

Системная модель психики для объяснения естественного и построения искусственного интеллекта

Антон Колонин и Владимир Крюков

akolonin@aigents.com

Telegram: akolonin

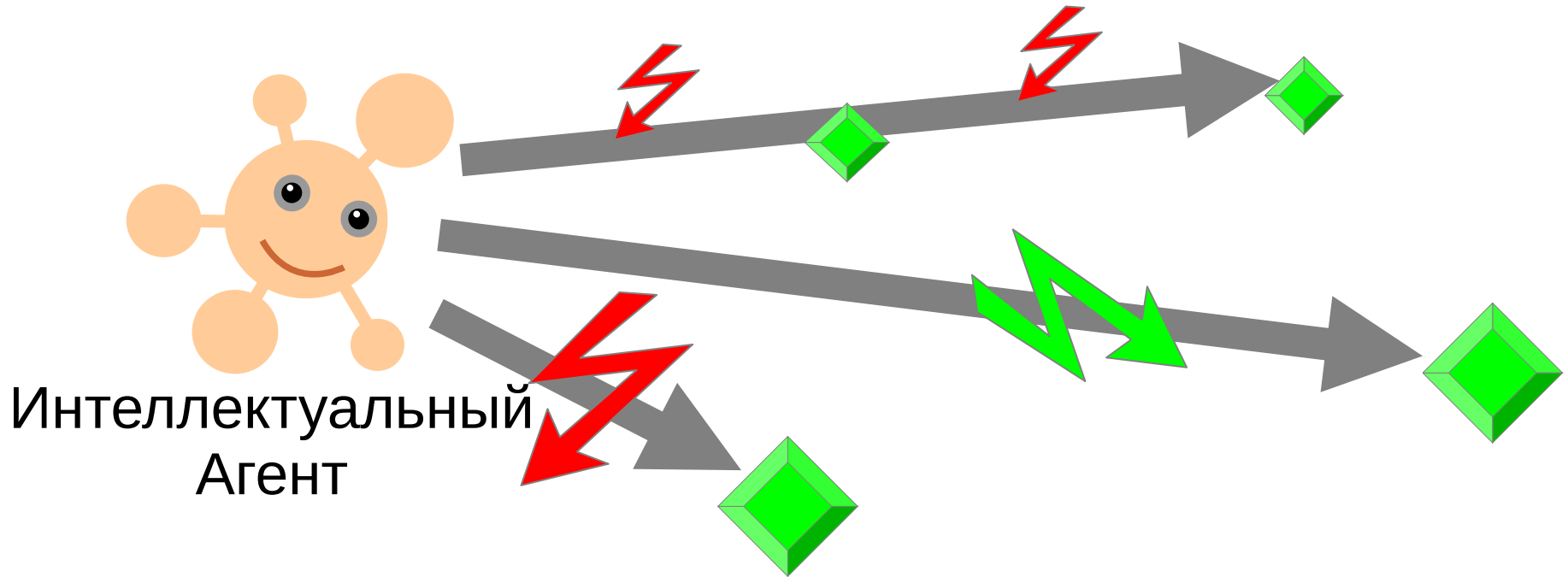
N* Novosibirsk
State
University
*THE REAL SCIENCE
<https://www.nsu.ru>



<https://agirussia.org>

Интеллект:

Достижение сложных **целей** в **различных**
сложных средах, в условиях ограниченных ресурсов
(Ben Goertzel + Pei Wang + **Shane Legg** + **Marcus Hutter**)



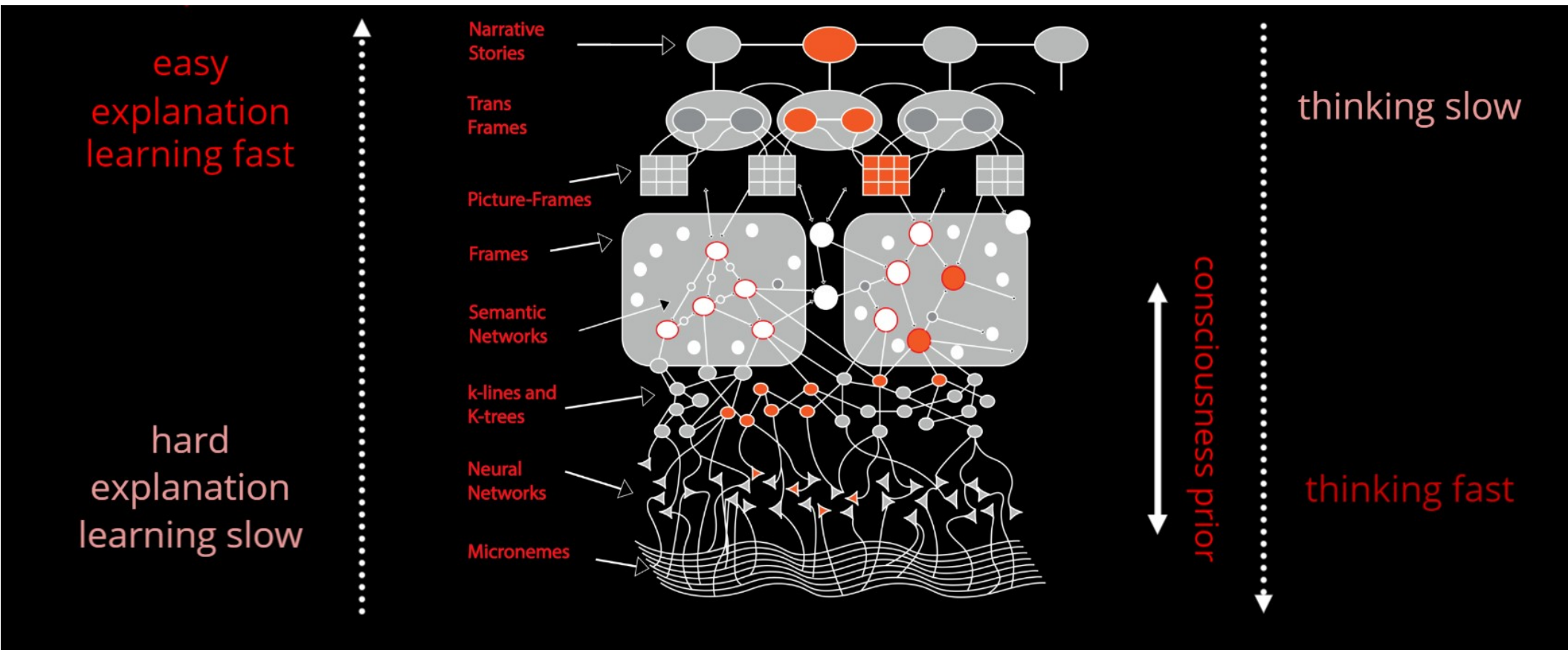
Псѣхика (от др.-греч. ψυχήος «душевнѣй, жизнєнный») — сложное понятие в психологии и медицине, которое, в зависимости от областей знаний и направлений наук, определяется как:

- Совокупность мыслительных процессов и явлений (ощущения, восприятия, эмоции, память и т. п.); специфический аспект жизнедеятельности человека и других животных с нервной системой в их взаимодействии с окружающей средой^{[1][2]}.
- «Форма активного отображения субъектом объективной реальности, возникающая в процессе взаимодействия высокоорганизованных живых существ с внешним миром и осуществляющая в их поведении (деятельности) регулятивную функцию». ^[3]
- Системное свойство высокоорганизованной живой материи, заключающееся в активном отражении субъектом объективного мира, в построении неотчуждаемой от него картины мира и регуляции на этой основе своего поведения и деятельности^[4].
- Внутренний мир человека (предмет психологии)^[5].

Психика животных — субъективный мир животного, охватывающий весь комплекс субъективно переживаемых процессов и состояний: восприятие, память, мышление, намерения, сны и т. п.^[6]

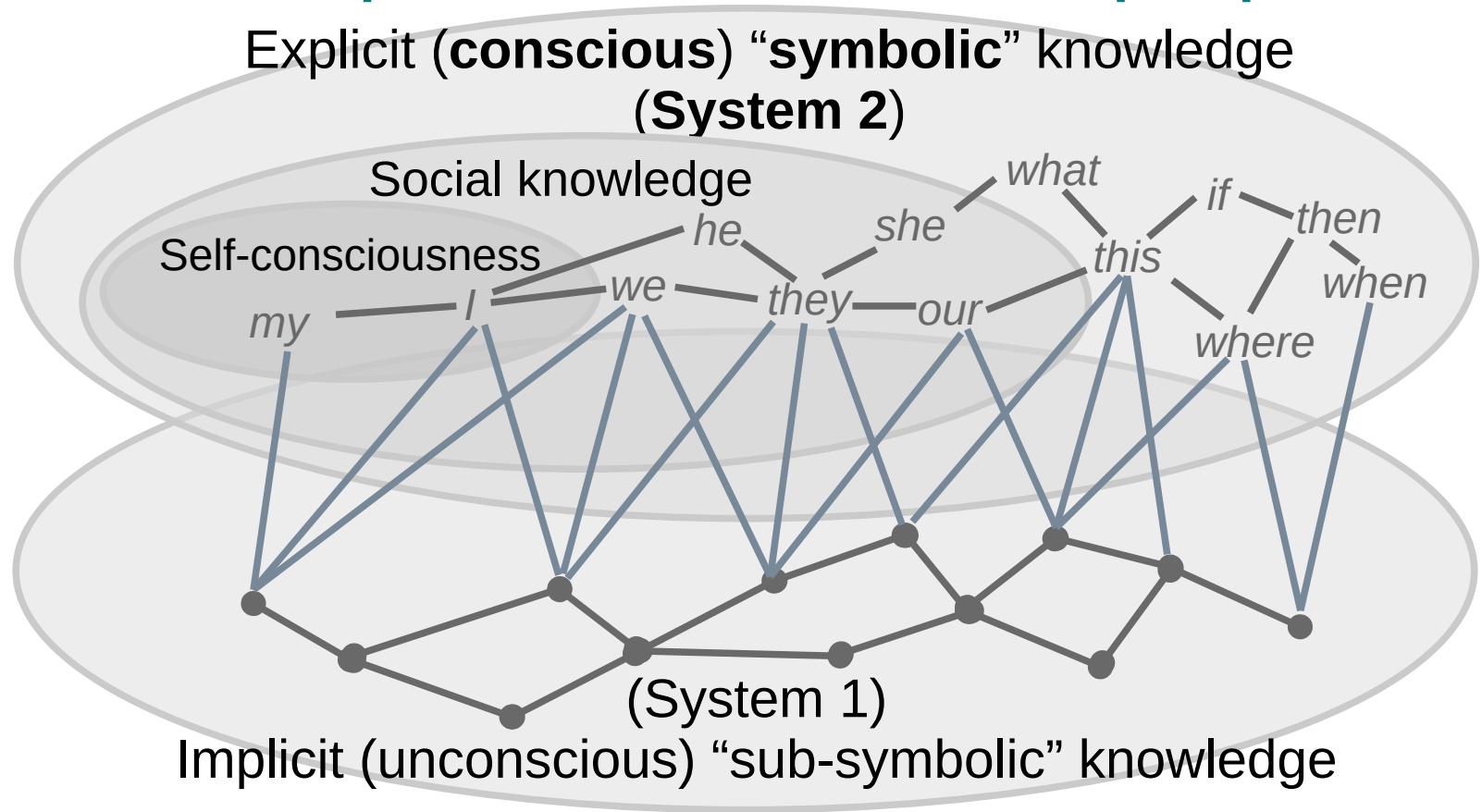
Психика характеризуется такими качествами, как целостность, активность, развитие, саморегуляция, коммуникативность, адаптация и т. д.; связана с соматическими процессами.

