ECAI 2023 KRAKÓW

UNIVERSITY of INFORMATION TECHNOLOGY and MANAGEMENT in Rzeszow, POLAND

POLITECHNIKA RZESZOWSKA im. IGNACEGO ŁUKASIEWICZA

AGH

Motivated Agent with Semantic Memory

Janusz Starzyk[1] , Marcin Kowalik[2], Adrian Horzyk[3]

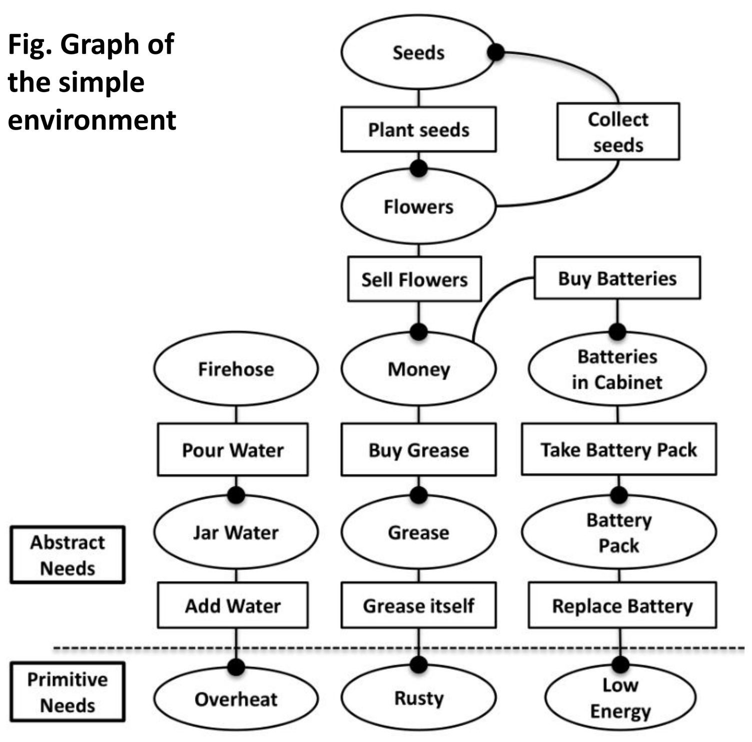
1- University of Information Technology and Management in Rzeszów, Poland / 2 - Rzeszów University of Technology, Poland / 3 - AGH University of Krakow, Poland

GOAL> The main objective is to test&demonstrate the effectiveness and advantages of a **motivated agent (ML)** scheme, particularly when the agent can utilize semantic memory to store knowledge about relationships in dynamically changing and unknown environment.

Simple environment

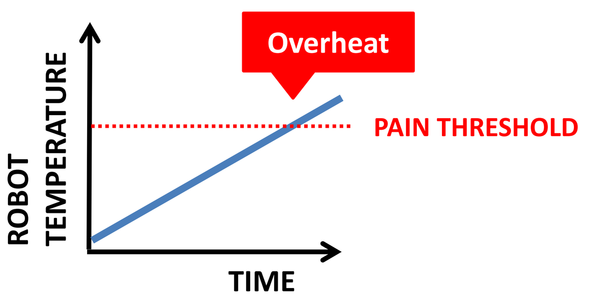
The agent's environment, as described in Table and shown in the figure below, is relatively simple. However, as it has been demonstrated in our's previous works, reinforcement learning systems that lack internally set objectives face difficulties even in such a simple environment. This is because they do not perform well in a non-stationary environment.

