Biologically motivated neural network UBAL in cognitive robotics

Kristína Malinovská and Jakub Mišovský

Centre for Cognitie Science, DAI FMPI,
Comenius University in Bratislava, Slovakia
Mlynská dolina, 84248 Bratislava
Email: kristina.malinovska@fmph.uniba.sk

The novel neural network model UBAL represents a more biologically plausible alternative to error-backpropagation. Since UBAL is a heteroassociator with a particular mechanism of activation propagation including self loops, it demonstrates interesting emergent phenomena such as the ability to generate meaningful reconstructions of the learned patterns. Cognitive robotics uses the paradigm of understanding by building and in this spirit explores and models the neural correlates of cognitive capacities in humanoid robots. In our modeling we tackled the understanding of motor actions and associations between different modalities such as vision, touch, and proprioception. In this paper, we outline the history of our modeling efforts as well as new prospects for our cognitive robotic models.

1 Introduction

Cognitive robotics aims at studying cognition of humans via "understanding by building" (Pfeifer & Scheier, 1999) employing connectionist bio-inspired models in robots that interact with and learn from complex and multimodal environments. It is well known the cognitive functions in the brain are built up in a hierarchical manner, the brain areas involved are processing information on lower levels such as low-level feature detection or motor control and pass it on to higher, so called association areas. The common coding theory (Hommel et al., 2001) suggests that there is a common representational base for perception and action (motor performance). The perception of action automatically activates its motor component and vice versa. The common coding framework might also be considered the means for sensorimotor simulation (Barsalou, 1999). A phenomenon closely related to common coding is the so called mirror neuron theory (Rizzolatti & Sinigaglia, 2010) which claims our brain has special mechanisms that allow us to resonate with or step into the shoes of another person thus providing us with deeper understanding of the observed actions.

Modeling of cognitive capacities in humanoid robots within the embodied paradigm of cognition builds up on robots own senses, competences and experience. In our research we have focused on building and connecting these capacities in a modular way, utilizing different neural network architectures and learning paradigms. This is something in contrast with current trend of end-to-end deep architectures, but unlike standard deep models, it offers bio-inspired mechanisms and relates more to the processes in the brain.

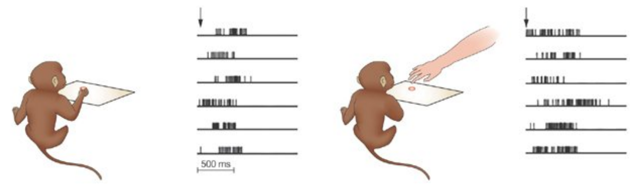

Fig. 1: Illustration of macaque monkey's mirror neuron responses - motor neurons are firing due to visual stimulation (Rizzolatti et al., 2009).

2 Our models

2.1 Robot MNS

First of our models that are building high-level multimodal representations was our robotic mirror neuron system (MNS) model (Rebrová et al., 2013). It consists of several modules: the core and the topmost part is the mirror neuron circuit itself, which is connected to association areas and the low-level modules for execution of movement and gathering of visual information.

In our modeling we assume the sensory-motor links are established between higher level representations, rather than directly between low-level representations of the movement. To encompass the temporal nature of the action sequences (i.e. joint angles of the robot changing in time), the association areas are made of the special type of recurrent self-organizing maps, namely the merge-SOMs (MSOM, Strickert & Hammer, 2005). For our experiments we used the simulated version of the humanoid iCub robot (Tikhanoff et al., 2008) trained to perform three different grasping actions. These are processed by the higher level association area (MSOM) and get self-organized on the resulting maps. We observed that this organization is according to the type of the grasp. Similarly, the higher level association area for vision received information from low-level vision module in terms of joint positions from different viewpoints.

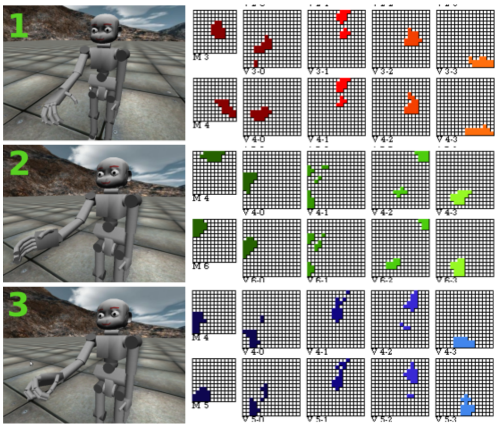

Fig. 2: The left side of the figure shows the examples of iCub's grasping actions of three types and on the right side activations of the trained motor MSOM of 12×12 neurons and from the visual MSOM of 16×16 neurons for each view-point, both binarized with $k = 16$.

