1-2010

Motivated learning as an extension of reinforcement learning

Janusz STARZYK

Pawel RAIF

Ah-hwee TAN
Singapore Management University, ahtan@smu.edu.sg

Citation

STARZYK, Janusz; RAIF, Pawel; and TAN, Ah-hwee. Motivated learning as an extension of reinforcement learning. (2010). *4th International Conference on Cognitive Systems, CogSys 2010: January 27-28, Zurich: Proceedings*. 1-1.

Available at: https://ink.library.smu.edu.sg/sis_research/6573

Motivated Learning as an Extension of Reinforcement Learning

Janusz A. Starzyk, Paweł Raif, Ah-Hwee Tan

Ohio University, Athens, OH, USA *Silesian University of Technology, Gliwice, POLAND* *Nanyang Technological University, SINGAPORE*

Motivated Learning

Motivated Learning (ML)

◆ML paradigm uses neural structures that self-organize and form 'pain centers' which correspond to internal motivations.

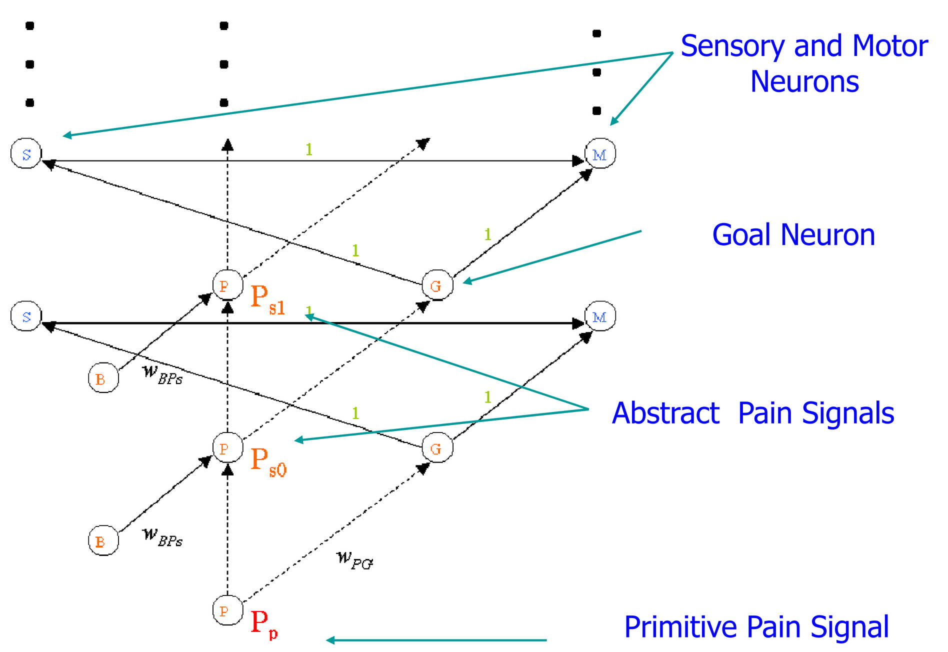
◆A ML system uses artificial curiosity to explore, and creation of **abstract motivations** to learn efficiently and purposefully. It increases the internal complexity of representations and skills.

◆At every step, the agent finds an action (actions) that satisfies its abstract pains and this may result in new motivations. Gradually, the agent learns values of various states and actions for various motivations.

◆ML effectively implements and manages a hierarchy of goals without explicit reward for different stages of hierarchy.

◆Any form of reinforcement learning (e.g. hierarchical reinforcement learning with subgoal discovery) can be used to resolve abstract pains.

◆ML enables active **learning through interaction** with the environment.