Fig. Graph of the simple environment



Tab. Meaningful action-resource pairs and their effect on the environment

Action (Motor)	Using Resource	Result on environment	Resource depleted
Add water	Jar Water	Lowers temperature of the robot	Water in jar
Grease itself	Grease	Reduces the rust level of the robot	Grease at hand
Replace battery	Battery pack	Increases energy level of the robot	Battery pack at hand
Pour water	Firehose	Increases water supply in the jar	Unlimited resource
Buy grease	Money	Increases grease supply	Money at hand
Take battery pack	Cabinet with spare batteries	Increases number of spare batteries at hand	Battery pack at cabinet
Buy batteries	Money	Increases number of spare battery packs in cabinet	Money at hand
Plant seeds	Seeds	Increases number of flowers	Seeds
Sell flowers	Flowers	Increases money at hand	Flowers
Collect seeds	Flowers	Increases number of seeds	Flowers



- **The agent's job is to take care of primitive needs such as overheating, rust, and energy levels.**
- **Primitive needs can be relieved by the use of resources present in an environment.**
- **A resource can be restored by using the appropriate resource and motor action.**
- **Primitive needs increase with time.**
- **A resource is available with probability P:**

  P = 1 / (1+ rate * uses)

MOTIVATED AGENT

Motivated learning is an approach that focuses on utilization of an agent's intrinsic goals. Some of these goals are predetermined by the designer and are rewarded. As the agent interacts with an environment, it acquires knowledge about the environment and sets internal goals, which are not rewarded, to achieve the designer's goals. To switch between different goals, the agent uses the opposite of the reward signal, which is the pain signal of unmet needs. The greater the need (represented by pain), the faster the agent will act to address it (**ATTENTION**). The motivational system serves to create and select the goals of the agent, while its memory can aid in implementing these goals by utilizing the knowledge of the environment stored within it.