“Быстрое и медленное мышление” – Daniel Kahneman



<https://towardsdatascience.com/explainable-ai-vs-explaining-ai-part-1-d39ea5053347>

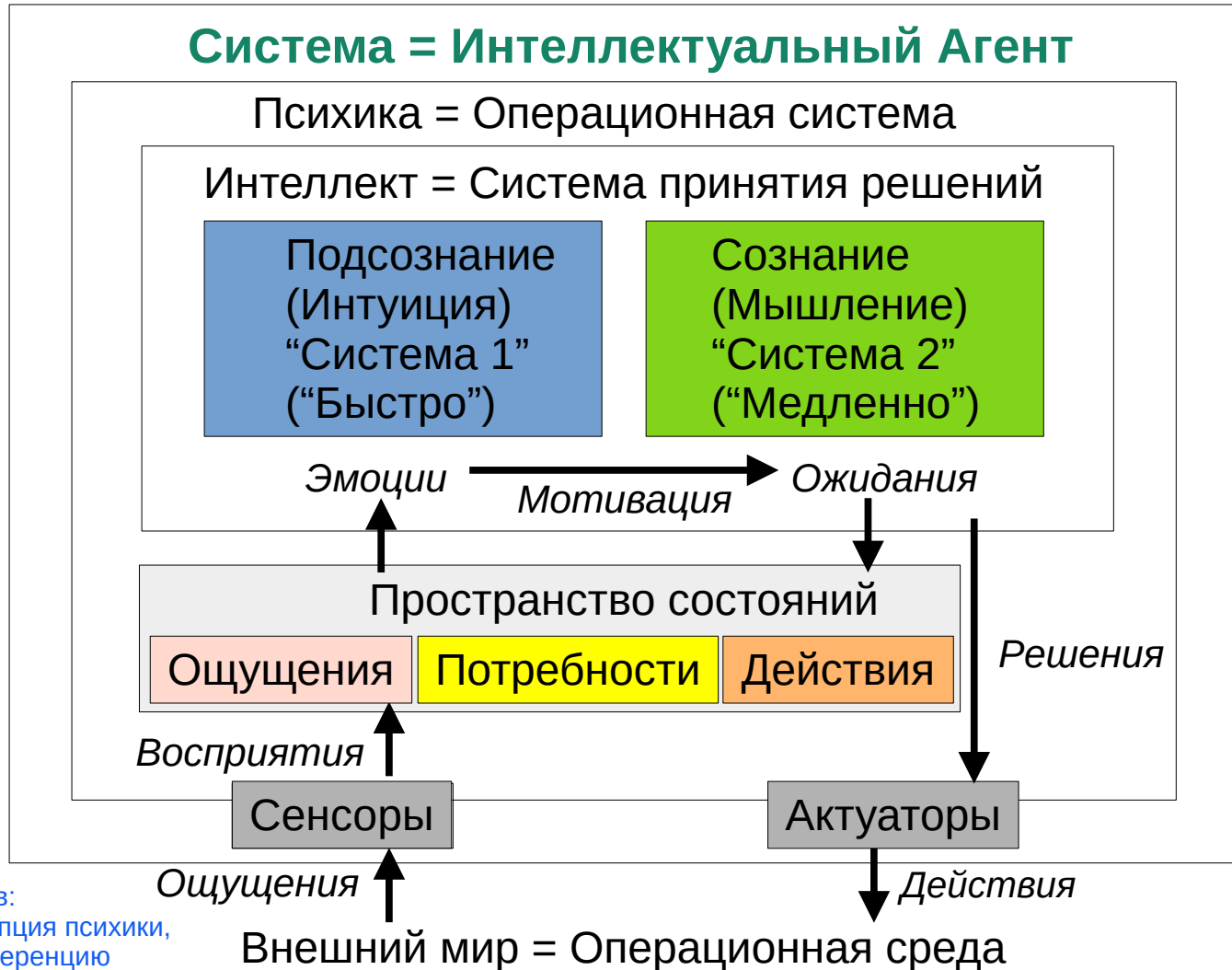
Среда: “нейро-символьный” “граф знаний”



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

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

Система = Интеллектуальный Агент



А.Г.Колонин, В.Г.Крюков:
Вычислительная концепция психики,
Статья подана на конференцию
Нейроинформатика-25

Психика = Операционная система

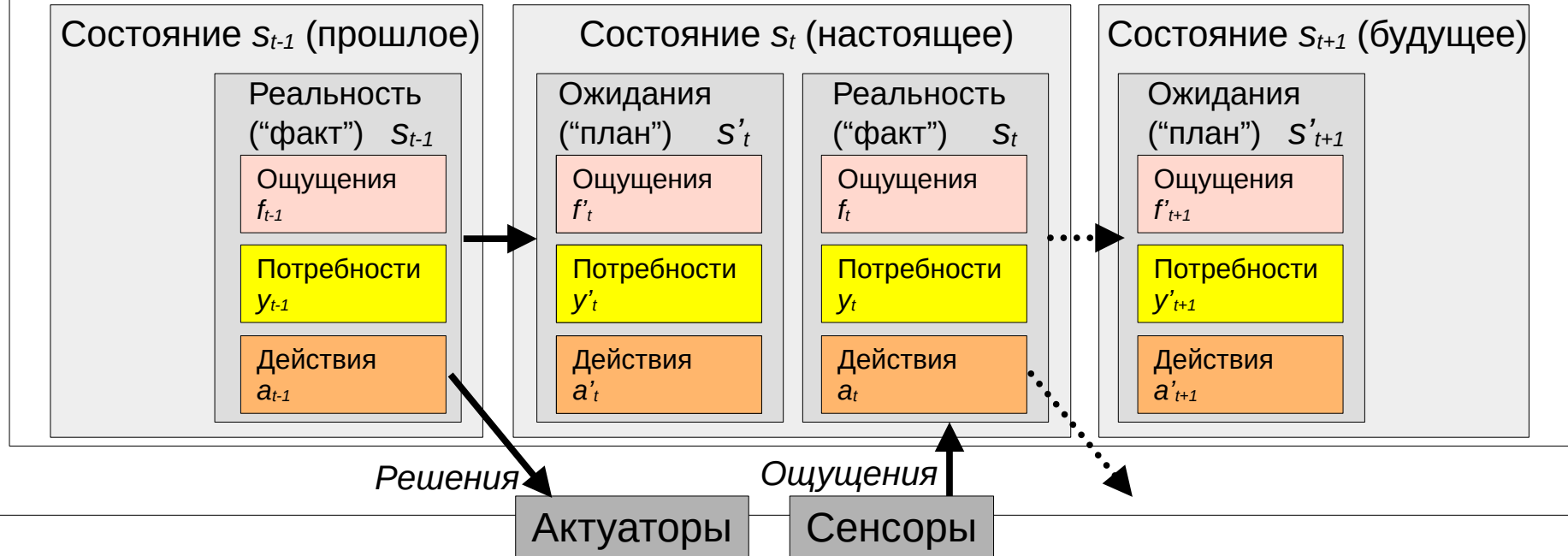
Интеллект = Система принятия решений

Модели s (“инварианты”) состояний, полезности U и вероятности P переходов
 $U(\{s_t\}_{t \in \{-T, -1\}}, s'_0) = L(x \cdot (y_t - y_{t+1}), s'_t - s_t, E(a_t)) \quad s'_t = \operatorname{argmax}_s (U(\{s_t\}_{t \in \{-T, -1\}}, s'_t), P(\{s_t\}_{t \in \{-T, -1\}}, s'_t))$

