

# Computational Concept and Architecture of Artificial Psyche

Anton Kolonin

[akolonin@aigents.com](mailto:akolonin@aigents.com)

Telegram: akolonin

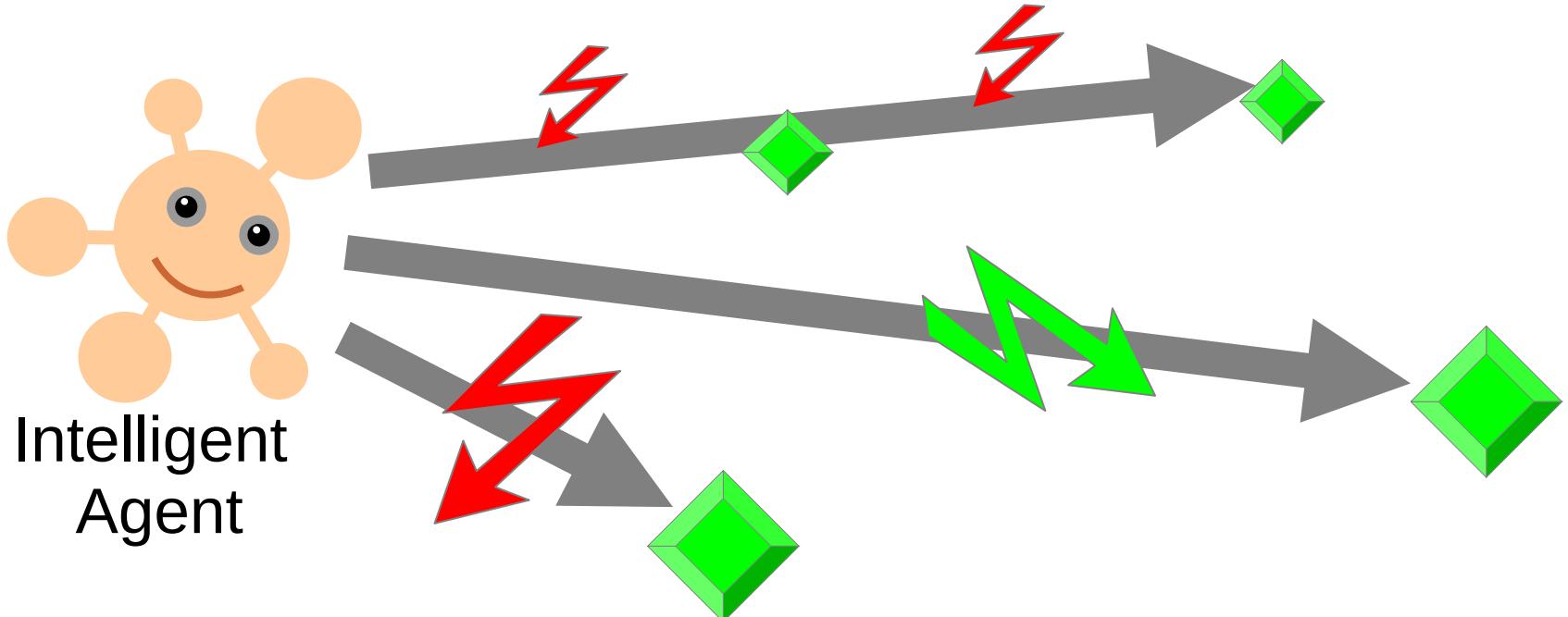


<https://agirussia.org>

# General Intelligence:

## Reaching complex goals in different complex environments, using limited resources under uncertainty

(Ben Goertzel + Pei Wang + Shane Legg + Marcus Hutter)

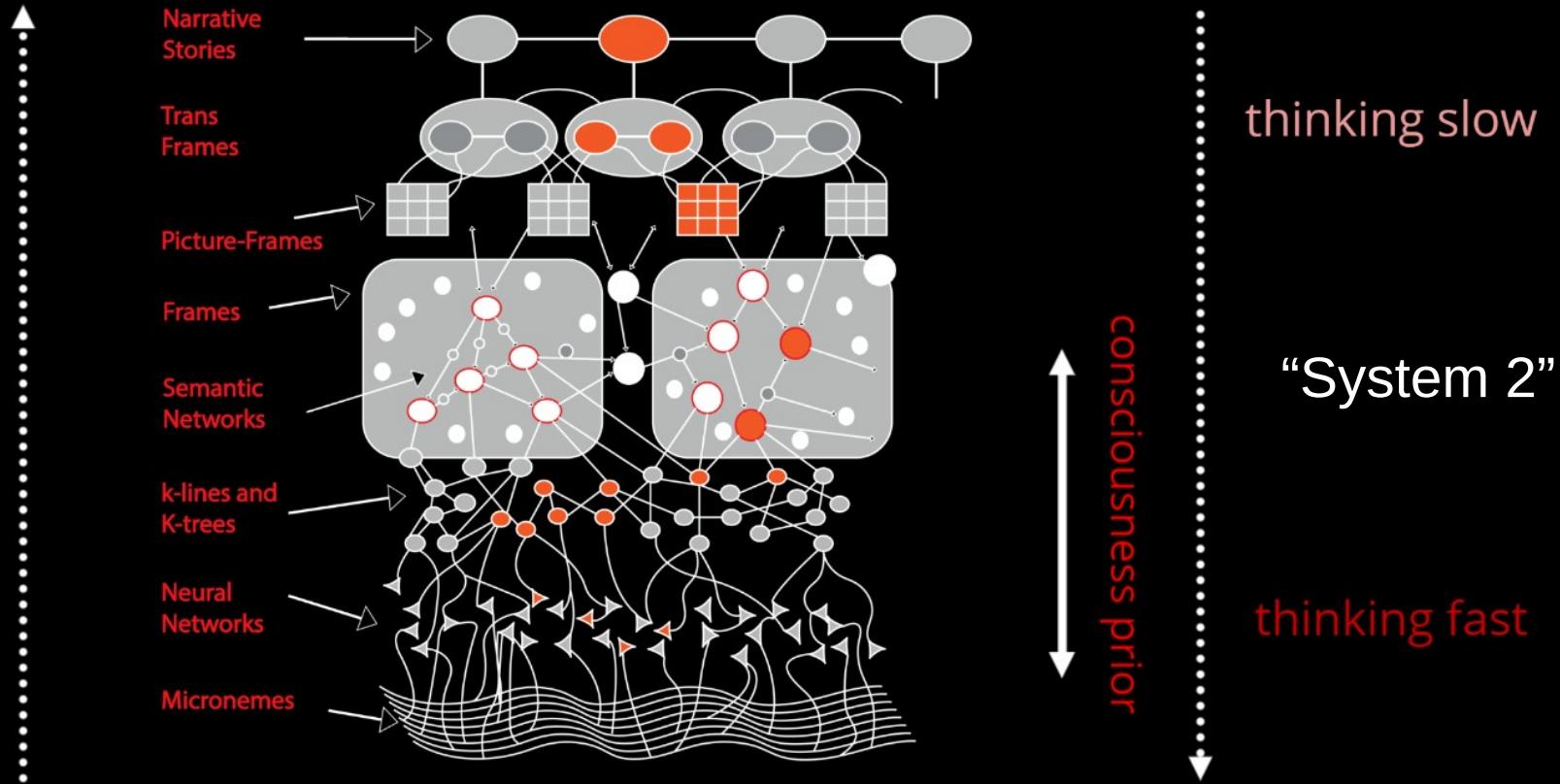


# “Fast and Slow Thinking” – Daniel Kahneman

easy  
explanation  
learning fast

“System 1”

hard  
explanation  
learning slow



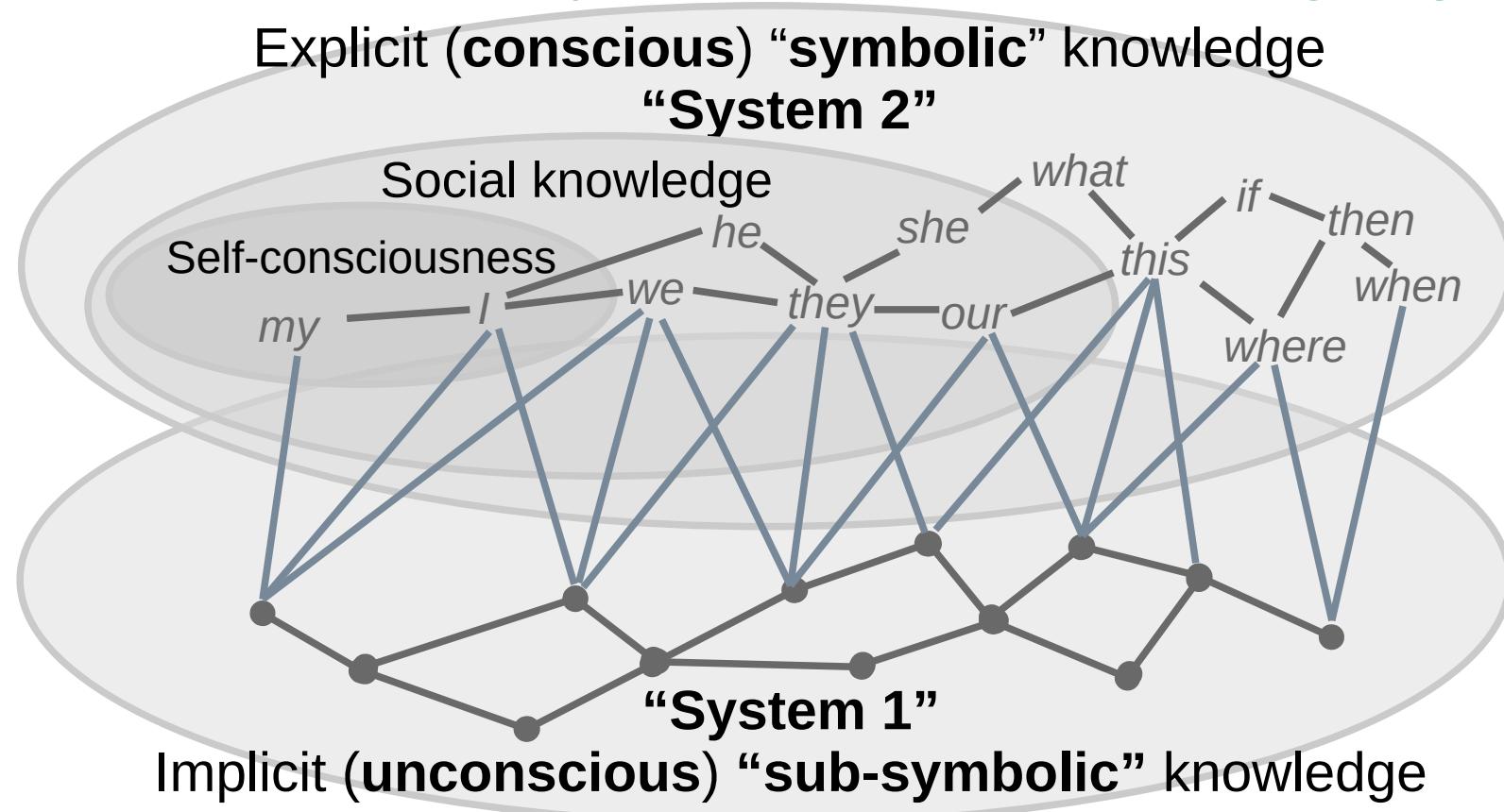
<https://www.linkedin.com/pulse/explainable-ai-vs-explaining-part-1-ahmad-haj-mosa/>

Xing, F., Cambria, E., Welsch, R. (2019). Theoretical Underpinnings on Text Mining. In: Intelligent Asset Management. Socio-Affective Computing, vol 9. Springer, Cham.  
[https://doi.org/10.1007/978-3-030-30263-4\\_3](https://doi.org/10.1007/978-3-030-30263-4_3)