RL vs ML

Reinforcement Learning	Motivated Learning
Single value function	Multiple value functions
The same for all goals	One for each goal
Measurable rewards	Internal rewards
Can be optimized	Cannot be optimized
Predictable	Unpredictable
Objectives set by designer	Sets its own objectives
Maximizes the reward	Solves minimax problem
Potentially unstable	Always stable
Learning effort increases with complexity	Learns better in complex environment
Always active	Acts when needed

http://derekallard.com/img/post_resources/draft_robot_revision_3.jpg

Computational Model and Simulation

Motivated Learning or Reinforcement Learning

◆Simulation Framework

We have developed a unified framework to conduct computational experiments with both learning systems:
•**Motivated learning** based on Goal Creation System, and
•**Reinforcement learning** using RL Q-Learning Algorithm.
Our goal was to compare their performance in terms of **learning speed** and **task completion ability**.

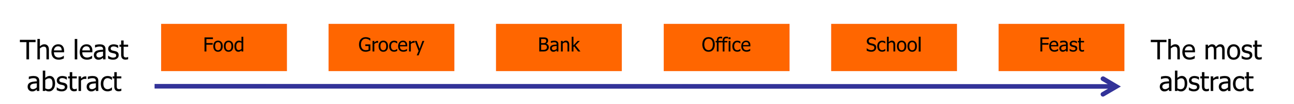
Meaningful sensory-motor pairs and their effect on the environment.

Id	SENSORY	MOTOR	INCREASES	DECREASES	PAIR Id
0	**Food**	**Eat**	Sugar level	Food supplies	0
1	**Grocery**	**Buy**	Food supplies	Money at hand	7
2	**Bank**	**Withdraw**	Money at hand	Spending limits	14
3	**Office**	**Work**	Spending limits	Job opportunities	21
4	**School**	**Study**	Job opportunities	Social contacts	28
5	**Feast**	**Play**	Social contacts	-	36

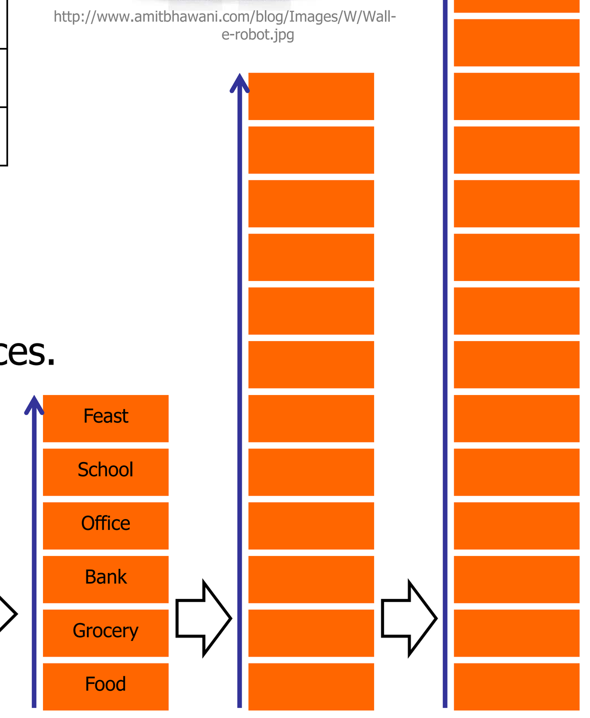
http://www.amitbhawani.com/blog/Images/W/Wall-e-robot.jpg

◆Task Specification

•**Complex, dynamically changing environment**
In base experiment environment consist of six different categories of resources. Five of them have limited availability. One, the most abstract resource is inexhaustible.



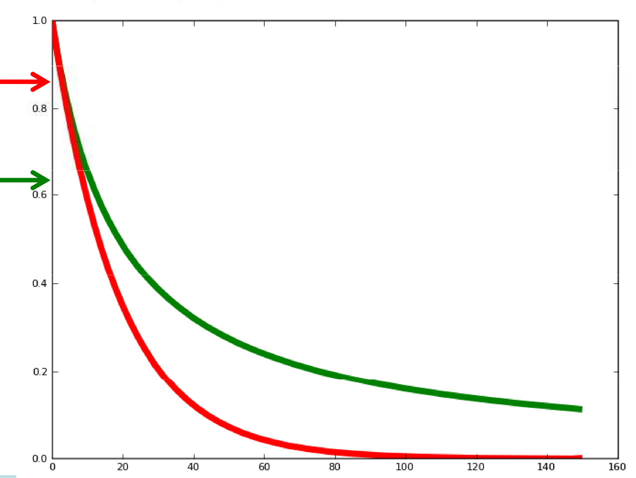
In further experiments there are more kinds of available resources.