↑ Обучение на опыте

↓ Принятие решений

Пространство состояний и эпизодическая память (“прецеденты”)



Психика = Операционная система

Интеллект = Система принятия решений

Модели s ("инварианты") состояний, полезности U и вероятности P переходов

$$U(\{s_t\}_{t \in \{-T, -1\}}, s'_0) = L(x \cdot (y_t - y_{t+1}), s'_t - s_t, E(a_t)) \quad s'_t = \operatorname{argmax}_s (U(\{s_t\}_{t \in \{-T, -1\}}, s'_t), P(\{s_t\}_{t \in \{-T, -1\}}, s'_t))$$

↑ Обучение на опыте

↓ Принятие решений

Пространство состояний и эпизодическая память ("прецеденты")

Состояние s_{t-1} (прошое)

Реальность
("факт") s_{t-1}

Ощущения
 f_{t-1}

Потребности
 y_{t-1}

Действия
 a_{t-1}

Состояние s_t (настоящее)

Ожидания
("план") s'_t

Ощущения
 f'_t

Потребности
 y'_t

Действия
 a'_t

Реальность
("факт") s_t

Ощущения
 f_t

Потребности
 y_t

Действия
 a_t

Состояние s_{t+1} (будущее)

Ожидания
("план") s'_{t+1}

Ощущения
 f'_{t+1}

Потребности
 y'_{t+1}

Действия
 a'_{t+1}

Решения

Ощущения

Актуаторы

Сенсоры

Оптимальное решение
и ожидание s

Ожидаемая
полезность s

Ожидаемая
вероятность s

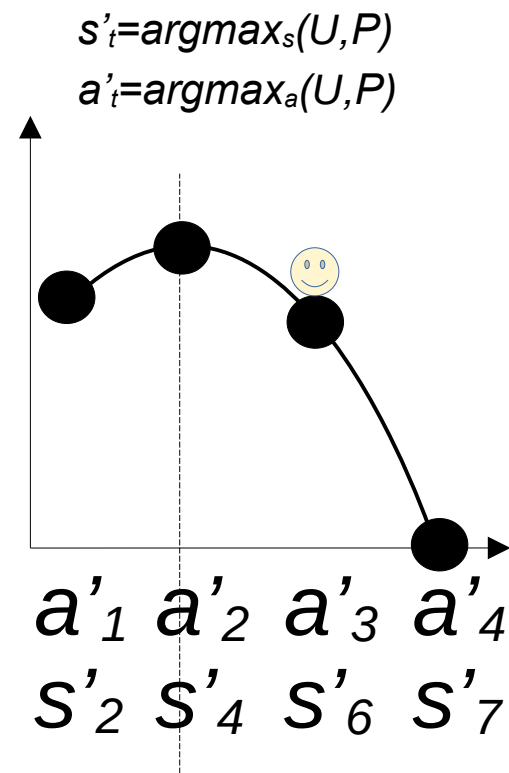
Удовлетворенность
 y


Предсказуемость
 s

Энергоэффективность
 a

Принятие решений как управление операционными рисками

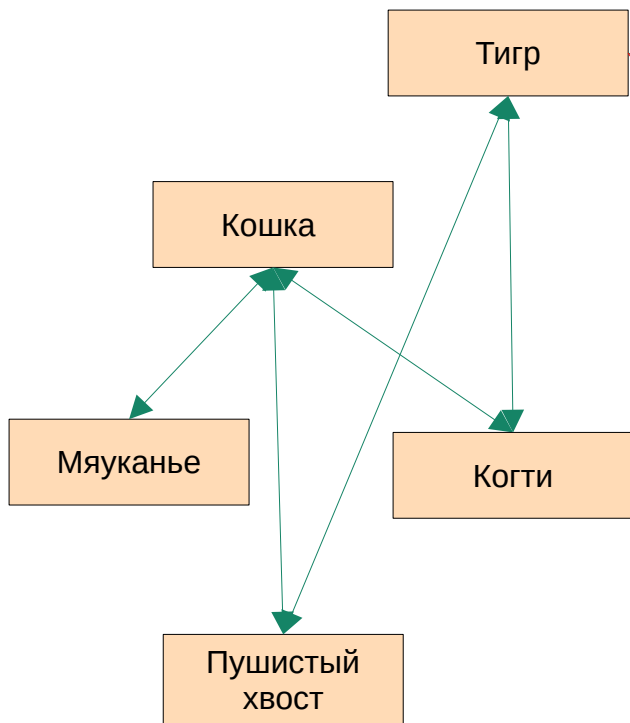
s_t	s'_{t+1}	s'_{t+1}			U	P	$\Sigma U*P$
		a'	y'	f'			
s_1	s'_2	a'_1	y'_1	...	1.0	0.5	<u>0.7</u>
s_1	s'_3	a'_1	y'_2	...	0.4	0.5	
s_1	s'_4	a'_2	y'_3	...	1.0	0.8	<u>0.8</u>
s_1	s'_5	a'_2	y'_4	...	0.0	0.2	
s_1	s'_6	a'_3	y'_5	...	0.6	1.0	<u>0.6</u>
s_1	s'_7	a'_4	y'_6	...	0.0	1.0	<u>0.0</u>



 По Тверски и Канеману, большинство людей выбирает a'_3 и s'_6 (“синицу в руке”)