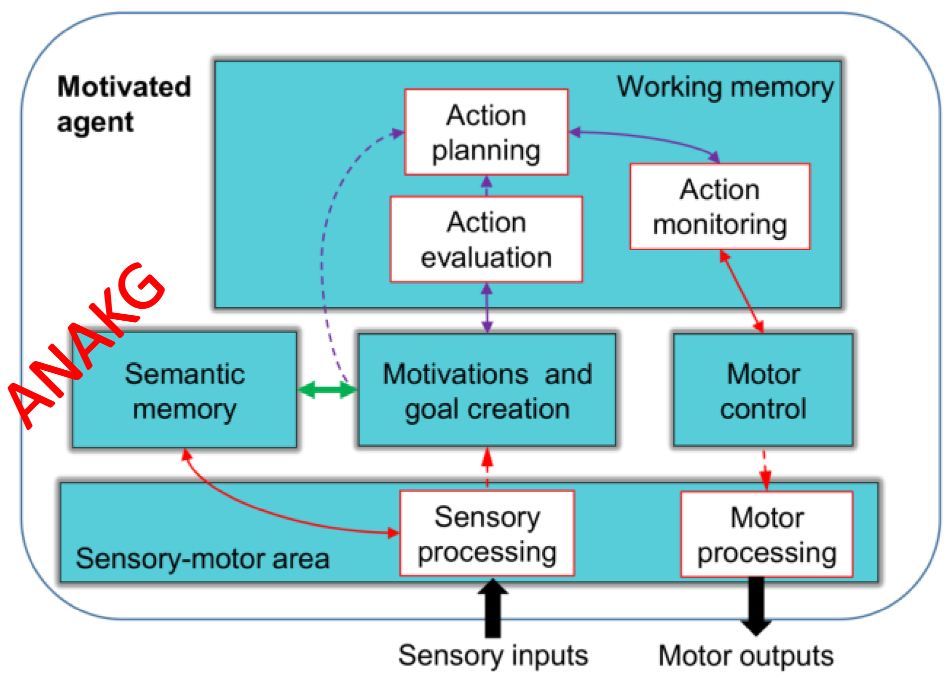


Fig. Simplified diagram of the motivated agent

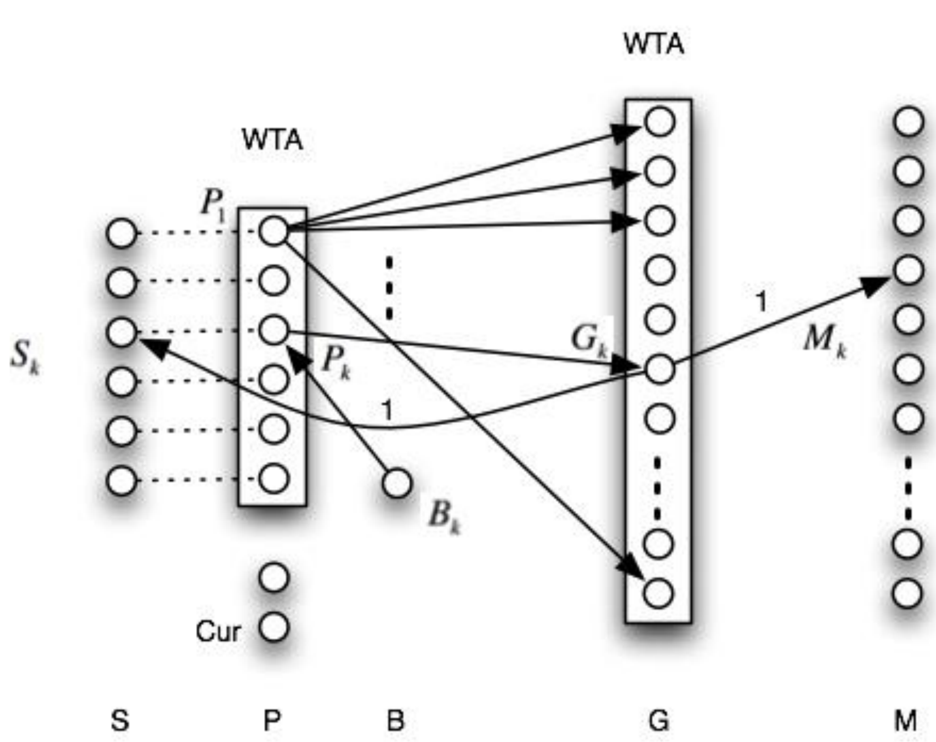


Fig. Connections between sensory S, motor M, bias B, pain P and goal neurons G. WTA – Winner Takes All

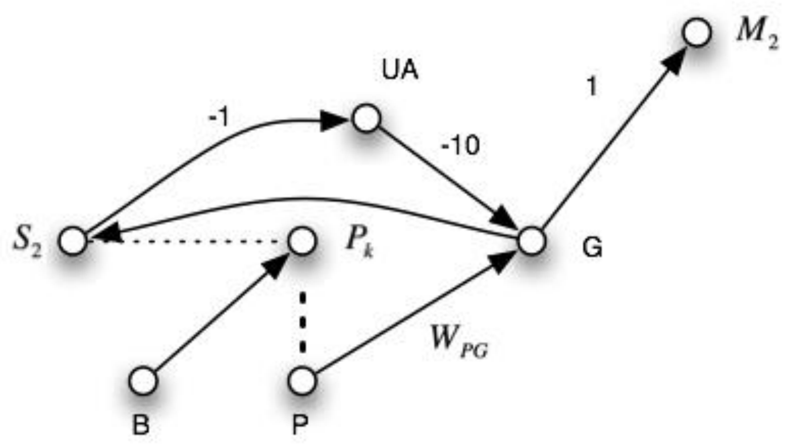
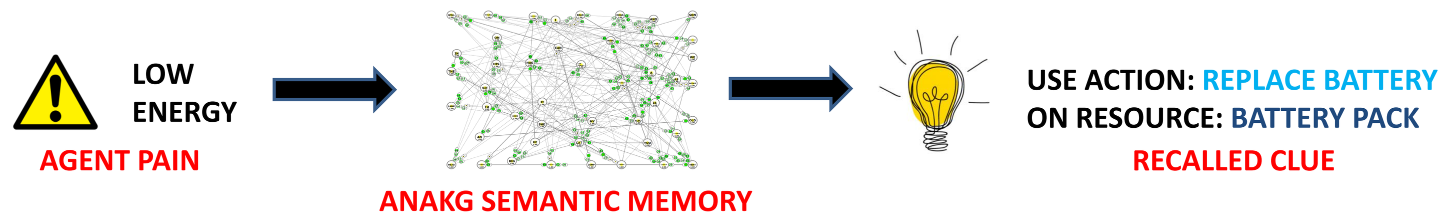