•**Hostility of environment**
Environment is not only complex and dynamic. It is also hostile. It means that amount of available resources is limited. We control environment's hostility using two different functions:

$$f_{ci}(k_{ci}) = \frac{1}{1+\frac{k_c}{\tau_c}} \qquad f_{ci}(k_{ci}) = e^{-\frac{k_c}{\tau_c}}$$

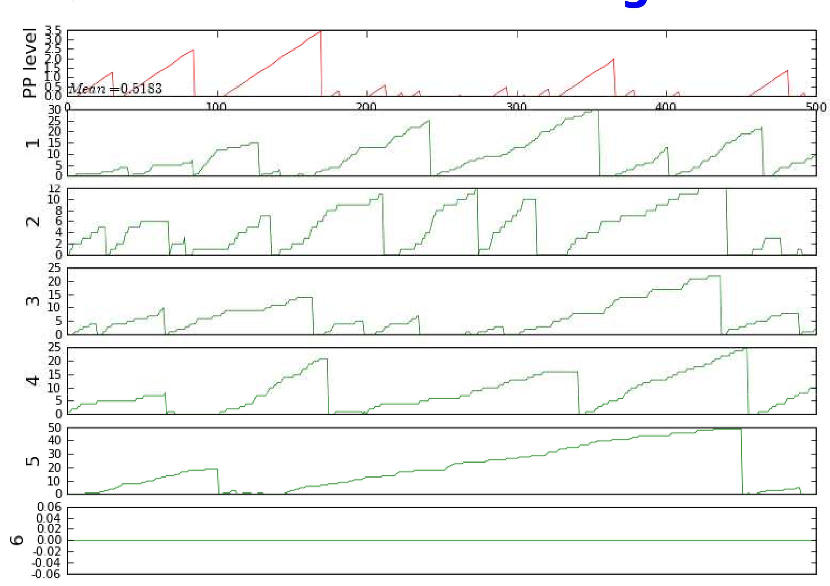
where:
τ_c – scaling factor - describes a resource declining rate
kc – number of times a resource was used



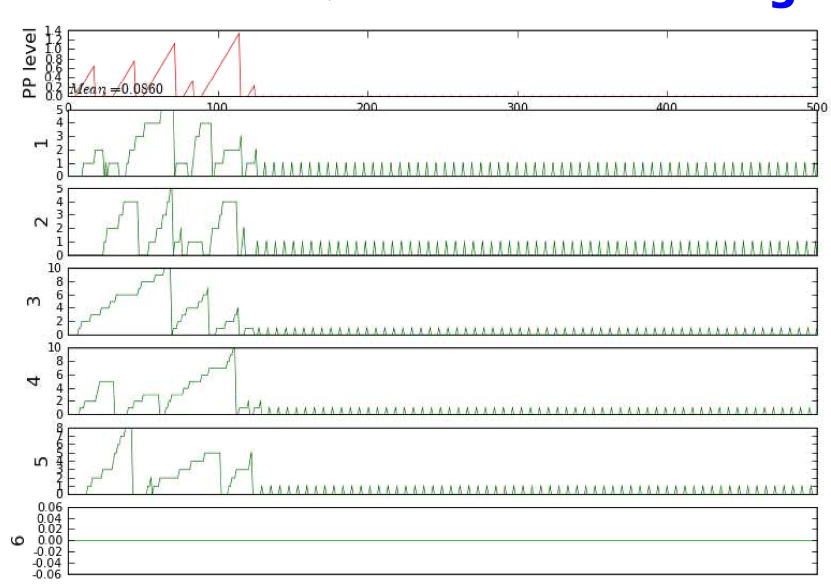
•**Problem**: determine which action should be performed at specific situation and renew this resource which is most needed at this very moment by performing selected action.

Find mappings from sensory inputs to motor output.