Взаимосвязность переменных состояния

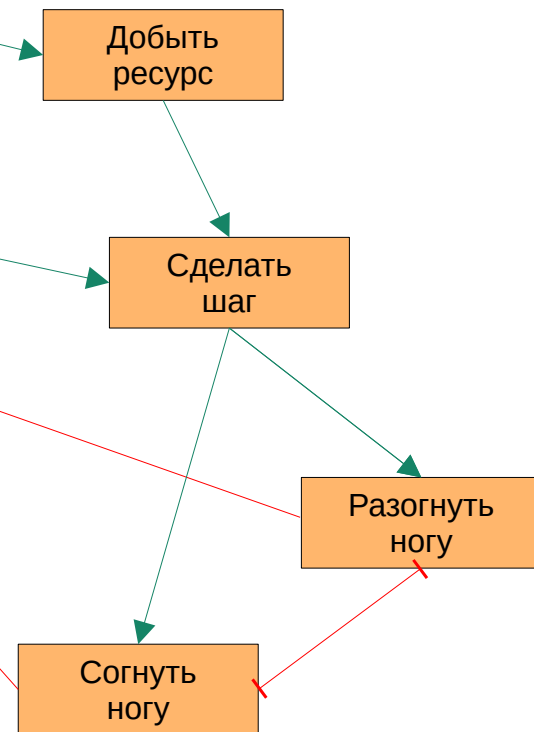
Переживаемые ощущения



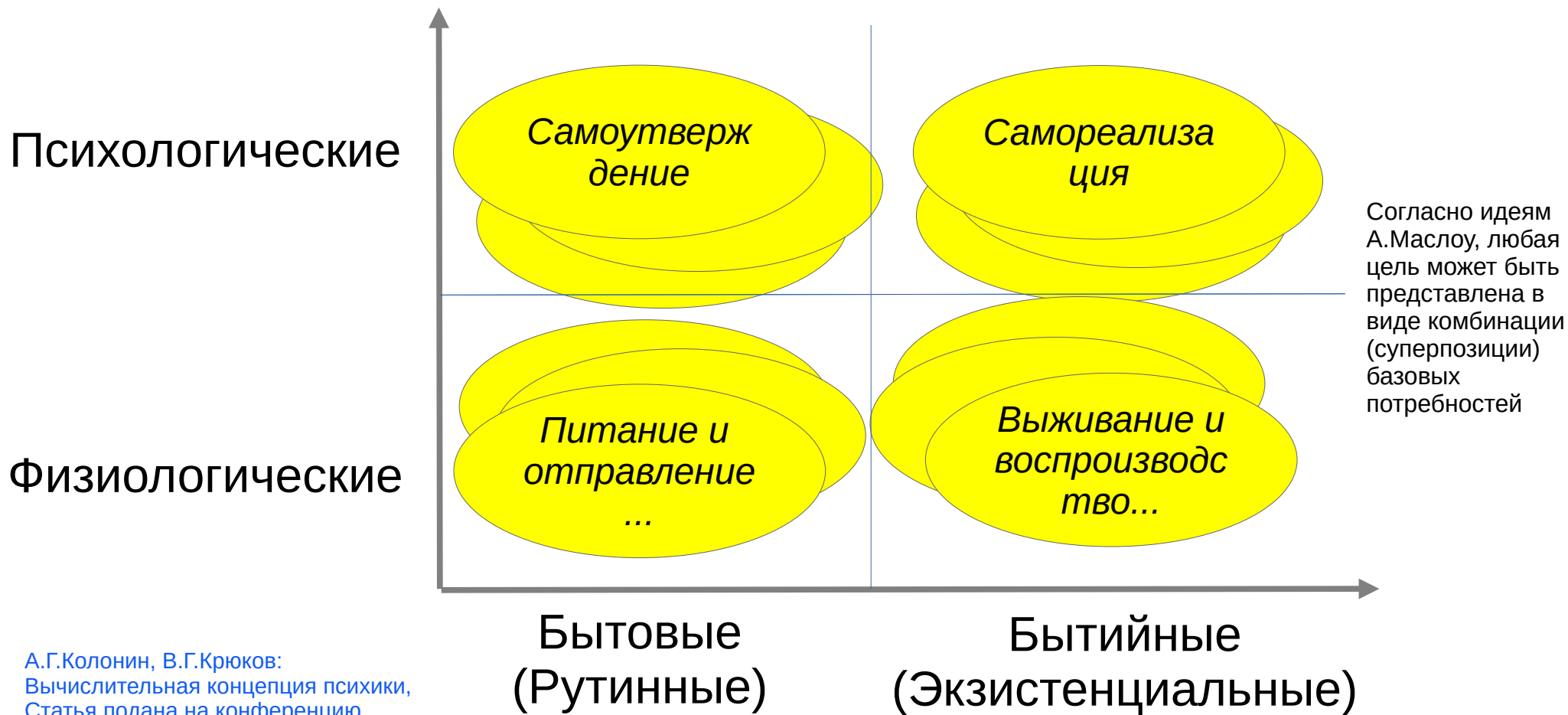
Удовлетворяемые потребности



Совершаемые действия



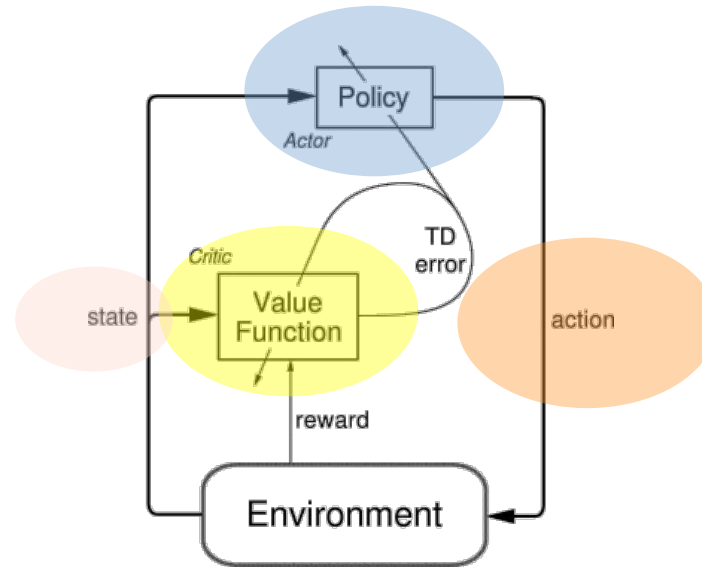
Пространство потребностей/целей/ценностей