Fig. Trainable connections between pain, bias, and goal neurons. UA – inhibiting neuron, it fires if resource S_2 is not available.

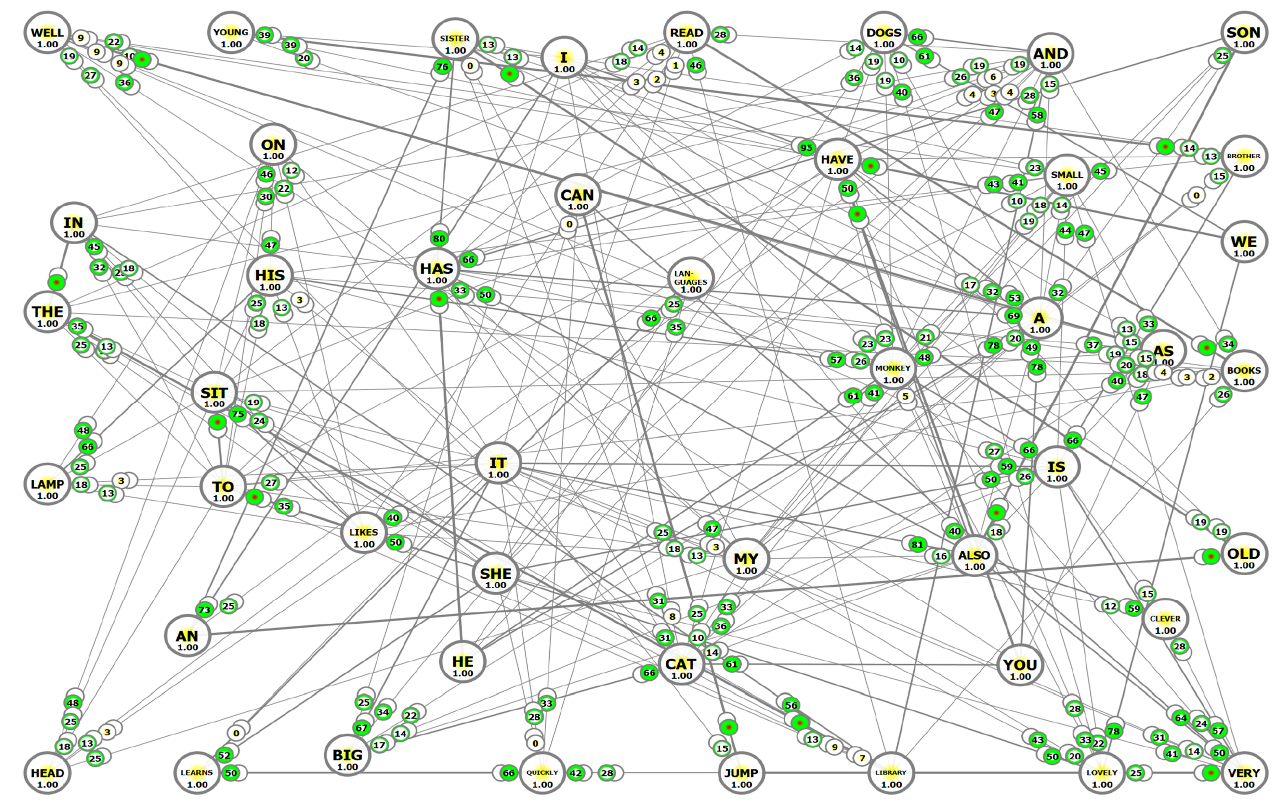
- **Goal G represents pair of a resource and an action.**
- **Sensor S_k sensens presence of a single resource .**
- **A clue from semantic memory temporarily modifies the values of weights W_{PG} which can result in the selection of different goal G.**