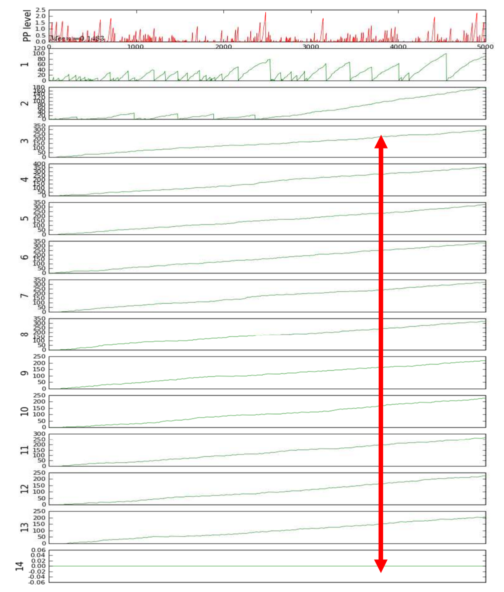
◆Reinforcement Learning



◆Motivated Learning



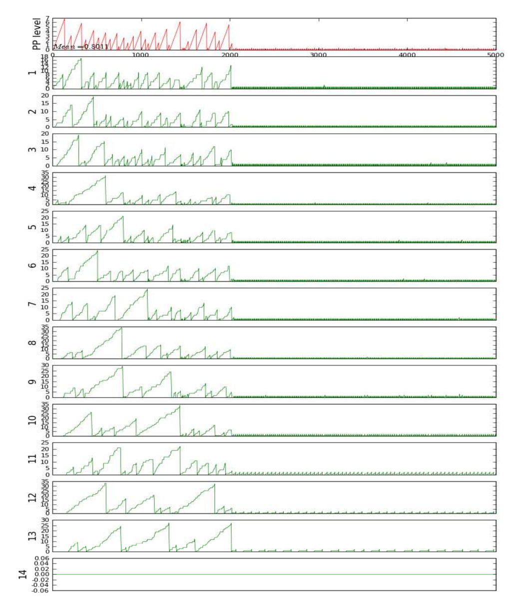
Simulation Results



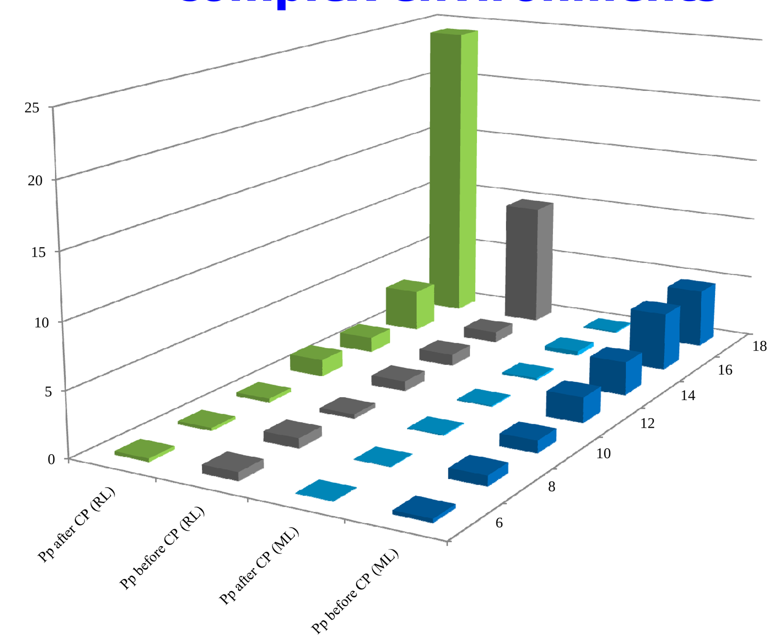
◆Reinforcement Learning

In a complex, dynamically changing environment, the ML works fine, while the reinforcement learning fails to perform and learn these complex relations.