This MSOM gets organized primarily according to the viewpoints and secondarily according to the grasp types (Fig. 2).

During the production of the movement, the motor information and the visual information from the self-observation perspective are binarized via the k-WTA method[1] and associated bidirectionally using the BAL (Bidirectional Activation-based Learning) neural network (Farkaš & Rebrová, 2013). We assume that the robot observes another robot producing the same actions and creates visual representations of those actions from viewpoints and associates them with the motor representations as well. After the model is trained, the action observed by the robot elicits the motor representation of the action. These are projected to the topmost module together with visual information, where the motor information helps to form the view-independent representations (Fig. 3).

In this previous work we proposed the BAL model, but soon found out that it was unable to learn one-to-many mappings required to associate different visual representations from different viewpoints with one motor representation. With a motivation to overcome this problem and also to provide a universal account on learning in the brain, the UBAL (Universal Bidirectional Activation-based Learning) model was born.

2.2 UBAL

The fruit of our research endeavors of cognitive modeling is a novel neural network model UBAL (Malinovská et al., 2019). It represents an alternative to classical, effective, yet biologically implausible error-

[1] k-WTA or k winner takes all: k winners on the neuron map are set to 1 and others to 0

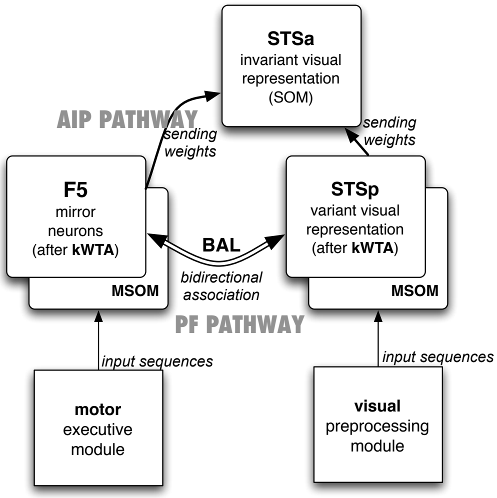

Fig. 3: The robot MNS model (Rebrová et al., 2013).

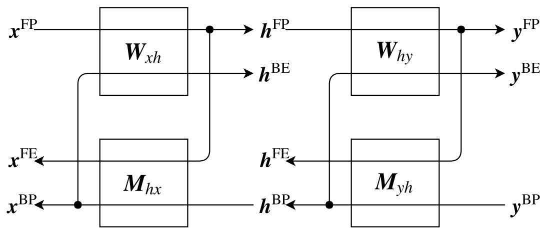

Fig. 4: Schematic deptions of UBAL. The neural activation in the model propagates bidirectionally (F and B) between layers x and y via weights W and M. Activation states computed by the network are also propagated backwards via *echo* connection denotes with E.

backpropagation learning. As well as its predecessor BAL, inpired by GeneRec model (O'Reilly et al., 2012), UBAL is heteroassociative network that does not learn by gradient descent. Unlike any other model UBAL maintains separate weight matrices for two different activation propagation directions between inputs and outputs (Fig. 4). As in the brain the signal only travels through the axons of the neurons in one way via separate synaptic weights. Additionally, UBAL enhances contrastive Hebbian learning rule with an internal echo mechanism enabling self-supervised learning. Therefore, it is able to master various tasks such as association (memory), denoising and classification based on its hyperparameter setup.

Since UBAL approaches any problem as a bidirectional heteroassociation, it has intriguing emergent properties, such as generation of patterns while trained for classification (Malinovská & Farkaš, 2021) as shown in Fig. 5. Such a neural network model can also be understood as a model cognition and the representations it creates on the input level without the stimulation input can be seen as prototypes of the learned categories.