M. Minsky, The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind (Simon & Schuster Paperbacks, Princeton, 2007)

Copyright © 2025 Anton Kolonin, Agents®

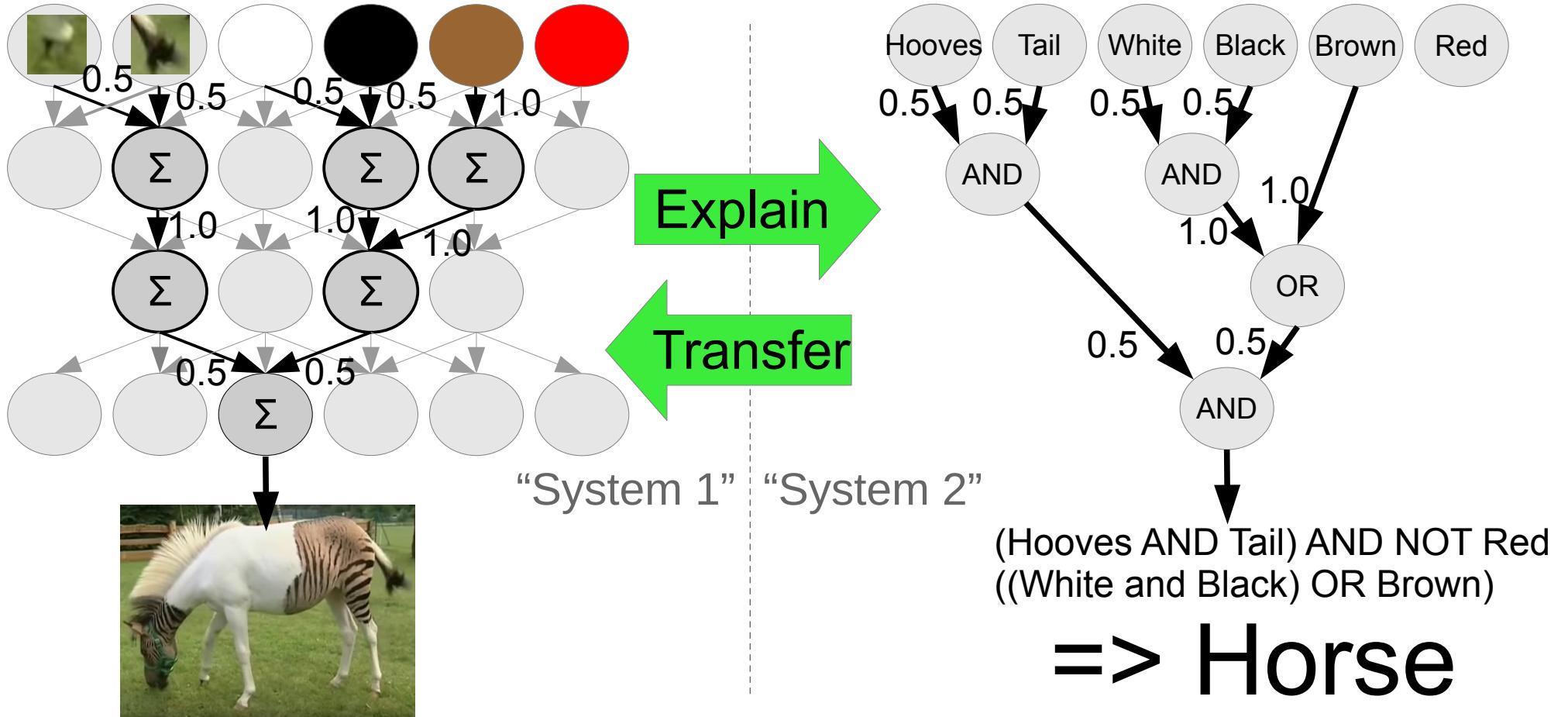
# Medium: “neuro-symbolic” “knowledge graph”



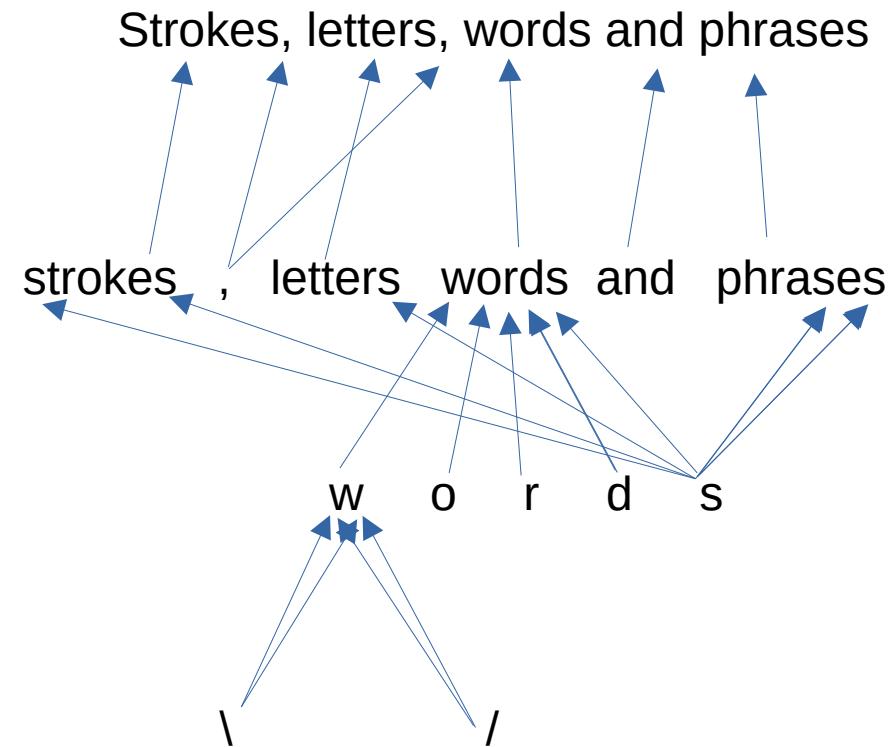
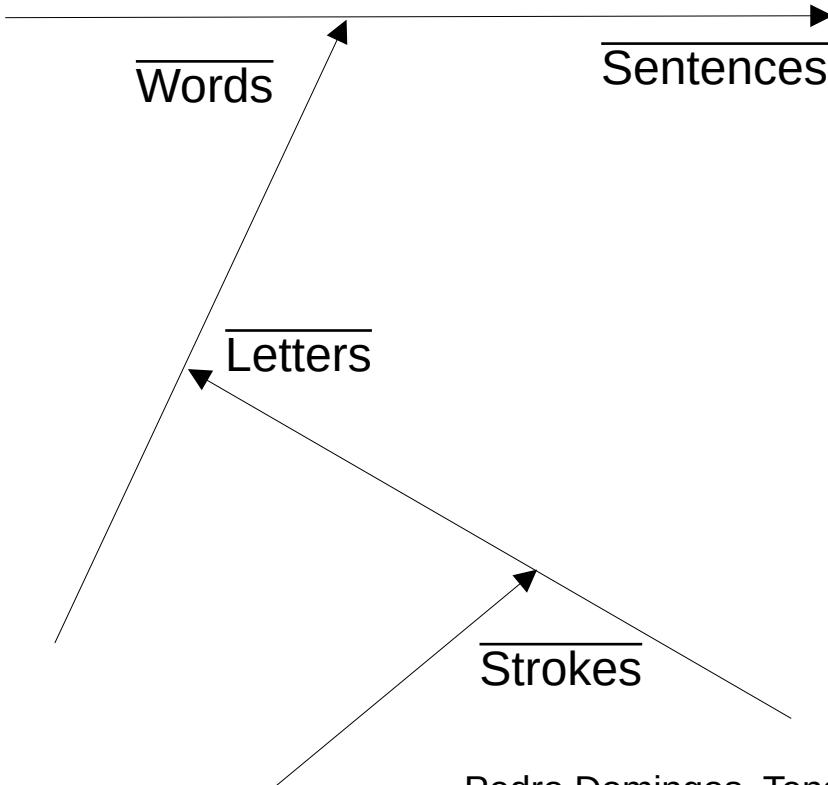
<https://www.amazon.com/Thinking-Fast-Slow-Daniel-Kahneman/dp/0374533555>

<https://amit02093.medium.com/atomspace-hyper-graph-information-retrieval-system-450cab9d751e>

# Neuro-Symbolic Integration for Interpretable AI



# Functional equivalence of neural network tensor and graph (symbolic) models

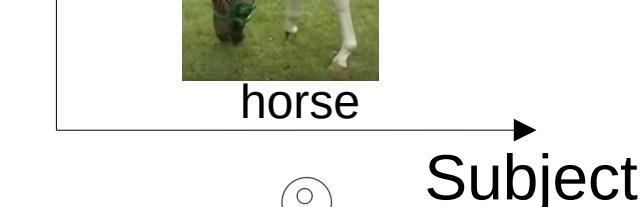
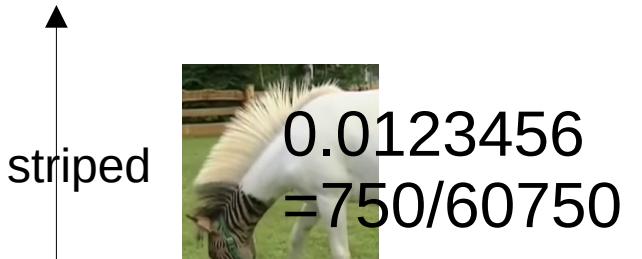


Pedro Domingos, Tensor Logic: The Language of AI  
<https://arxiv.org/pdf/2510.12269>

# Typed tensor logic for different kinds of AI-s (logical, sub-symbolic, probabilistic/non-axiomatic)

**Truth-Value Tensor**  
(NARS/PLN)

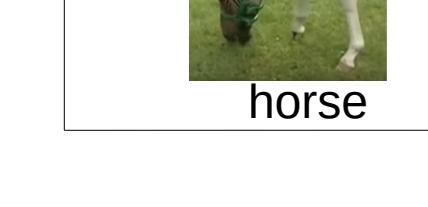
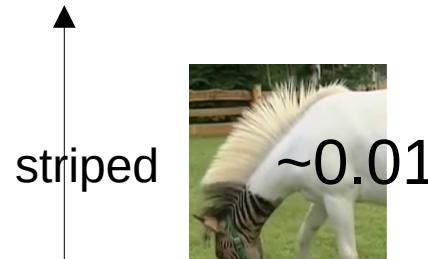
Property



Life-long  
learning?

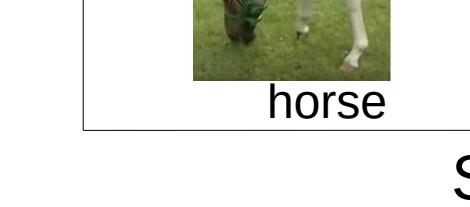
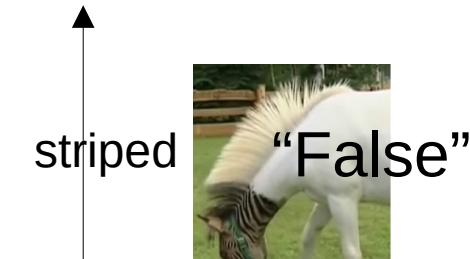
**Numerical Tensor**  
(ANN/Bayesian Logic)

Property



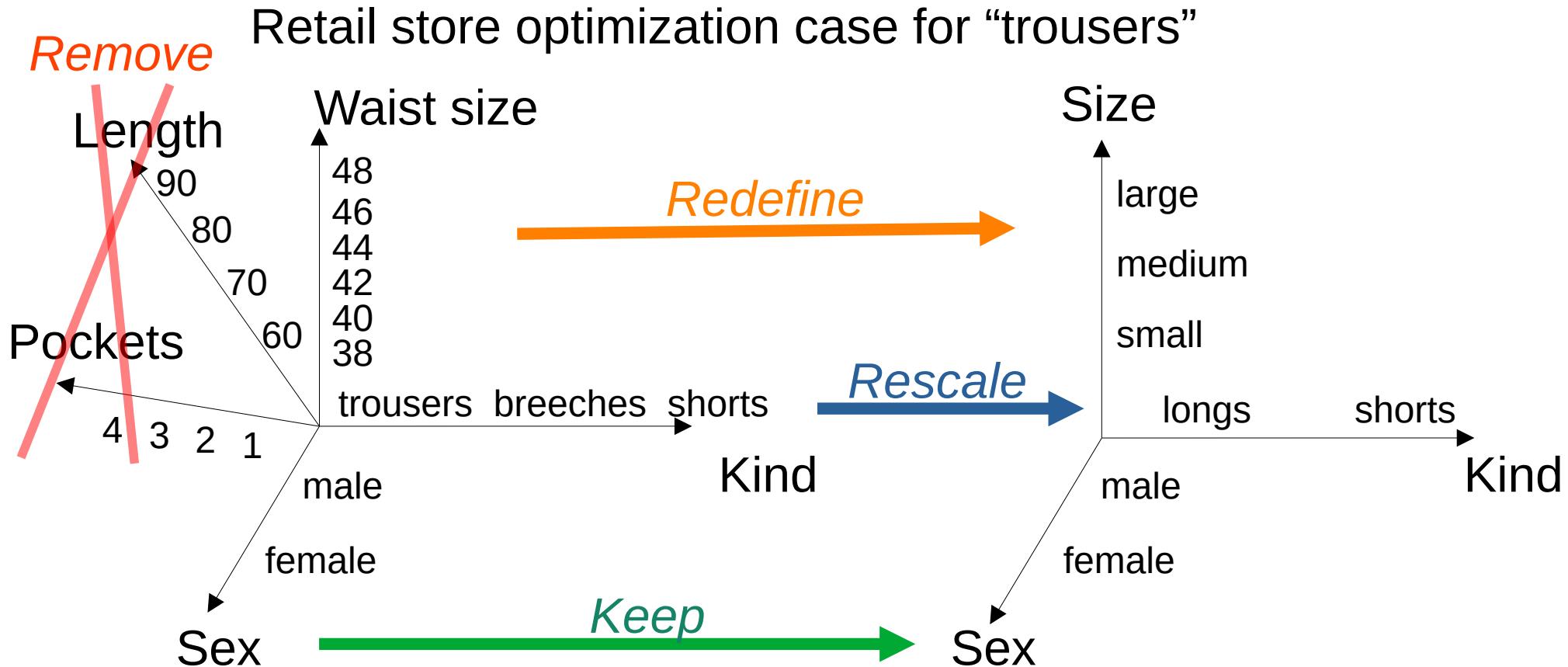
**Boolean Tensor**  
(Boolean Logic)

