.
- [7] C. R. Huyck, “*Combining Symbolic and Connectionist Models in Artificial Intelligence*”, 1999.
- [8] A. R. Luria, “*The Working Brain - An Introduction in Neuropsychology*”, Basic Books, 1973.
- [9] M. Solms and O. Turnbull, “*The Brain and the Inner World – An Introduction to the Neuroscience of Subjective Experience*”, Other Press New York, 2002.
- [10] R. Velik, “A Model for Multimodal Humanlike Perception based on Modular Hierarchical Symbolic Information Processing, Knowledge Integration, and Learning”, *Proc. 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems*, 2007, pp. 8.
- [11] R. Velik and G. Pratl and R. Lang, “Multi-Sensory, Symbolic, Knowledge-Base Model for Humanlike Perception”, *Proc. 7th IFAC International Conference on Fieldbuses and Networks in Industrial and Embedded Systems*, 2007, pp. 273–278.
- [12] R. Velik, “*A Bionic Model for Human-like Perception*”, PhD Thesis, Vienna University of Technology, 2008.
- [13] S. Wermter, “Hybrid Neural Symbolic Integration”, *Proc. International Conference on Neural Information Processing Systems*, Breckenridge, Colorado, 1998.
- [14] S. Wermter and R. Sun, “*Hybrid Neural Systems*”, Springer, 2000.