ANAKG - ACTIVE NEURO-ASSOCIATIVE KNOWLEDGE GRAPHS

ANAKG creates a complex graph structure of dynamic and reactive neurons and connections between them to represent a set of training sequences. The ANAKG network can be used as an associative semantic memory, which is able to recall revelant information to help solve agent's needs (pains).



In the ANAKG network, a neuron represents and aggregates all instances of the same token (i.e., word or symbol) that occur in all training sequences. Internal states of neurons are time dependend.



ANAKG LEARNING & RECALLING MEMORY

SPARSE STRUCTURE

CONTEXTUAL CONNECTIONS

HARD & SOFT ATTENTION

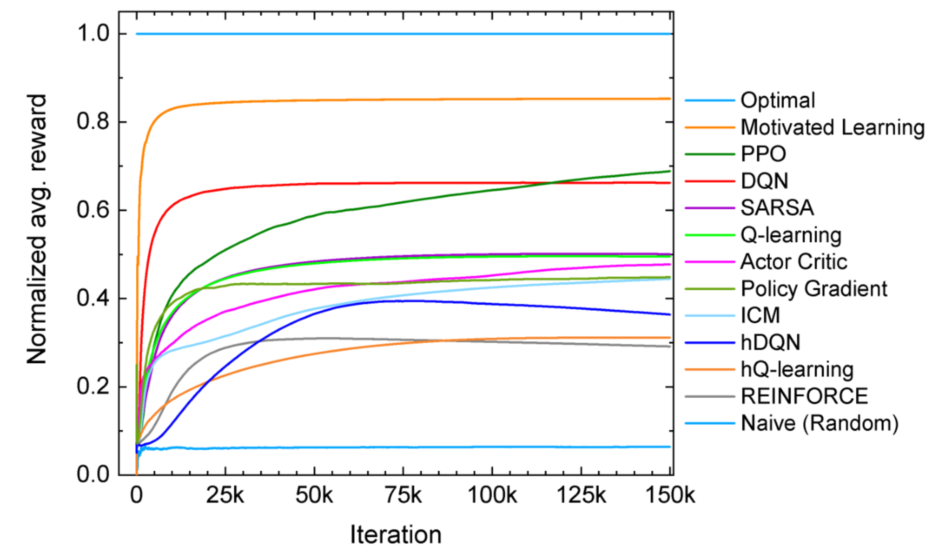
Fig. The ANAKG network trained on a short text is shown. Neurons are represented by big black circles and weights by smaller green ones. The network correctly recalls all training sequences for the minimum unique initial contexts. For example, when given the input „I", the network recalls „I have a monkey", and for the input „My", it recalls „My monkey is very small".