Property



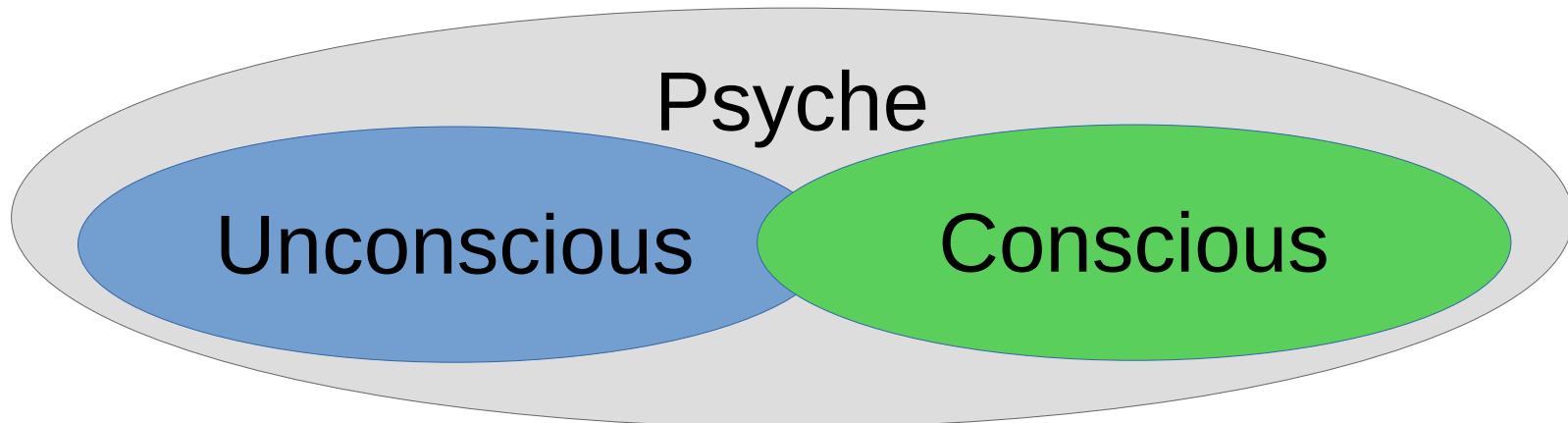
Pedro Domingos, Tensor Logic: The Language of AI  
<https://arxiv.org/pdf/2510.12269>

# Problem of dimensionality (reduction) and discreteness (increase)



The psyche /'saɪki/ is currently used to describe the totality of the **human mind**, **conscious** and **unconscious**.

<https://en.wikipedia.org/wiki/Psyche>



# System = Intelligent agent

Psyche = Operating system

Intelligence = Decision making system

Subconsciousness  
(Intuition)  
“System 1”  
 (“Fast”)

Consciousness  
(Reasoning)  
“System 2”  
 (“Slow”)

*Emotions* → *Motivation* → *Expectations*

Space of states

Sensations      Needs      Actions      Decisions

*Perceptions*

Sensors

Actuators

*Sensations*

*Actions*

Outer world = Operational environment

# Psyche = Operating system

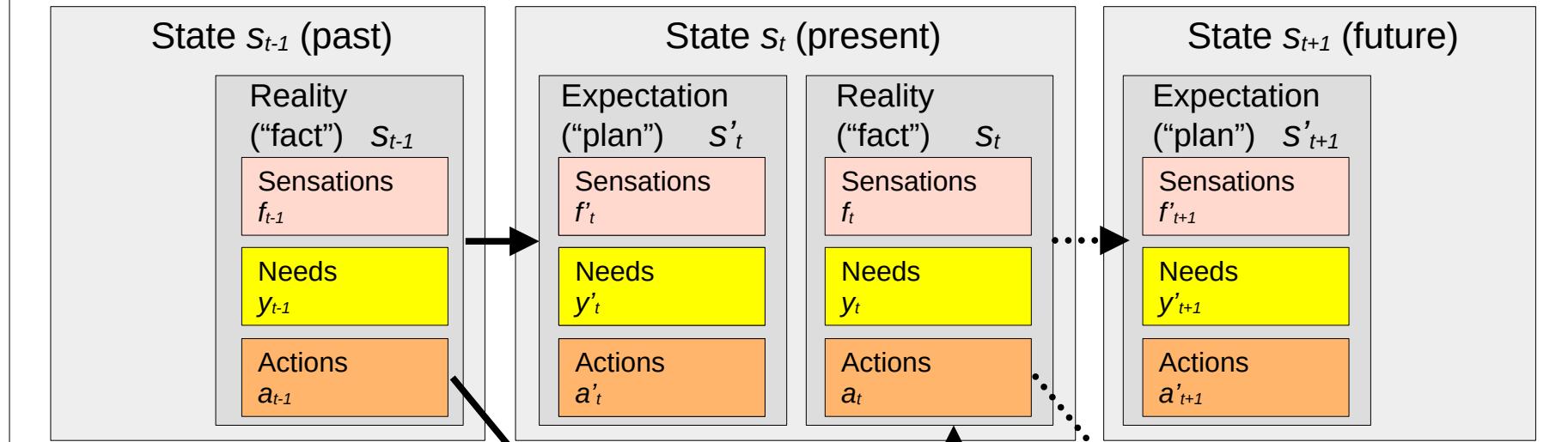
Intelligence = Decision making system

Models  $s$  ("invariants") of states with utilities  $U$  and probabilities  $P$  of transitions  
 $U(\{S_T\}_{T \in \{t-T, t-1\}}, S'_0) = L(x \cdot (y_t - y_{t-1}), (s'_t - s_t), E(a_{t-1}))$        $s'_t = \text{argmax}_s(U(\{S_T\}_{T \in \{t-T, t-1\}}, S'_t), P(\{S_T\}_{T \in \{t-T, t-1\}}, S'_t))$

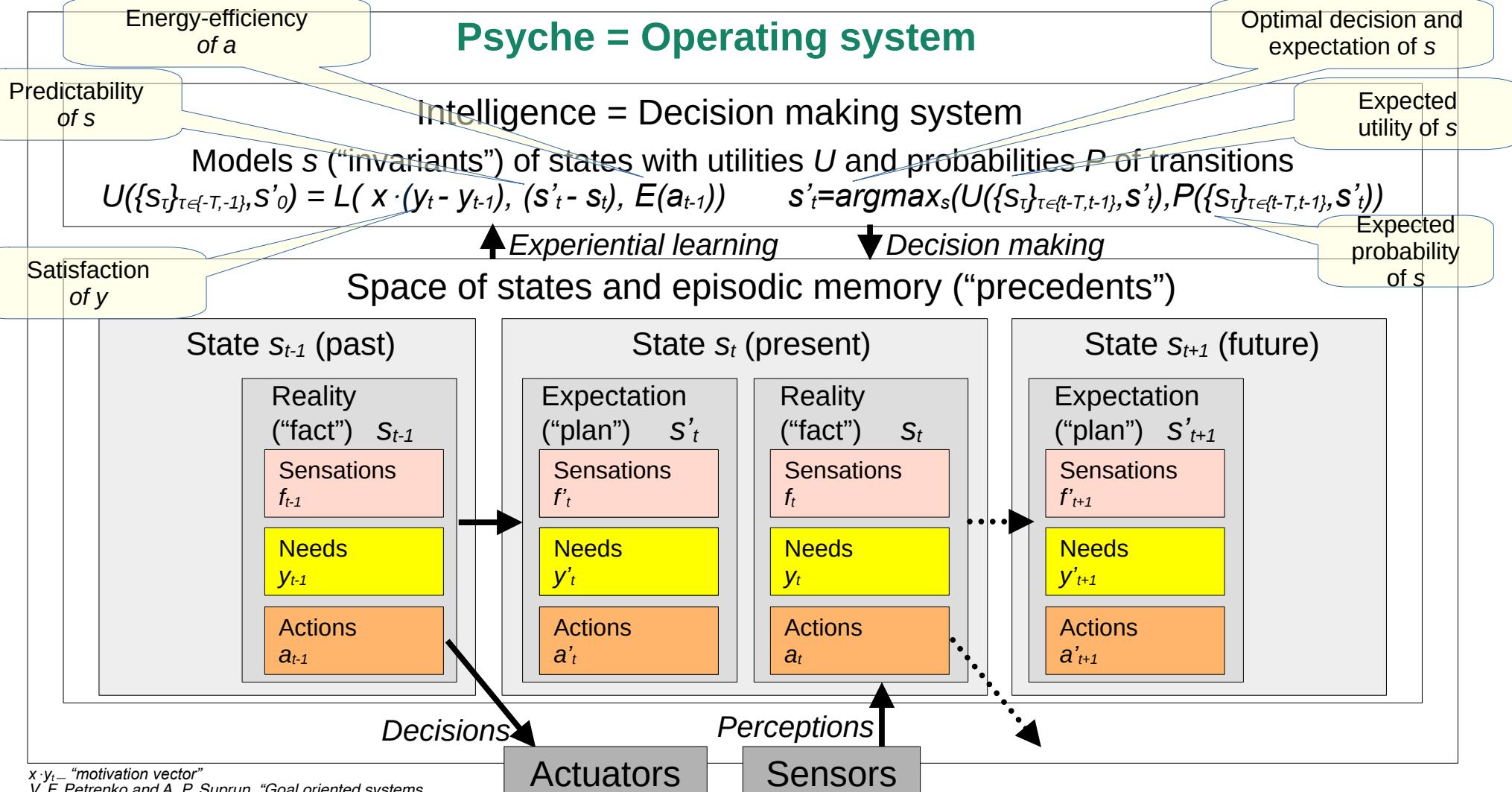
↑Experiential learning

↓Decision making

Space of states and episodic memory ("precedents")



# Psyche = Operating system

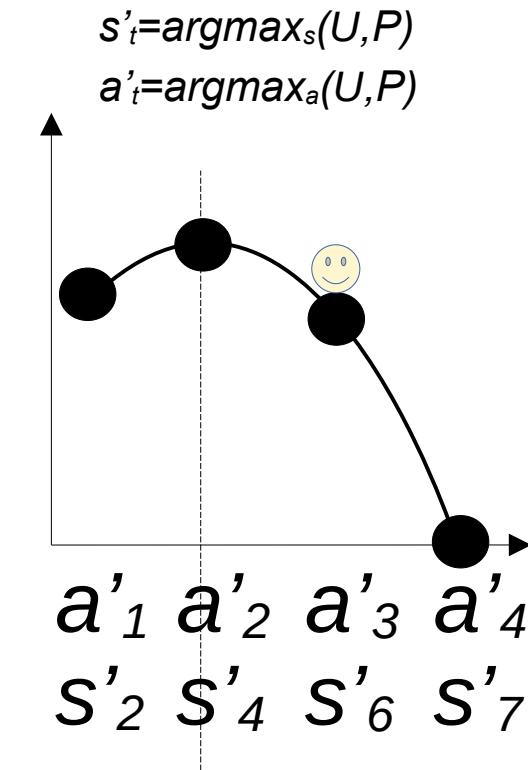


$x \cdot y_t$  – "motivation vector"

V. F. Petrenko and A. P. Suprun, "Goal oriented systems, evolution, and the subjective aspect in systemology," Tr. Inst. Sistem. Analiza RAN 62 (1) (2012)