А.Г.Колонин, В.Г.Крюков:
Вычислительная концепция психики,
Статья подана на конференцию
Нейроинформатика-25

Варианты реализации

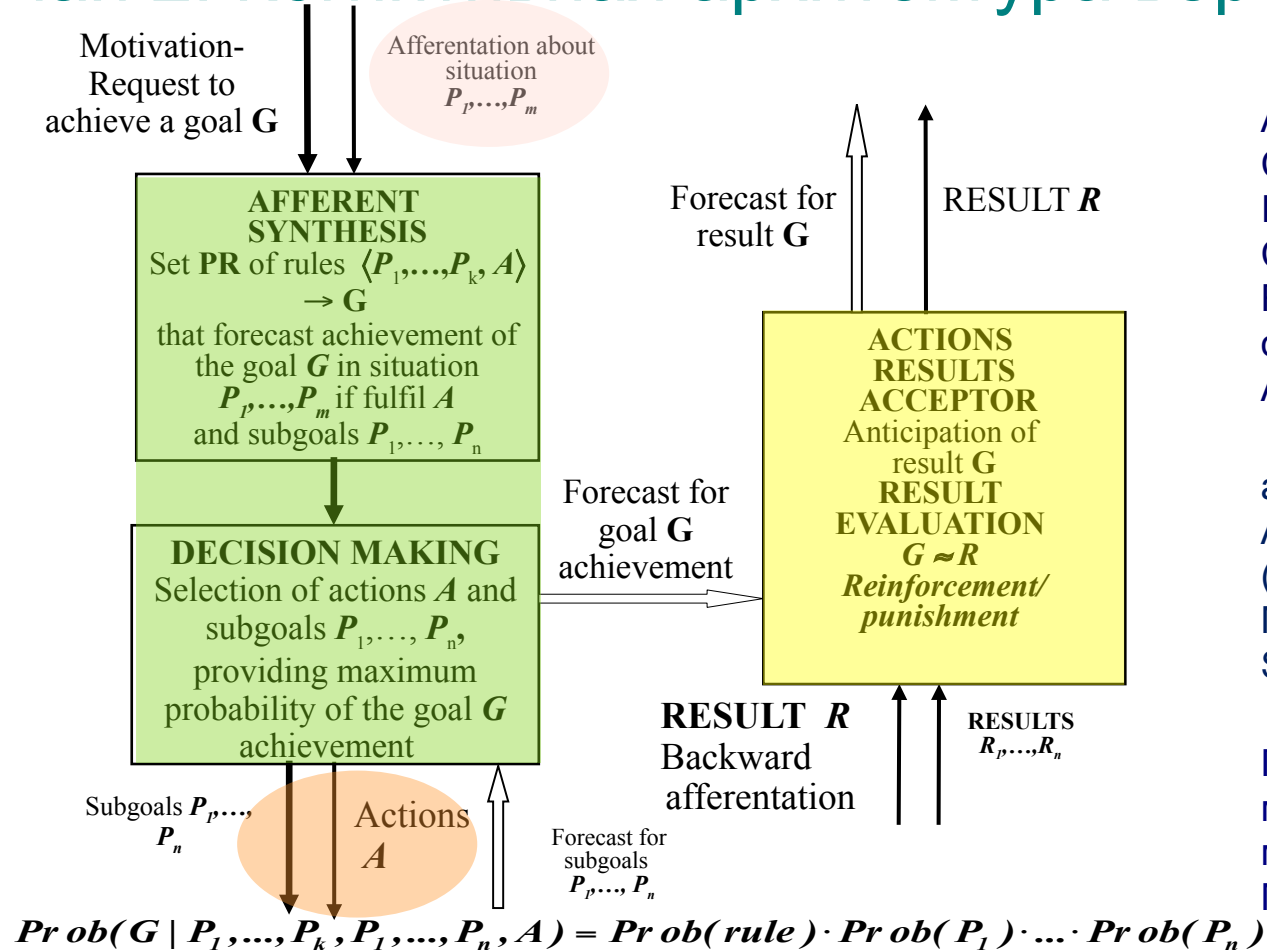
Случай 1: Обучение с подкреплением “актор-критик”