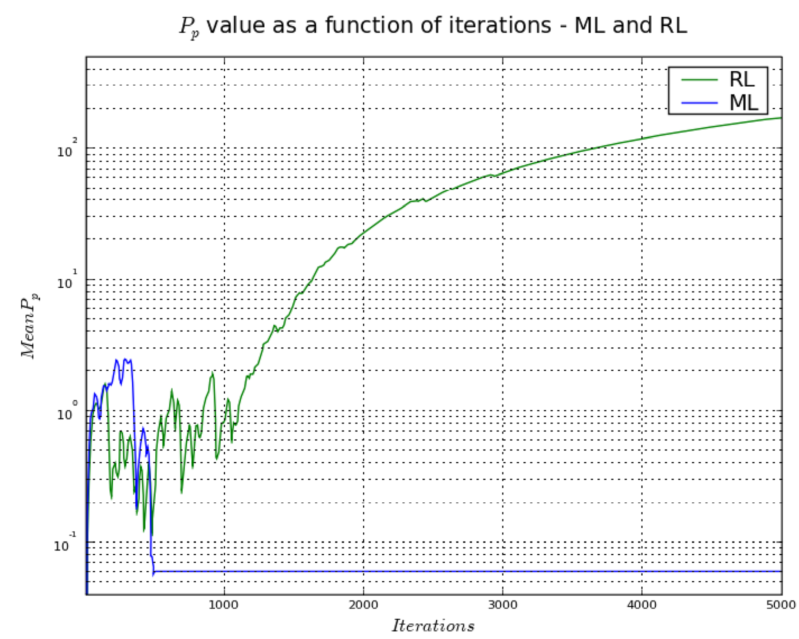
◆Motivated Learning



Task completion ability in more complex environments

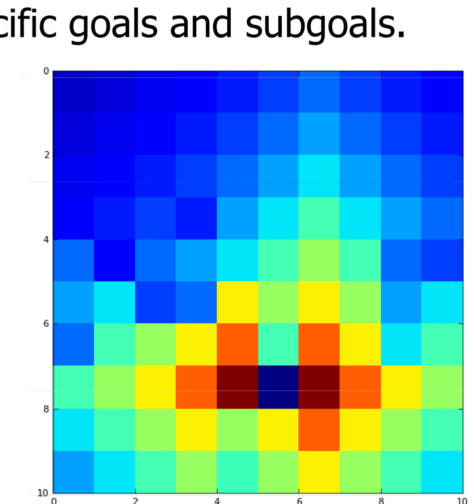


Task completion ability in more hostile environments



Conclusions and future work

•Future work includes combining motivated learning to set abstract motivations and manage goals with reinforcement learning to learn proper actions.
•Motivated learning will provide a self-organizing system of internal motivations and goal selection.
•Reinforcement learning will be used to train machine in solving specific goals and subgoals.
•This will allow to test motivated learning on typical reinforcement learning benchmarks with large dimensionality of the state/action spaces.
•Any form of reinforcement learning e.g. hierarchical reinforcement learning with subgoal discovery can be used.
•Other forms of learning can be used instead of RL, for instance Pavlovian learning proposed by O'Reilly [3].
•The proposed approach enriches machine learning by providing natural goal oriented motivation, that may lead to increase machine intelligence



This work was supported in part by the Singapore National Research Foundation Interactive Digital Media R&D Program, under research Grant NRF2008IDM-IDM004-037

References

[1] J. A. Starzyk, "Motivation in Embodied Intelligence" in Frontiers in Robotics, Automation and Control, I-Tech Education and Publishing, Oct. 2008, pp. 83-110.
[2] A-H. Tan, N. Lu and D. Xiao. Integrating Temporal Difference Methods and Self-Organizing Neural Networks for Reinforcement Learning with Delayed Evaluative Feedback. IEEE *Transactions on Neural Networks*, Vol. 9, No. 2 (February 2008), pp. 230-244.
[3] R. C. O'Reilly, M. J. Frank, T. E. Hazy, B. Watz, PVLV: The Primary Value and Learned Value Pavlovian Learning Algorithm, Behavioral Neuroscience. vol 121(1), Feb 2007, pp. 31-49.