# Decision making as operational risk management

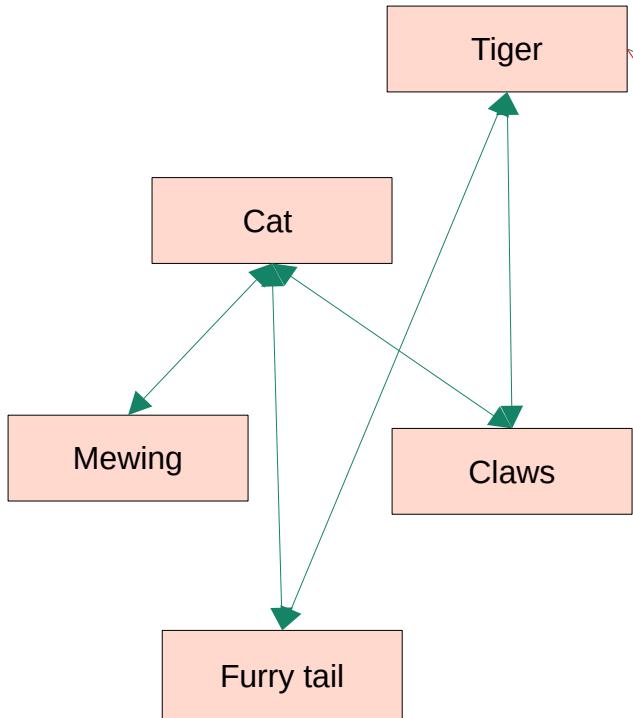
$s_t$	$s'_{t+1}$	$s'_{t+1}$			$U$	$P$	$\sum U^*P$
		$a'$	$y'$	$f'$			
$s_1$	$s'_2$	$a'_1$	$y'_1$	...	1.0	0.5	<u>0.7</u>
$s_1$	$s'_3$	$a'_1$	$y'_2$	...	0.4	0.5	
$s_1$	$s'_4$	$a'_2$	$y'_3$	...	1.0	0.8	<u>0.8</u>
$s_1$	$s'_5$	$a'_2$	$y'_4$	...	0.0	0.2	
$s_1$	$s'_6$	$a'_3$	$y'_5$	...	0.6	1.0	<u>0.6</u>
$s_1$	$s'_7$	$a'_4$	$y'_6$	...	0.0	1.0	<u>0.0</u>



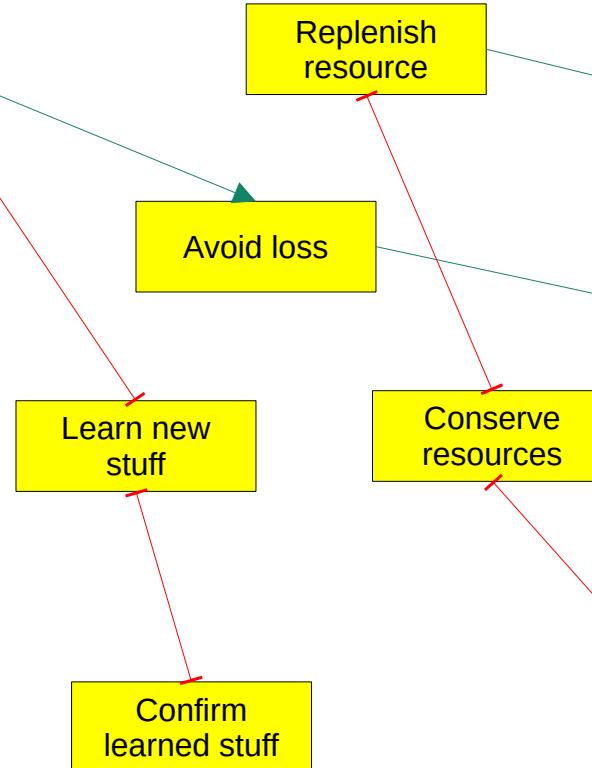
Tversky & Kahneman:  
most people choose  $a'_3$  и  $s'_6$   
("smaller profit with greater reliability")

# Mutual dependency of state variable subgraphs tensors

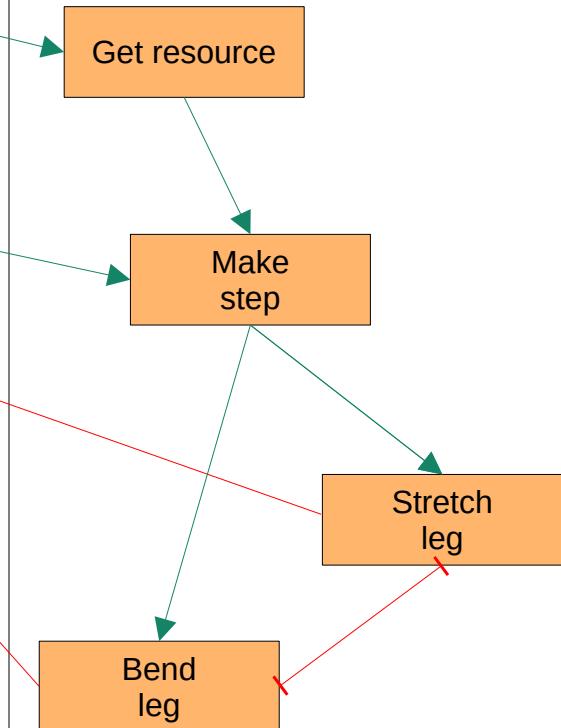
Percept feelings



Satisfied needs



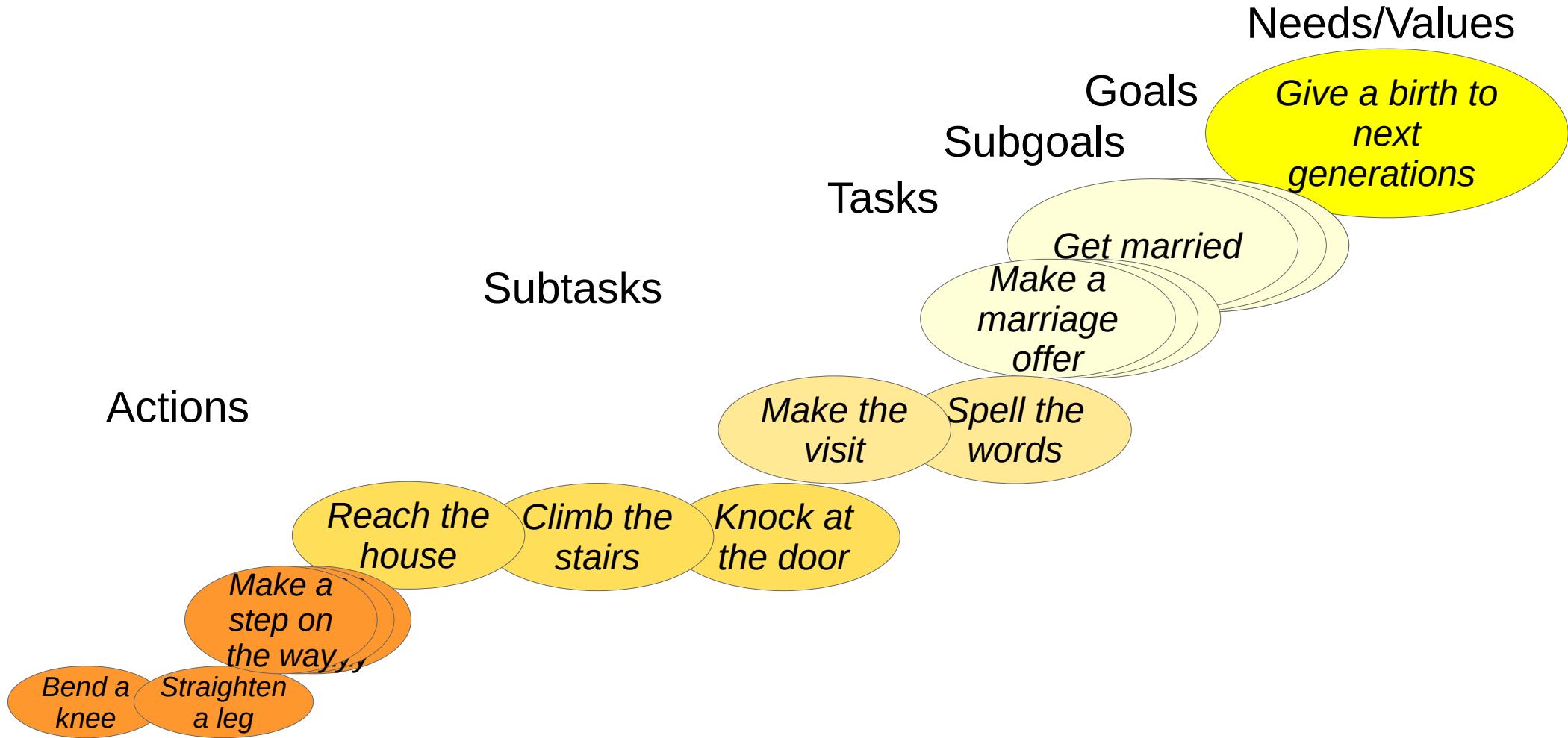
Decided & committed actions



Anton Kolonin & Vladimir Kryukov,  
Computational Concept of the Psyche,  
Neuroinformatics-2025

<https://arxiv.org/pdf/2509.07009>

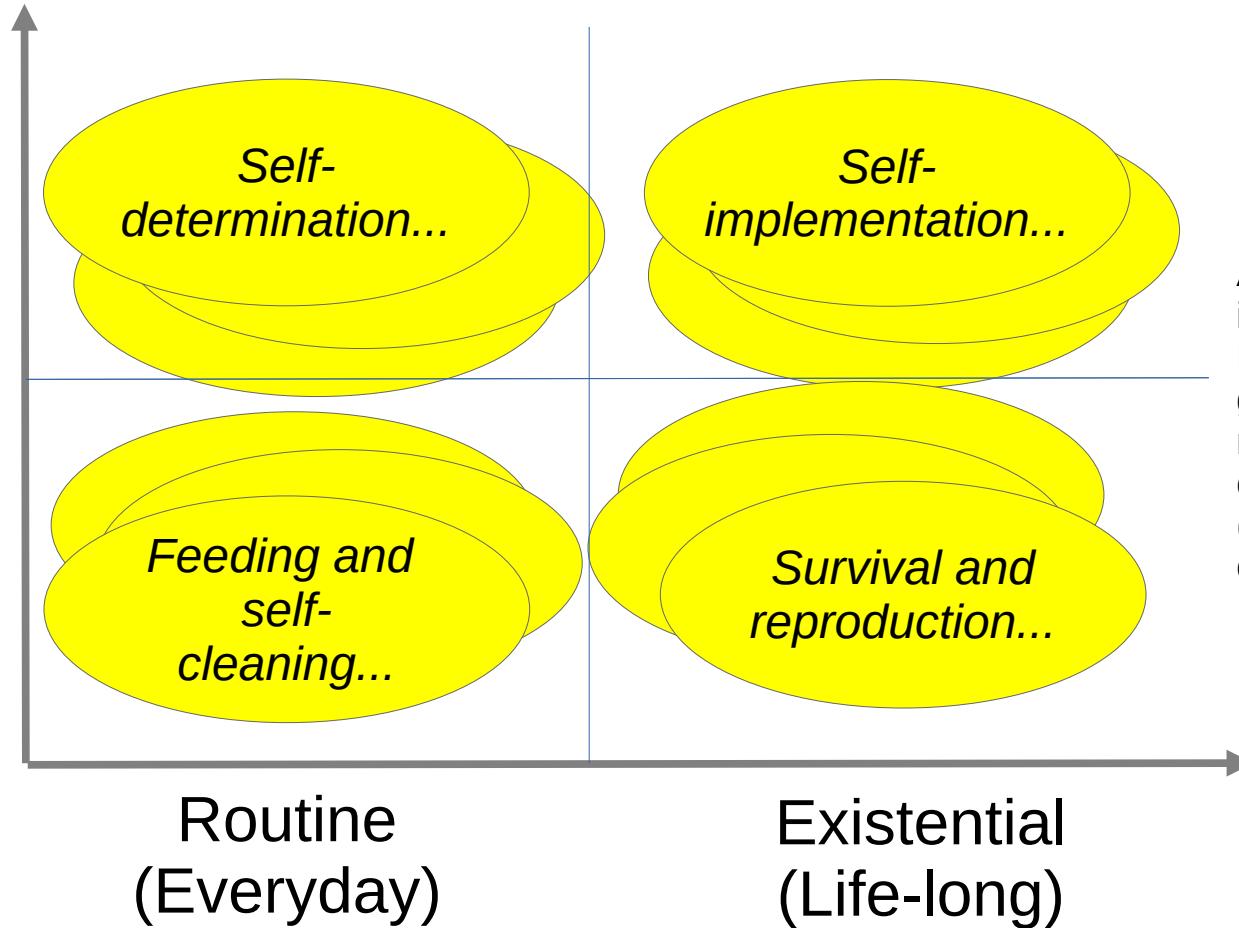
# Hierarchy of values/goals/subgoals/tasks/subtasks