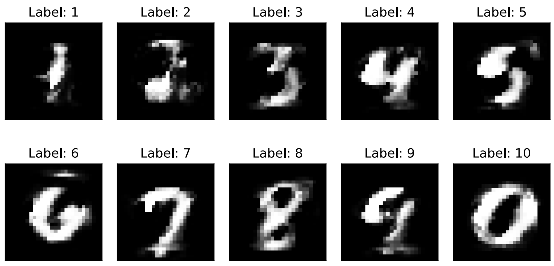

Fig. 5: Projections of learned digits from the MNIST dataset generated by UBAL.

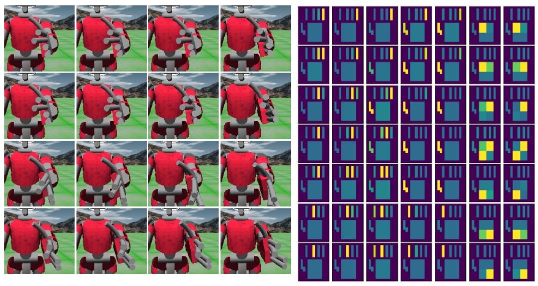

Fig. 6: Proprio (left) and tactile (right) self-organizing map-based association area representations.

2.3 Proprio-tactile associations

In a more recent work within the modular hierarchical modeling paradigm (Malinovská et al., 2022) we explored the phenomenon of the emergence of the body schema which built form our early infancy. Body representations have a multimodal nature that foremost comprises somatosensory, i.e. tactile and proprioceptive information, coupled in a hierarchical fashion in high-level multi-modal representations. In line with the cognitive robotics paradigm we proposed a simple connectionist model build from hierarchically connected neural networks that learns the proprioceptive-tactile representations for the humanoid iCub robot.

This model, along with the previous MNS model utilizes self-organizing maps (SOM, Kohonen, 1997) as association areas (Fig. 6, the k-WTA mechanism and the UBAL model at the topmost level. It has shown a quite good ability to predict touch and its location from proprioceptive information. Due to the generative properties of UBAL mentioned above we could also observe how the model predicts the body-configuration based on the information from the tactile modality, even though the association is many-to-one rather then one-to-one. This intriguing properties along with other aspects such as the quality of the high-level representations emerging on the hidden layer of the model are worth further investigation and modeling and are subject to ongoing research.

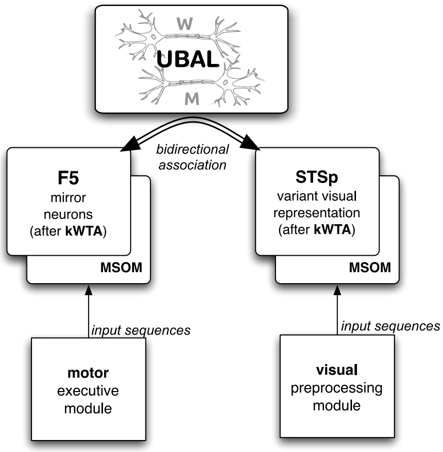

Fig. 7: The robot MNS model (Rebrová et al., 2013).

2.4 Robot MNS: a revival

Our path towards robot mirror neurons or robot common coding model is still unraveling. Our current effort is to revisit the old MNS model and replace the BAL module with the UBAL module which would in turn also replace the topmost association area of the model. Progressing with the current knowledge on the UBAL model we can also assume interesting phenomena if we turn the prediction task the other way round and let our model predict what would be seen when the action is executed.

The modular neural architecture (Fig. 7 is composed of low level execution modules that can be various different control models[2] and vision processing module which could be a deep neural network for extraction of the observed joint positions. The core of our model is representing the movements in motor (proprioceptive) and visual modalities as patterns in a topologically organized association area module utilizing the MSOM model (Strickert & Hammer, 2005). On the top these representations are connected via the UBAL model, capable of connecting multiple different visual representations with one motor representation (robot's own motor repertoire).