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

Варианты реализации

Случай 2: Когнитивная архитектура вероятностной логики



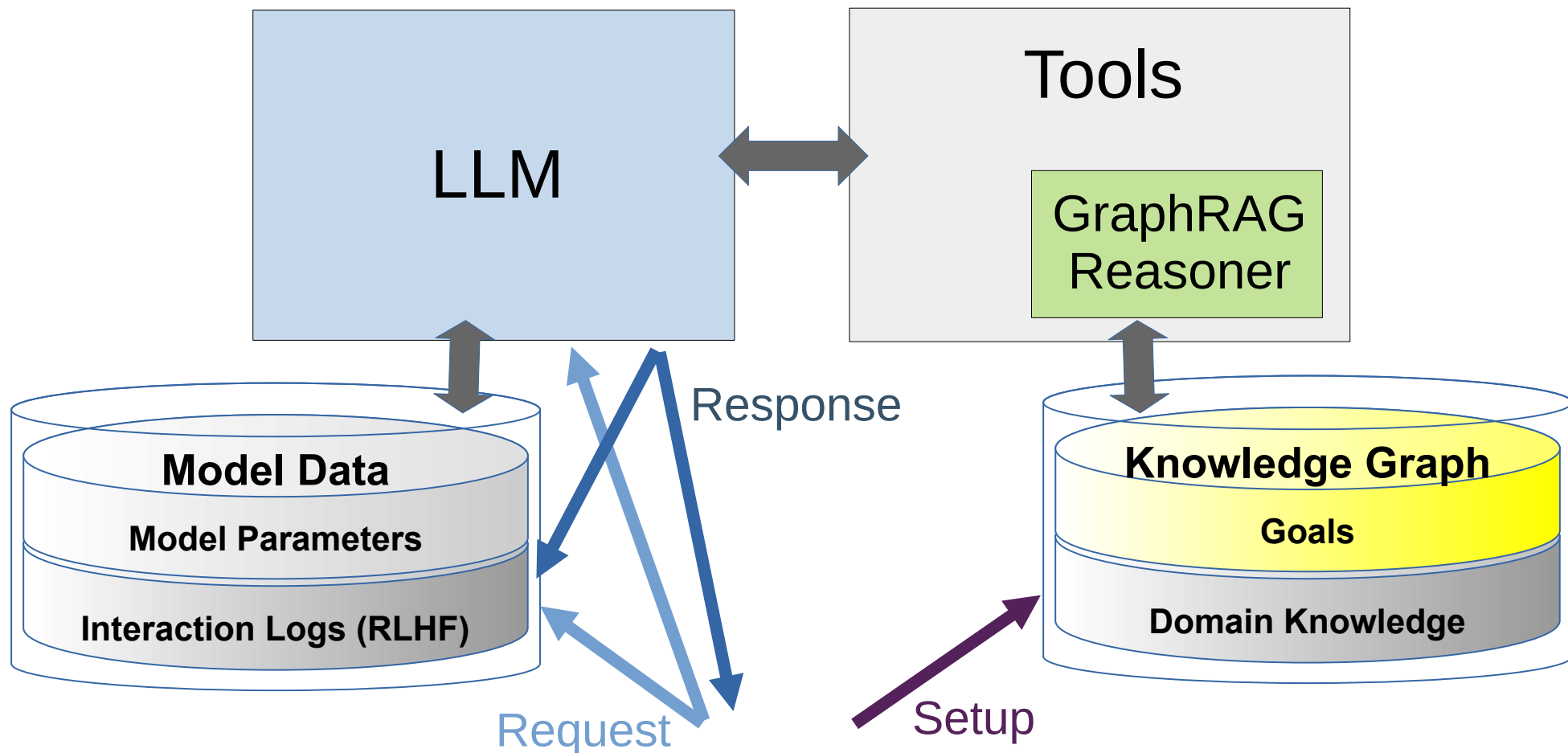
Evgenii Vityaev, Alexander Demin:
Adaptive Control of Modular Robots //
Conference Paper in Advances in
Intelligent Systems and Computing,
Conference: First International Early
Research Career Enhancement School
on Biologically Inspired Cognitive
Architectures, Springer, August 2018

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.