# Space of needs/goals/values

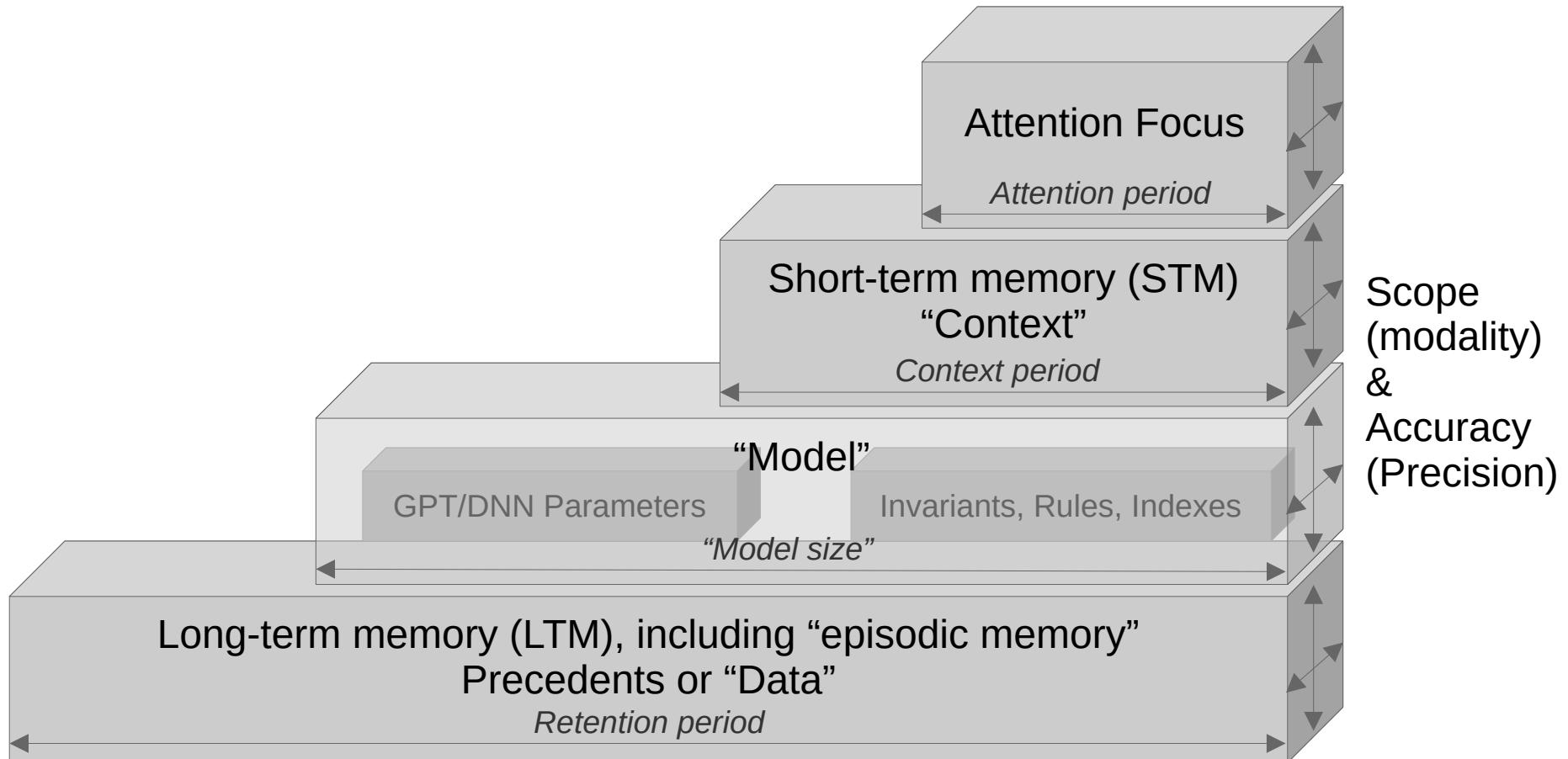
Psychological

Physical  
(Physiological)

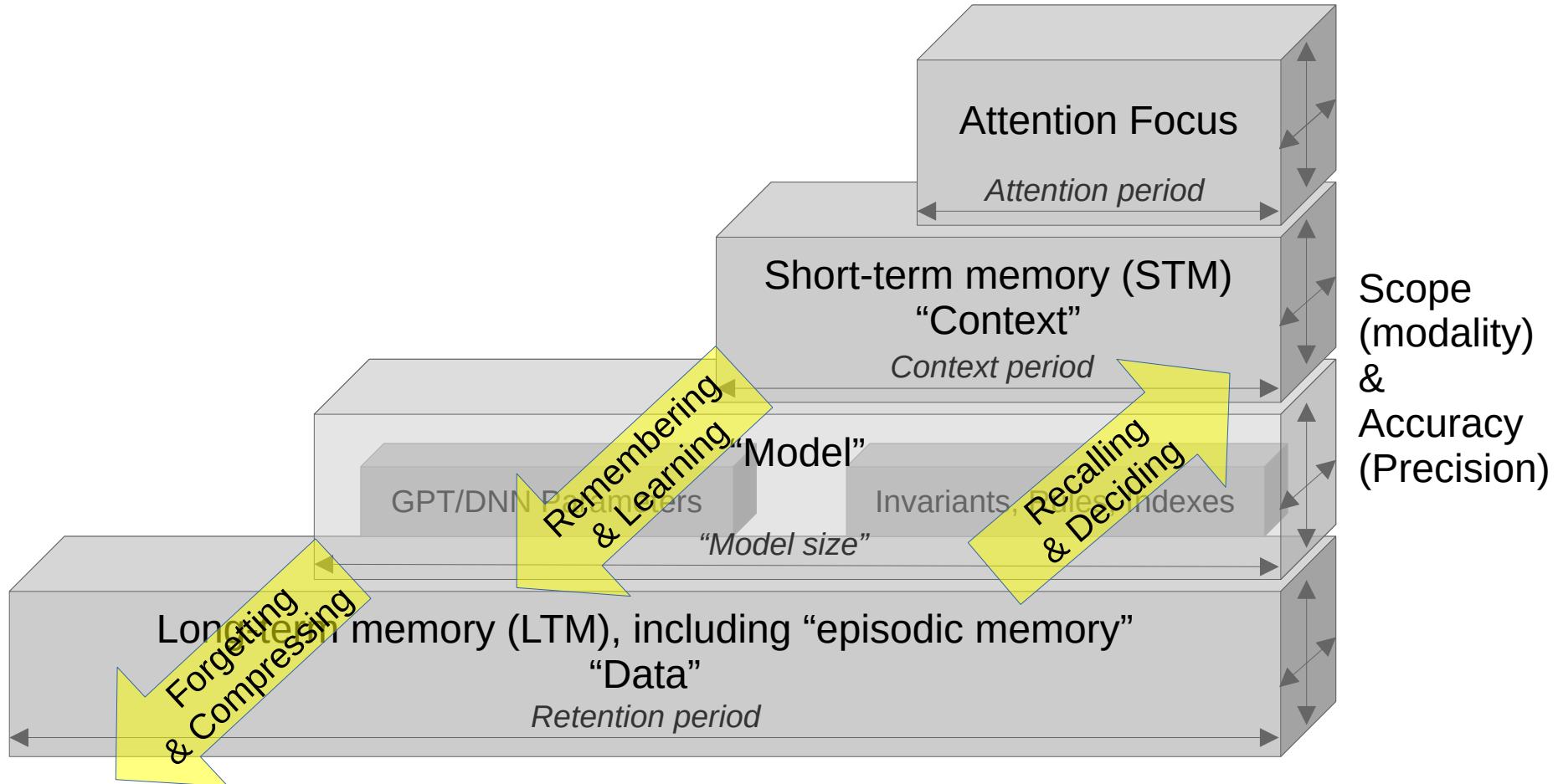


According to the ideas of A. Maslow, any goal can be represented as a combination (superposition) of basic needs

# Attention, context, model and memories

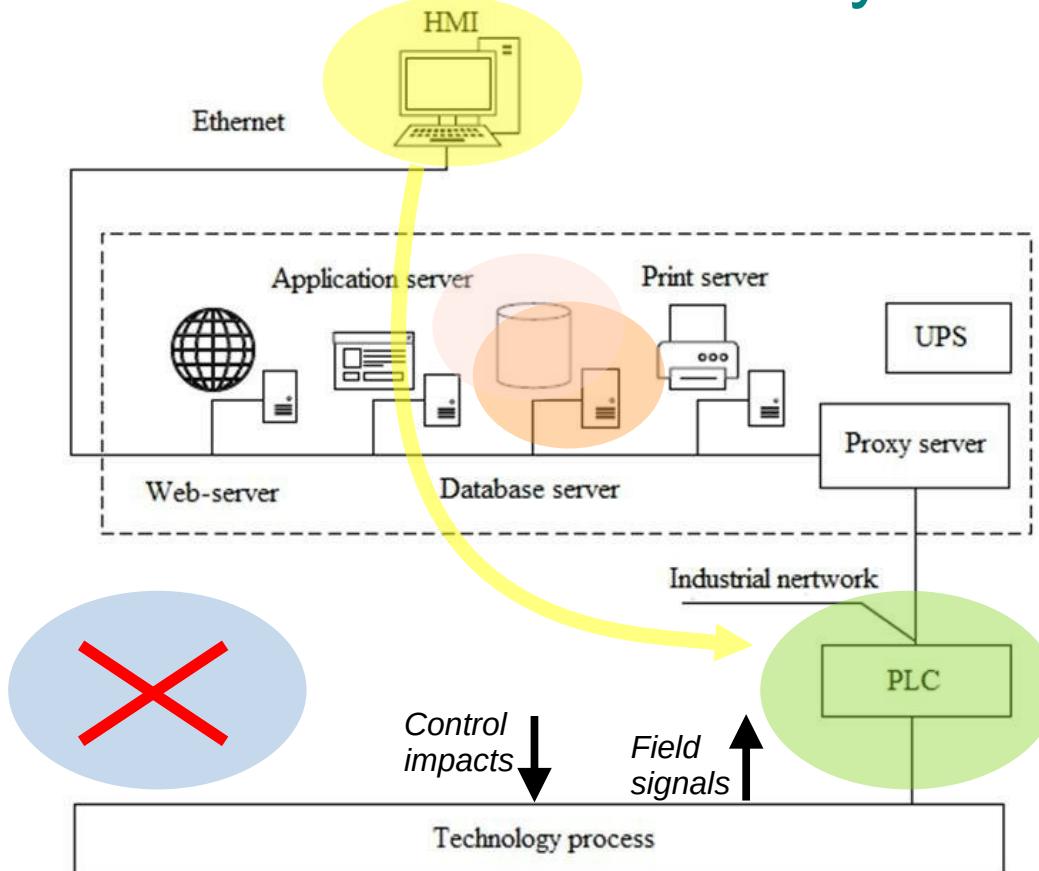


# Attention, context, model and memories



# Application cases

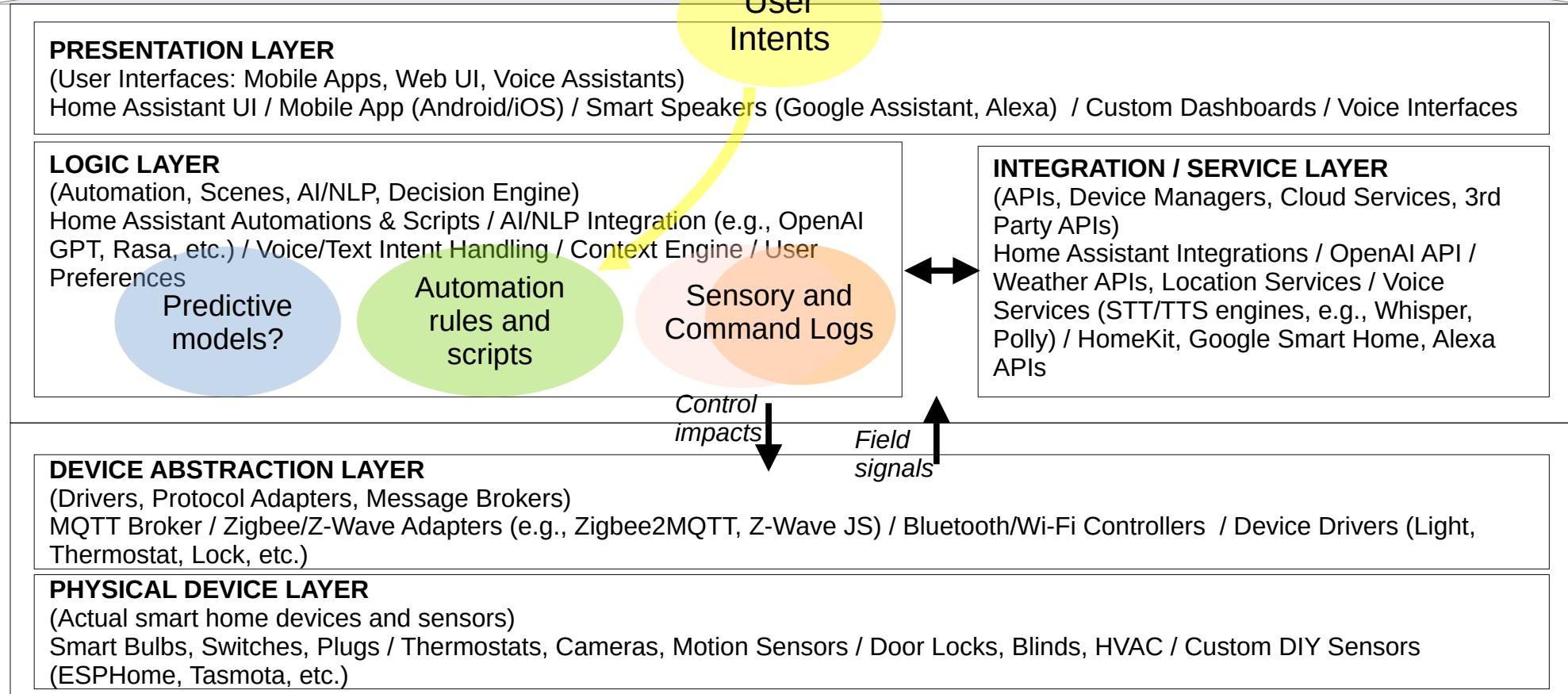
## A) Automated Process Control Systems (APCS)