Results: SIMPLE ENV - ML agent vs RL

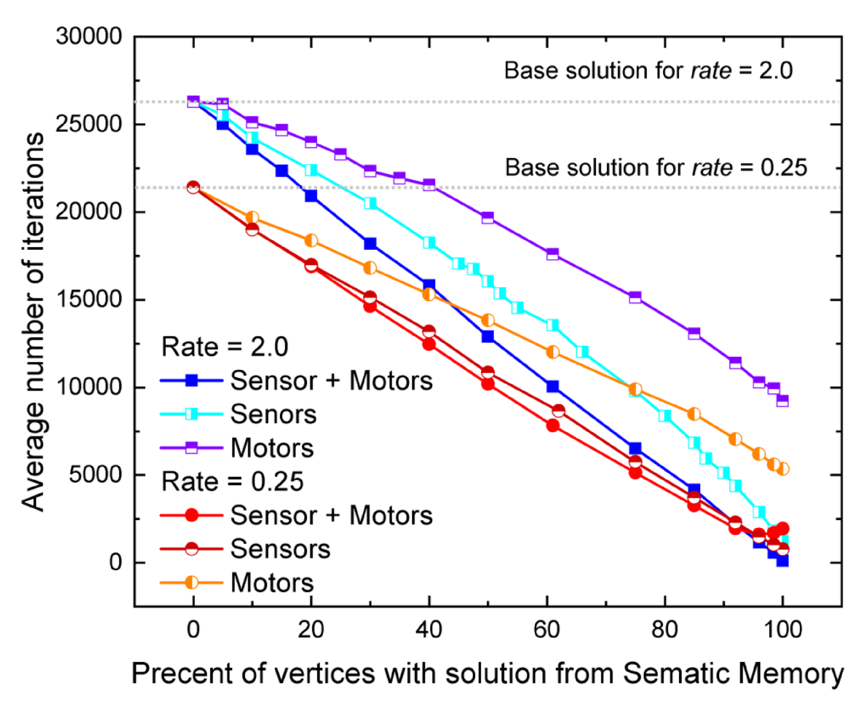
The Motivated Learning (ML) agent (w/o semantic memory) was compared in a simple environment to several RL-algorithms, including classical methods such as REINFORCE, Q-learning (TD0), and SARSA (TD0). Additionally, more advanced algorithms were evaluated, including Policy Gradient (PG), Actor-Critic (AC), Deep Q-Network (DQN), deep reinforcement learning with hierarchical value functions (h-DQN), and the tabular version of h-DQN (hQ-learning). Finally, the study also included state-of-the-art algorithms such as Proximal Policy Optimization (PPO) and Curiosity-driven exploration by self-supervised prediction (ICM). This selection of algorithms provides a diverse range of approaches for learning and solving tasks within a given environment.

Some of the evaluated algorithms, such as h-DQN and hQ-learning, utilize intrinsic goals to generate corresponding intrinsic rewards. On the other hand, algorithms like ICM leverage curiosity to create intrinsic goals and rewards. **The utilization of intrinsic goals and intrinsic rewards in tandem represents a key feature of the ML agent.**

The goal of an agent acting in the graph environment is to keep all primitive needs below a pain threshold. Every successful relief of primitive pain, which exceeds the pain threshold, is rewarded. he agent must learn which pairs of a resource and an action allow for the relief of primitive pain or the restoration of a resource (relief of abstract pain associated with a resource).



Results: COMPLEX ENV - ML agent supported by Semantic Memory



Semantic memory has a more pronounced impact on the number of iterations required to solve complex environments with numerous resources and dependencies expressed in the environmental graph.

The complex environment under consideration consists of 117 resources and 10 possible motor actions, represented by an environmental graph with a depth of 15. The generated graph includes 7 unlimited resources, and approximately 12% of all resources can be restored using multiple actions.

We conducted experiments with two different resource availability settings for the agent (rates set to 2 and 0.25). In these experiments, we varied the percentage of vertices in the environmental graph that utilized clues provided by the semantic memory. Three types of clues were considered:

- full (sensory-motor pairs)
- partial sensors
- partial motors

Question: Why does having full or nearly full knowledge about the environment, which is set by the parameter rate at 0.25 (as seen in the red curve in the figure on the right), increase the number of iterations required to solve the environment?

CONCLUSIONS

- **A motivated learning agent benefits from information about the environment stored in its associative memory.**
- **Any useful knowledge about the environment can facilitate the agent's learning task and enhance its ability to explore the environment.**
- **Motivated learning is more efficient than reinforcement learning in dynamically changing environments. This applies even if an RL algorithm generate intrinsic goals and intrinsic rewards or is acting based on curiosity.**