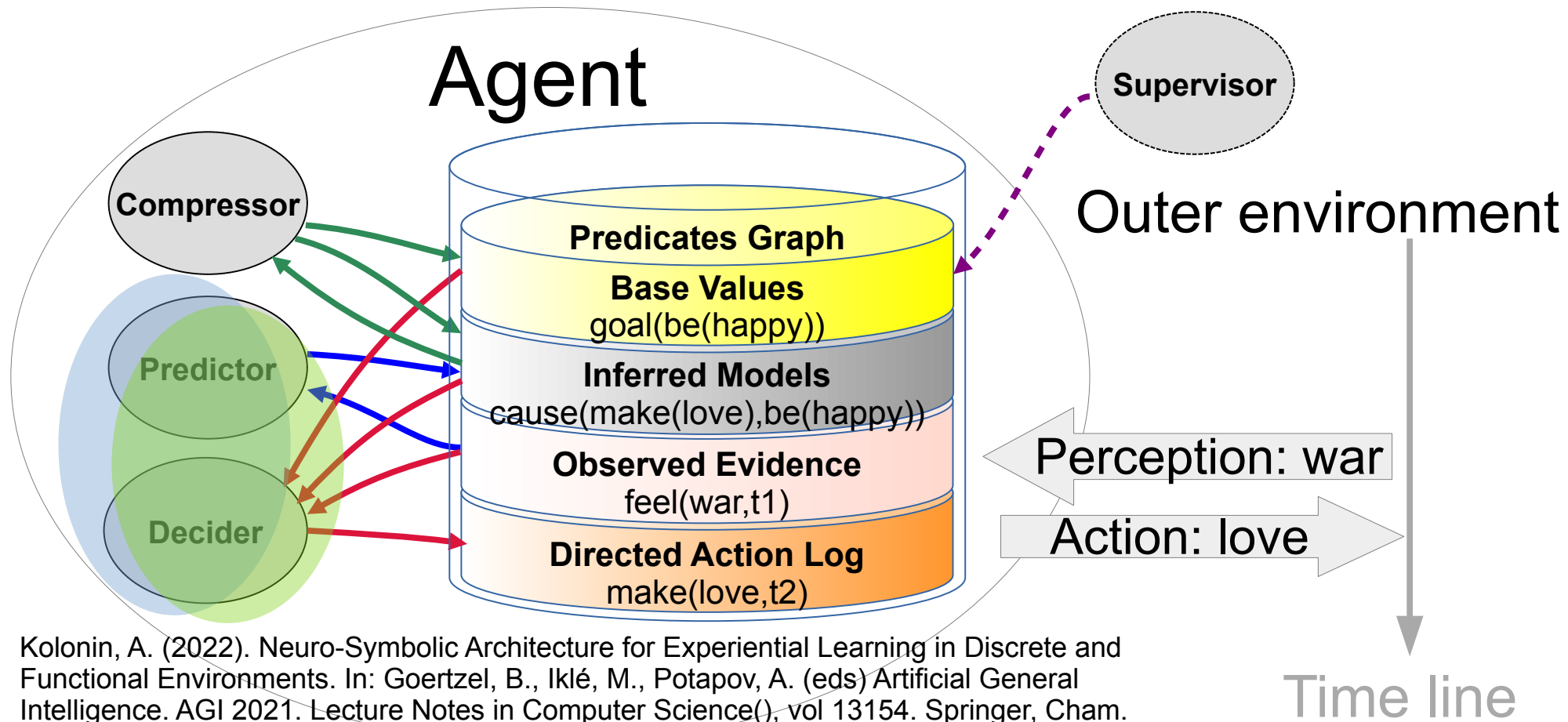
Варианты реализации

Случай 3: Когнитивная архитектура на основе LLM и GraphRAG



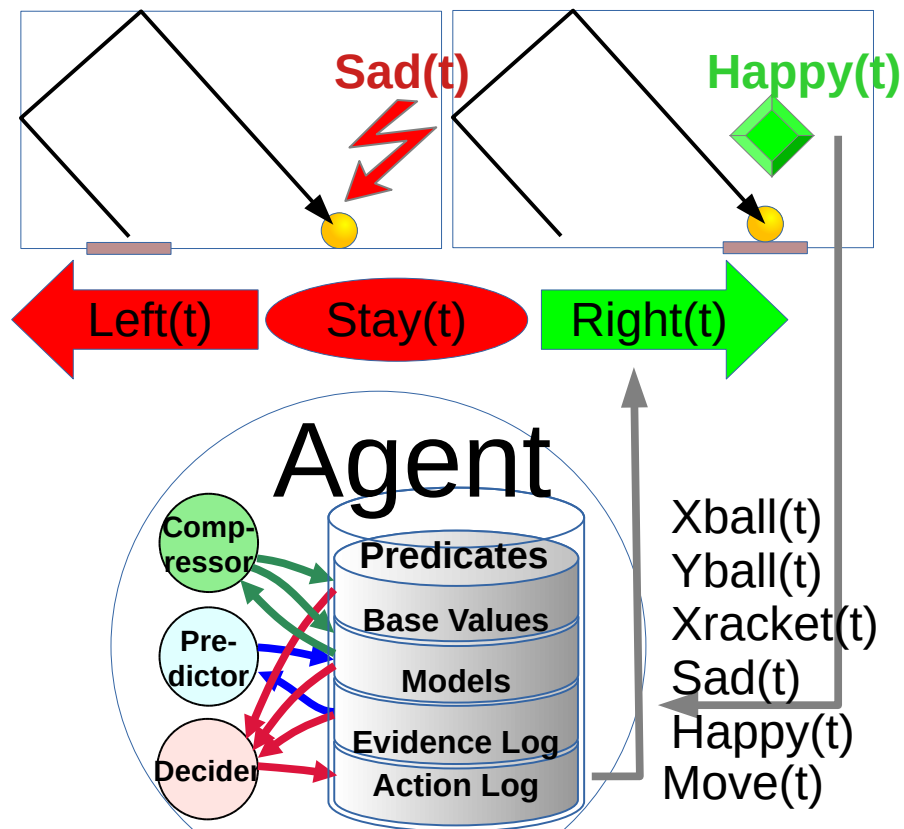
Варианты реализации

Случай 4: Когнитивная архитектура обучения на основе ценностей и опыта

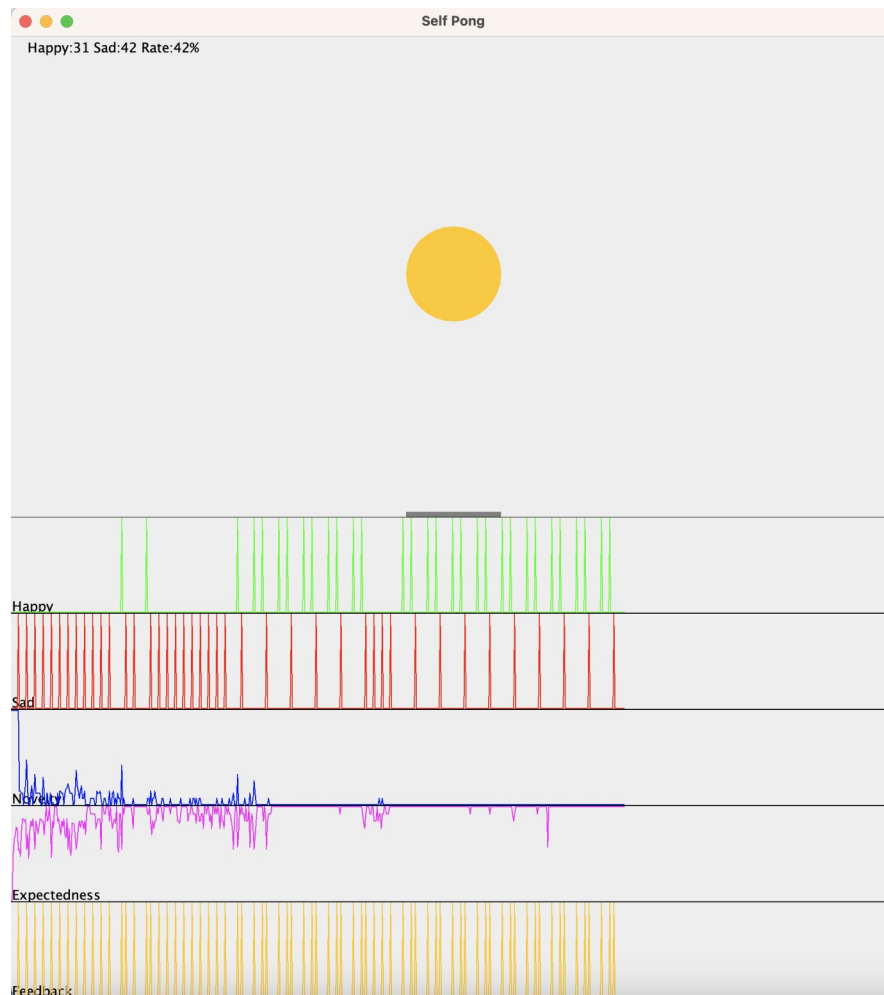


Kolonin, A. (2022). Neuro-Symbolic Architecture for Experiential Learning in Discrete and Functional Environments. In: Goertzel, B., Iklé, M., Potapov, A. (eds) Artificial General Intelligence. AGI 2021. Lecture Notes in Computer Science(), vol 13154. Springer, Cham.
https://doi.org/10.1007/978-3-030-93758-4_12

Когнитивная архитектура обучения на основе ценностей и опыта



А.Г.Колонин, В.Г.Крюков:
Вычислительная концепция психики,
Статья подана на конференцию
Нейроинформатика-25



Спасибо за внимание!

Вопросы?

Антон Колонин и Владимир Крюков

akolonin@aigents.com

Telegram: akolonin

N* Novosibirsk
State
University
*THE REAL SCIENCE
<https://www.nsu.ru>



<https://agirussia.org>