Programmable logic  
**IEC 61131** (synchronous)  
**IEC 61499** (asynchronous)

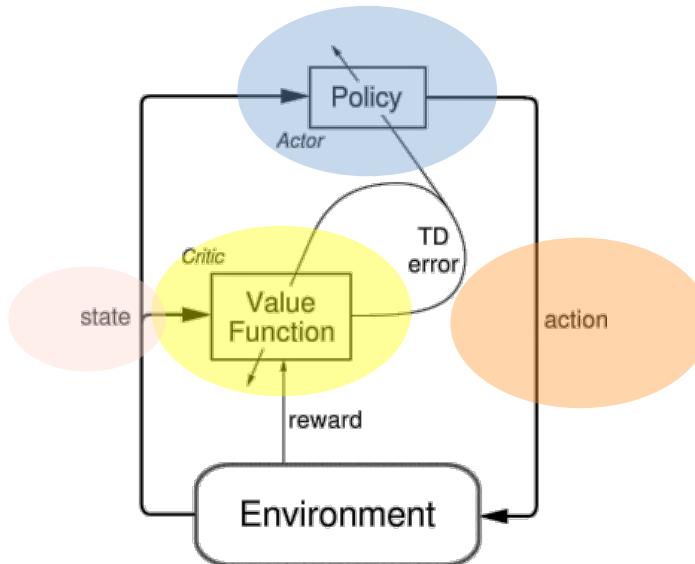
# Application cases

## B) “Smart Home”



# Implementation options

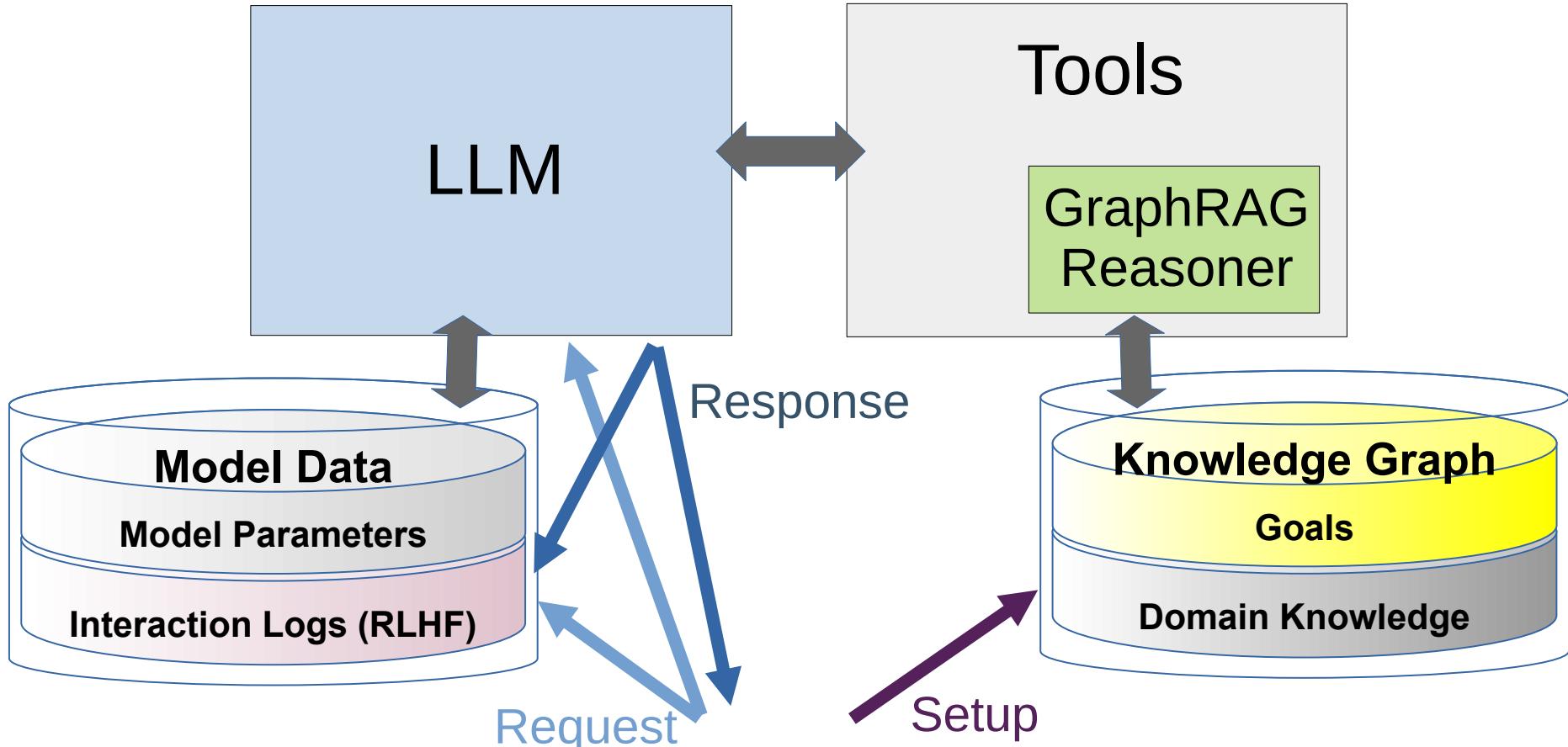
## 1) Reinforcement learning based on “actor-critic” model



<https://medium.com/intro-to-artificial-intelligence/the-actor-critic-reinforcement-learning-algorithm-c8095a655c14>

# Implementation options

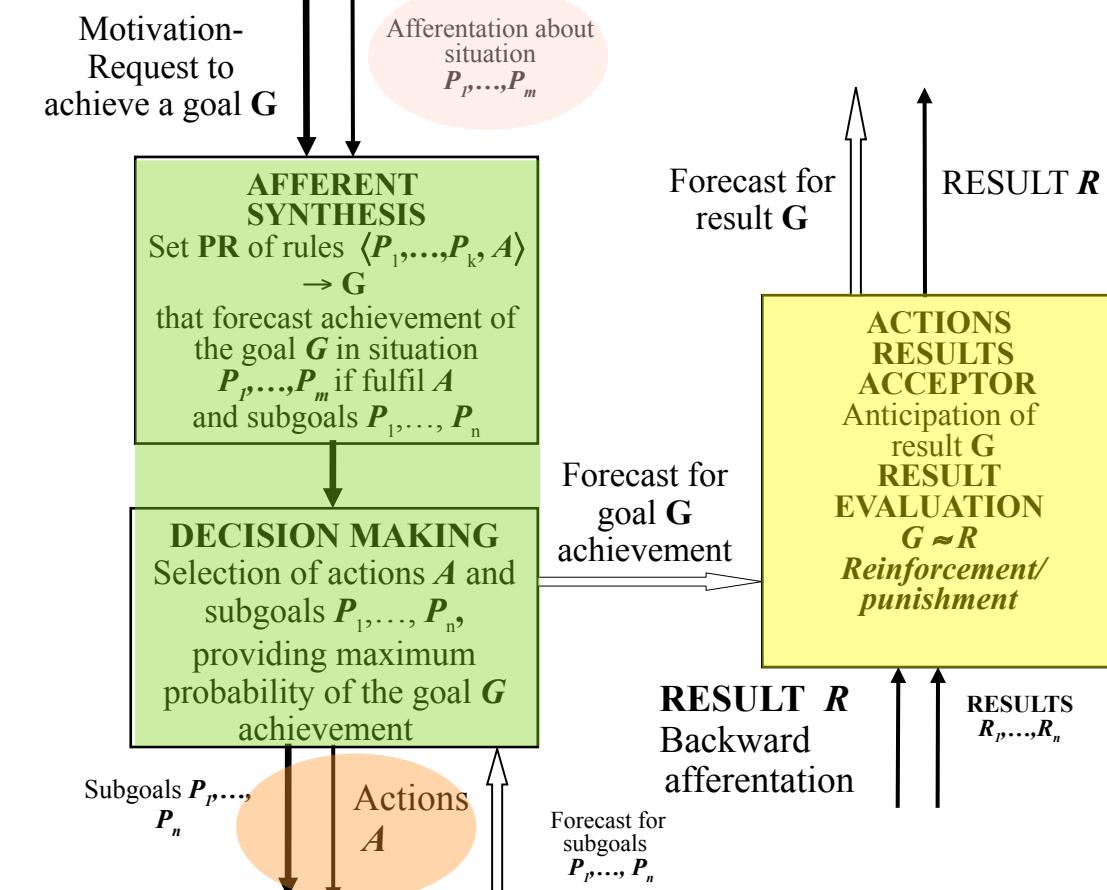
## 2) Cognitive architecture based on LLM and (active) GraphRAG



Leonie Monigatti, "The Evolution from RAG to Agentic RAG to Agent Memory", November 3, 2025  
<https://www.leoniemonigatti.com/blog/from-rag-to-agent-memory.html>

# Implementation options

## 3) Cognitive architecture based on probabilistic logic



Copyright © 2025 Anton Kolonin, Aigents®

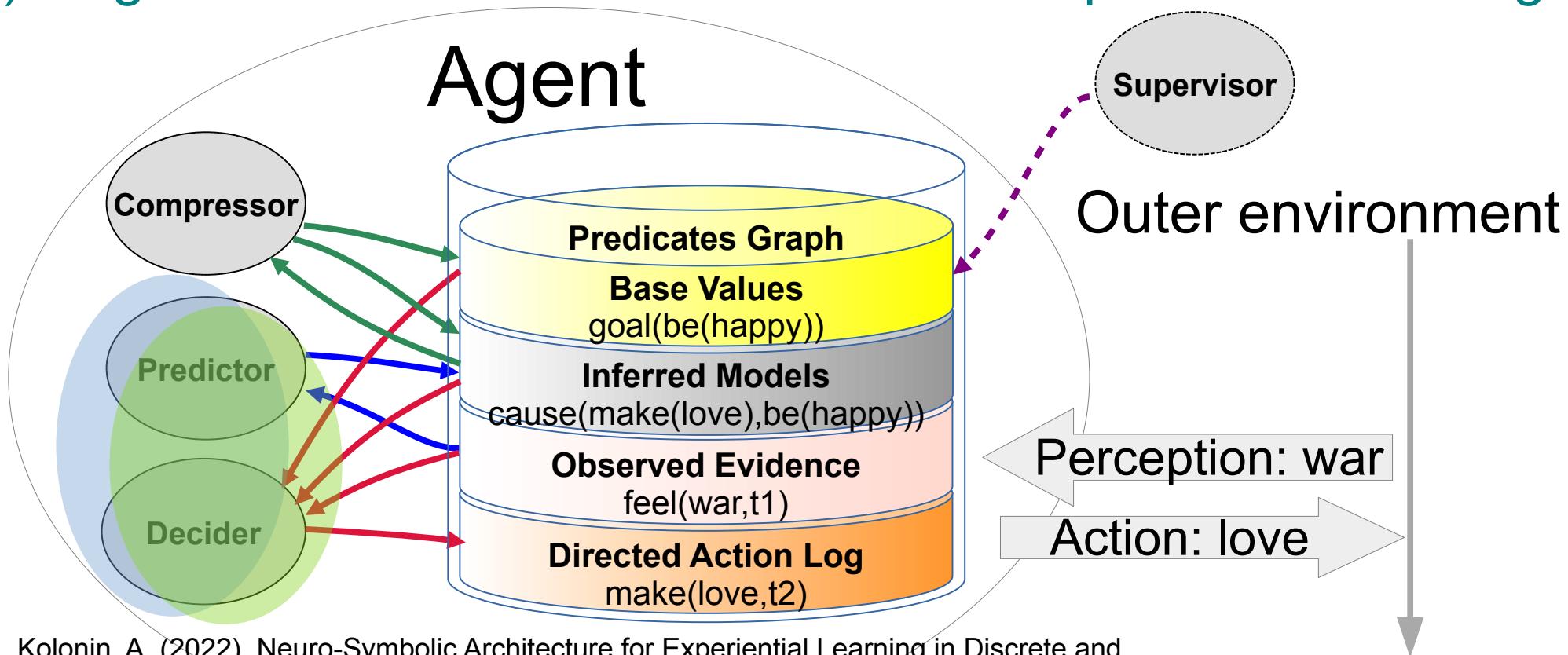
Evgenii Vityaev, Alexander Demin:  
Adaptive Control of Modular Robots // Conference Paper in Advances in Intelligent Systems and Computing, Conference: First International Early Research Career Enhancement School on Biologically Inspired Cognitive Architectures, Springer, August 2018

Evgenii E. Vityaev: Purposefulness as a Principle of Brain Activity // Anticipation: Learning from the Past, (ed.) M. Nadin. Cognitive Systems Monographs, V.25, Chapter No.: 13. Springer, 2015, pp. 231-254.