A vital task for this research is to collect an adequate sample of data from the robot platform and also to extend from the iCub robot to other platforms. In our case we are working with the small humanoid NICO robot (Kerzel et al., 2017), which we have also in physical form. The model that allows the robot to connect its own movement with the observed one could be beneficial for human-robot interaction and smooth implementation of robot's ability to imitate the human partner. The core idea behind is to use the hierarchical neural model that will allow transfer on the conceptual level, rather then on the concrete effector level, which is also

[2]Modules generating joint angle sequences for the robot such as an inverse kinematics module or RL-trained neural network such as the one used in the previous model

a known property of the mirror neurons (Rizzolatti & Sinigaglia, 2010). The aim for the next modeling would be to have the way to retrieve the robot's movement patterns from the association areas. Such model, connected to a helpful software processing the human moving in real time such as the MediaPipe (Lugaresi et al., 2019) can be then used for the interaction of the robot with the human in imitation learning and collaboration tasks.

3 Conclusion

Cognitive robotics aims at studying cognitive capacities of humans via building them in humanoid robots. Capabilities such as action understanding or body schema building are a result of hierarchical multi-modal processing and therefore we claim it is opportune to build them via modular hierarchical neural network architectures with use of biologically plausible neural learning mechanisms such as self-organizing maps and bio-inspired association models (UBAL). In our future work we aim at further developing and testing the model for action understanding and use it in the context of human-robot interaction in imitation and collaboration tasks.

Acknowledgement

This paper was written at the Centre for Cognitive Science at DAI FMPI Comenius Universty Bratislava, with the support of a grant VEGA 1/0373/23 and KEGA 022UK-4/2023. We also thank for support to the Slovak Society for Cognitive SSKV[3].

References

Barsalou, L. (1999). Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22(04):577–660.

Farkaš, I. & Rebrová, K. (2013). Bidirectional activation-based neural network learning algorithm. V *Proceedings of the International Conference on Artificial Neural Networks (ICANN), Sofia, Bulgaria*.

Hinton, G. E. & McClelland, J. L. (1988). Learning representations by recirculation. V *Neural Information Processing Systems*, str. 358–366. New York: American Institute of Physics.

Hommel, B., Müsseler, J., Aschersleben, G. & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(05):849–878.

Kerzel, M., Strahl, E., Magg, S., Navarro-Guerrero, N., Heinrich, S. & Wermter, S. (2017). Nico - neuro-inspired companion: A developmental humanoid robot platform for multimodal interaction. V *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, str. 113–120.

Kohonen, T. (1997). *Self-organizing Maps*. Springer.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M. & Grundmann, M. (2019). Mediapipe: A framework for building perception pipelines.

Malinovská, K. & Farkaš, I. (2021). Generative properties of Universal Bidirectional Activation-based Learning. V *Artificial Neural Networks and Machine Learning - ICANN 2021*, str. 80–83.

Malinovská, K., Farkaš, I., Harvanová, J. & Hoffmann, M. (2022). A connectionist model of associating proprioceptive and tactile modalities in a humanoid robot. V *IEEE International Conference on Development and Learning (ICDL)*.

Malinovská, K., Malinovský, L., Krsek, P., Kraus, S. & Farkaš, I. (2019). UBAL: a Universal Bidirectional Activation-based Learning Rule for Neural Networks. V *Proceedings of the 2019 2nd International Conference on Computational Intelligence and Intelligent Systems*, str. 57–62.

O'Reilly, R. C., Munakata, Y., Frank, M., Hazy, T. et al. (2012). *Computational Cognitive Neuroscience*. PediaPress.

Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence*. The MIT Press, Cambridge, MA.

Rebrová, K., Pecháč, M. & Farkaš, I. (2013). Towards a robotic model of the mirror neuron system. V *Proceedings of the 3rd joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*. Osaka, Japan.

Rizzolatti, G., Fabbri-Destro, M. & Cattaneo, L. (2009). Mirror neurons and their clinical relevance. *Nature Clinical Practice Neurology*, 5(1):24–34.

Rizzolatti, G. & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11(4):264–74.

Strickert, M. & Hammer, B. (2005). Merge SOM for temporal data. *Neurocomputing*, 64:39–71.

Tikhanoff, V., Cangelosi, A., Fitzpatrick, P., Metta, G., Natale, L. & Nori, F. (2008). An open-source simulator for cognitive robotics research: The prototype of the icub humanoid robot simulator. V *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*, str. 57–61. ACM.

[3] https://cogsci.fmph.uniba.sk/sskv/