Витяев Е.Е. Логика работы мозга. Подходы к моделированию мышления. (сборник под ред. д.ф.-м.н. В.Г. Редько). УРСС Эдиториал, Москва, 2014г., стр. 120-153.

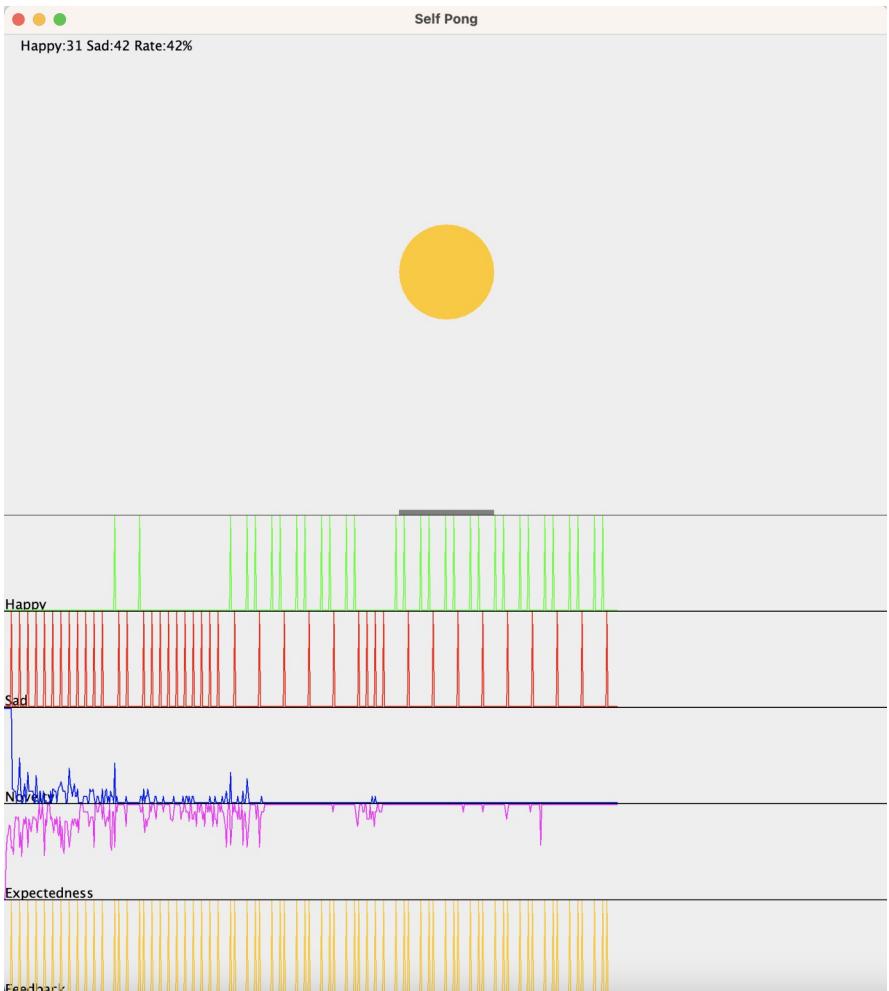
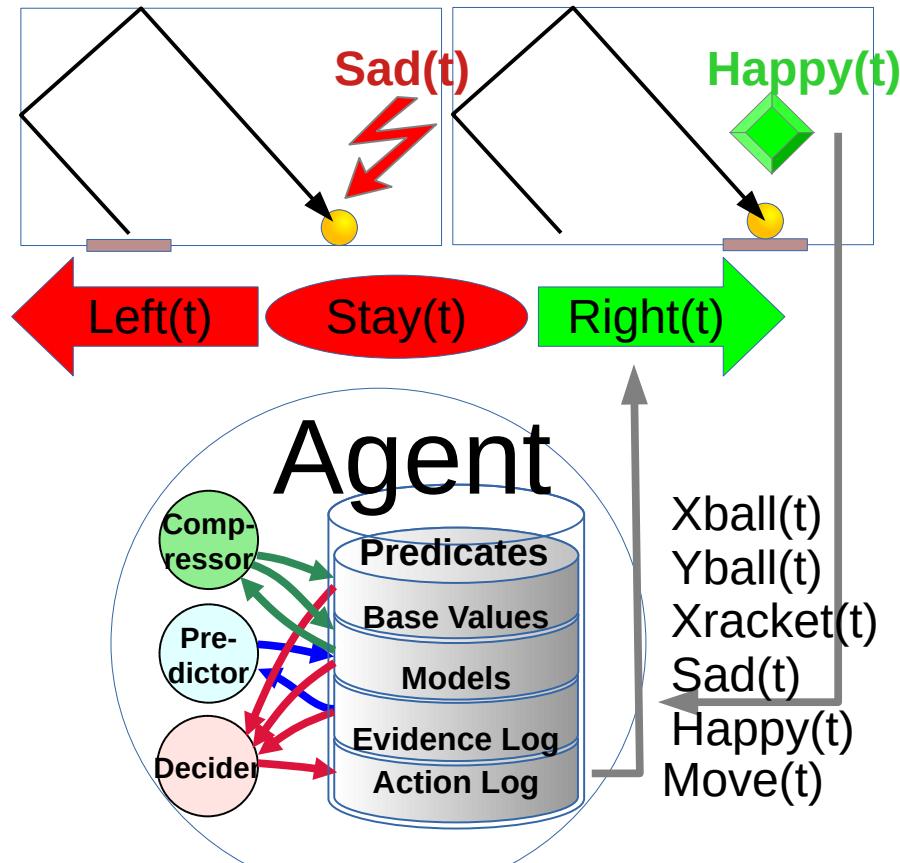
# Implementation options

## 4) Cognitive architecture of value-based experiential learning



Kolonin, A. (2022). Neuro-Symbolic Architecture for Experiential Learning in Discrete and Functional Environments. In: Goertzel, B., Iklé, M., Potapov, A. (eds) Artificial General Intelligence. AGI 2021. Lecture Notes in Computer Science(), vol 13154. Springer, Cham.  
[https://doi.org/10.1007/978-3-030-93758-4\\_12](https://doi.org/10.1007/978-3-030-93758-4_12)

# Cognitive architecture of value-based experiential learning



Anton Kolonin & Vladimir Kryukov, Computational Concept of the Psyche

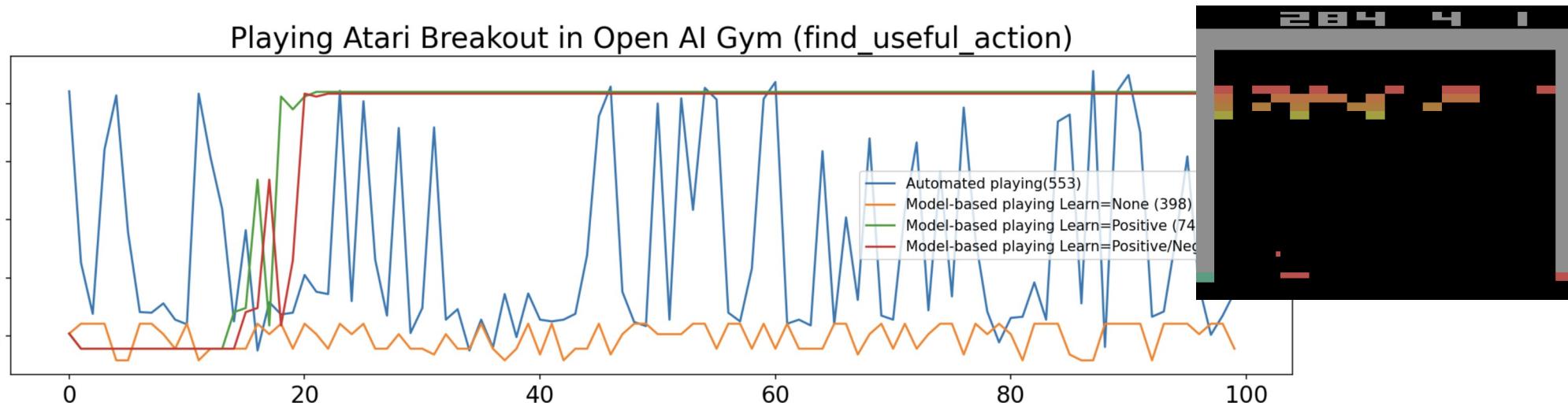
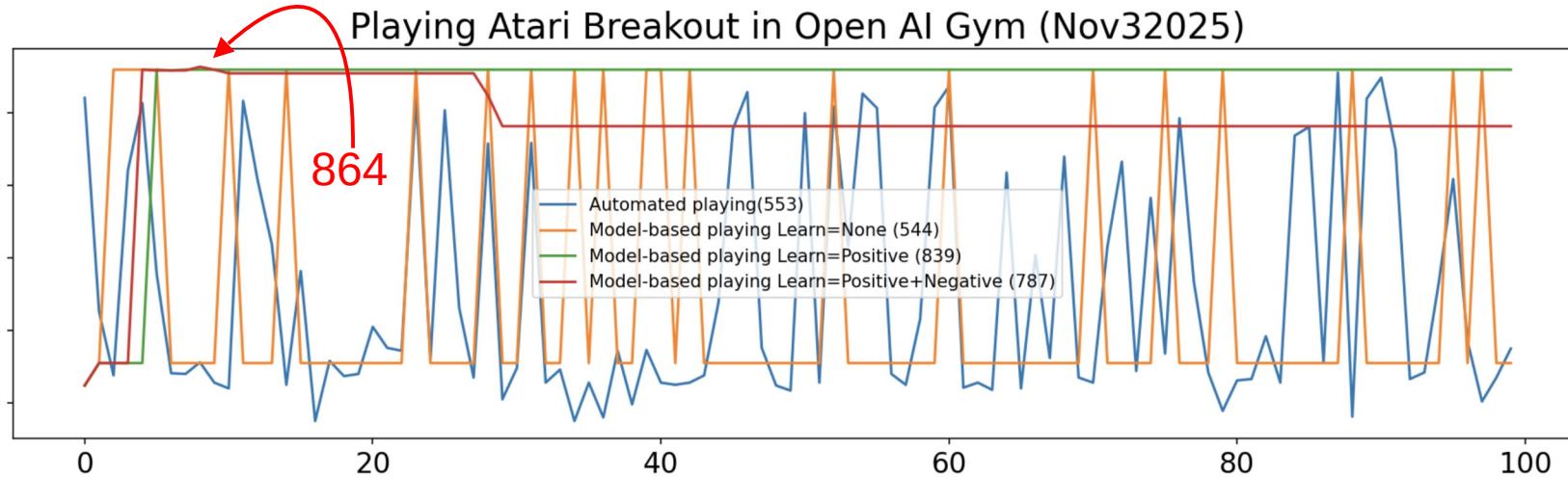
<https://arxiv.org/pdf/2509.07009>

Anton Kolonin, Neuro-Symbolic Architecture for Experiential Learning in Discrete and Functional Environments

[https://doi.org/10.1007/978-3-030-93758-4\\_12](https://doi.org/10.1007/978-3-030-93758-4_12)

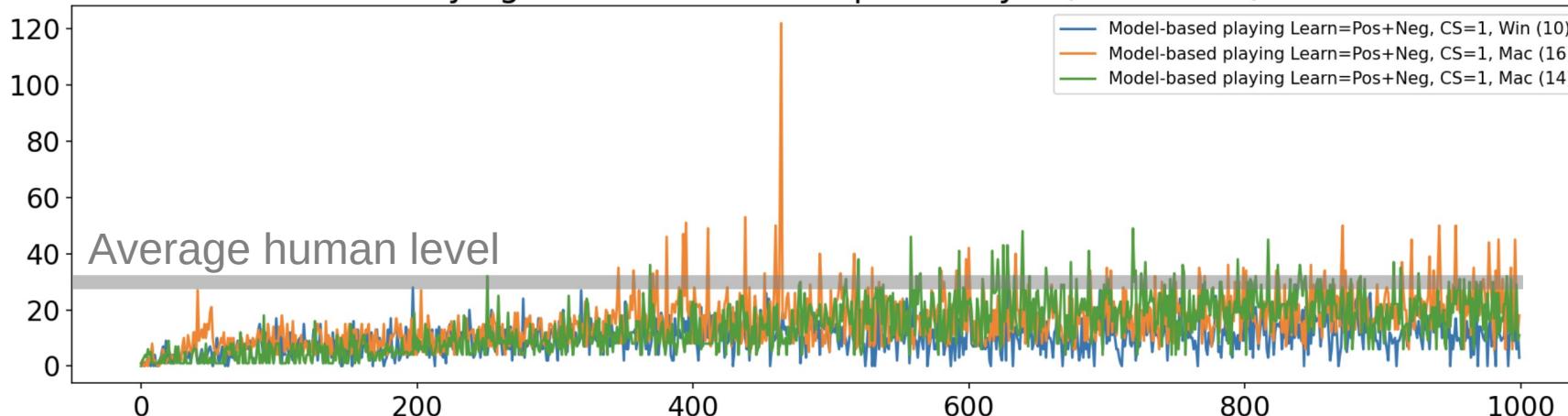
Copyright © 2025 Anton Kolonin

# Imitation learning – decision making on “pre-trained” model

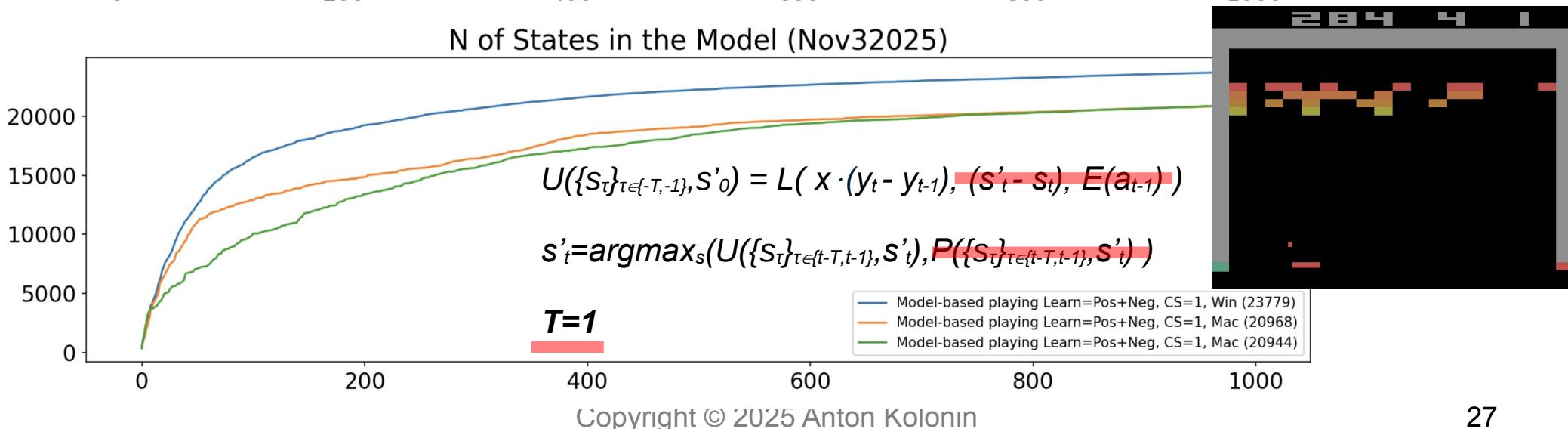


# Reinforcement learning – experiential learning and decision making

Playing Atari Breakout in Open AI Gym (Nov32025)

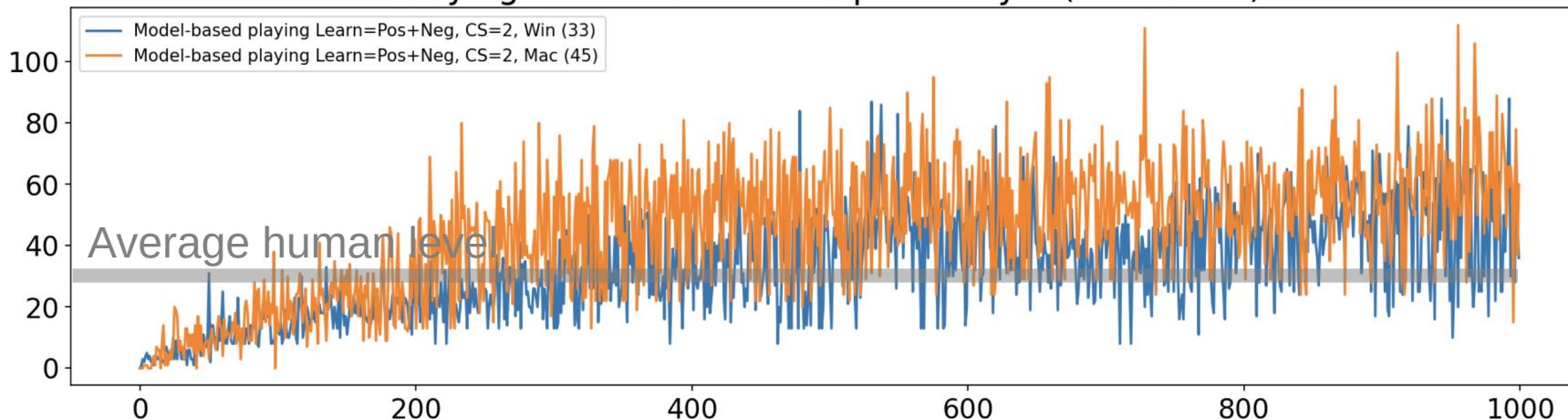


N of States in the Model (Nov32025)

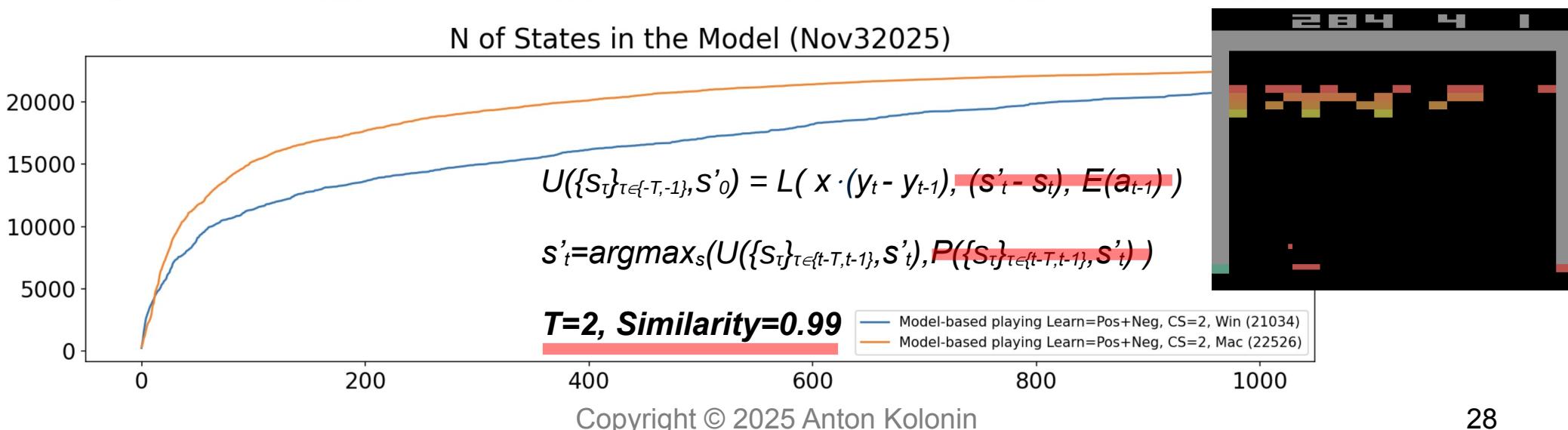


# Reinforcement learning – experiential learning and decision making

Playing Atari Breakout in Open AI Gym (Nov32025)

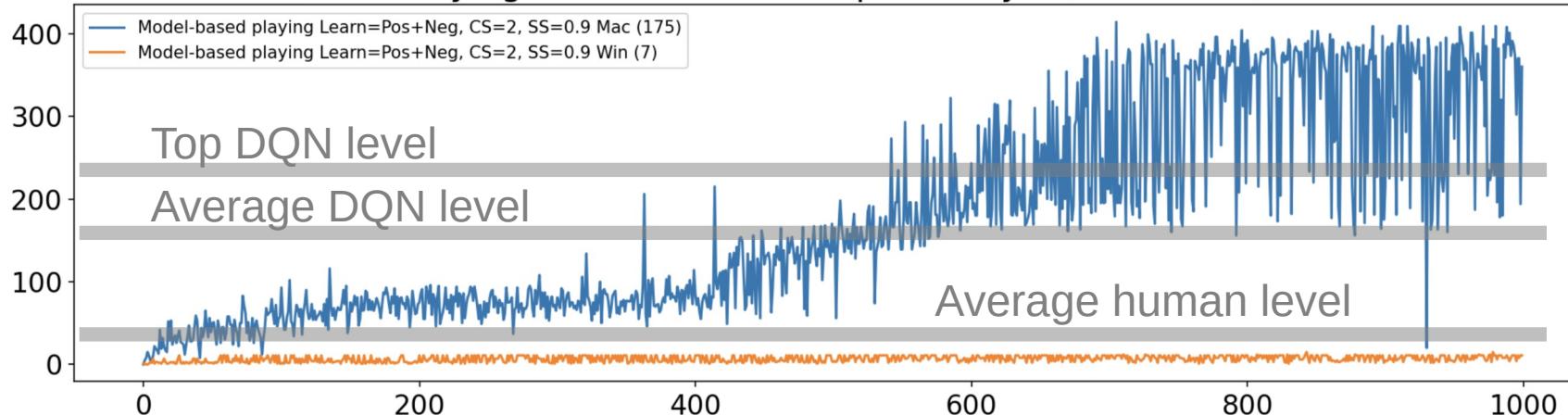


N of States in the Model (Nov32025)

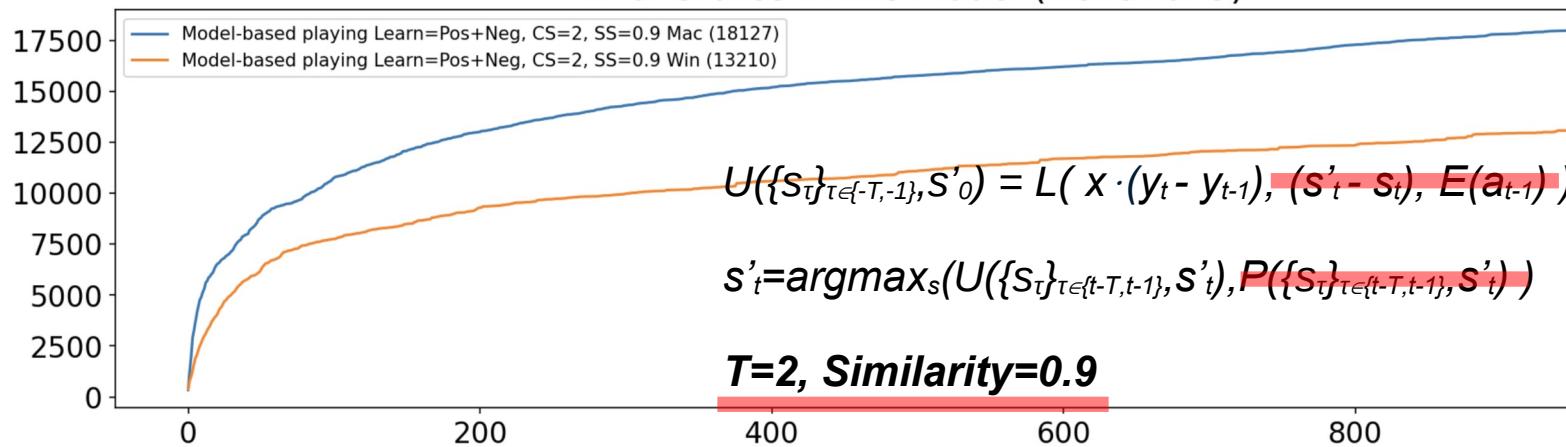


# Reinforcement learning – experiential learning and decision making

Playing Atari Breakout in Open AI Gym (Nov32025)



N of States in the Model (Nov32025)



# Thank you for attention! Questions?

Anton Kolonin

[akolonin@aigents.com](mailto:akolonin@aigents.com)

Telegram/GitHub: akolonin

Workshop recording  
on the subject



Anton Kolonin & Vladimir Kryukov,  
Computational Concept of the  
Psyche, Neuroinformatics-2025

