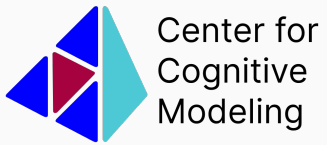
Распределенные вероятностные модели хеббовской временной памяти

Прохоренко Артём
prokhorenko.aiu@phystech.edu

Научно-исследовательский семинар
March 25, 2025



Motivation

- *Online sequence learning*: критичность в real-time адаптации модели
- Ускорение формирования моделей внешнего мира агентами
- Традиционные модели памяти чувствительны к шумным, динамическим средам
- Способность эффективно работать в условиях частичной наблюдаемости (POMDP)

Problem Statement

- **Среда:**
  - S — множество состояний среды.
  - A — множество действий агента.
  - $T = P(s_{t+1}|s_t, a_t)$ — функция динамики среды, определяющая вероятность перехода из состояния s_t в s_{t+1} при выполнении действия a_t.
- **Задача:**
  - Построение модели мира для сред (в т.ч. PO сред) с использованием моделей DHTM (Distributed Hebbian Temporal Memory) и tPC (Temporal Predictive Coding).
  - Предсказание вероятных будущих состояний в реальном времени на основе истории наблюдений.
  - Оптимизация модели для захвата как **явных**, так и **скрытых закономерностей** в последовательных данных.

Problem Statement

- **Метод для оптимизации DHTM: Belief Propagation**
  - Для факторного графа $G = (V, F, E)$
  - Задача: $p(x_i)$ для каждой скрытой переменной x_i.
  - Метод: итеративное обновление вероятностей с помощью передачи «сообщений» между узлами X_i и X_j.

$$m_{i \to j}(x_j) = \sum_{x_i} \psi_{ij}(x_i, x_j) \prod_{k \neq j} m_{k \to i}(x_i)$$

  где $\psi_{ij}(x_i, x_j)$ — функция взаимосвязи между узлами, $m_{k \to i}(x_i)$— «сообщение» между узлами.

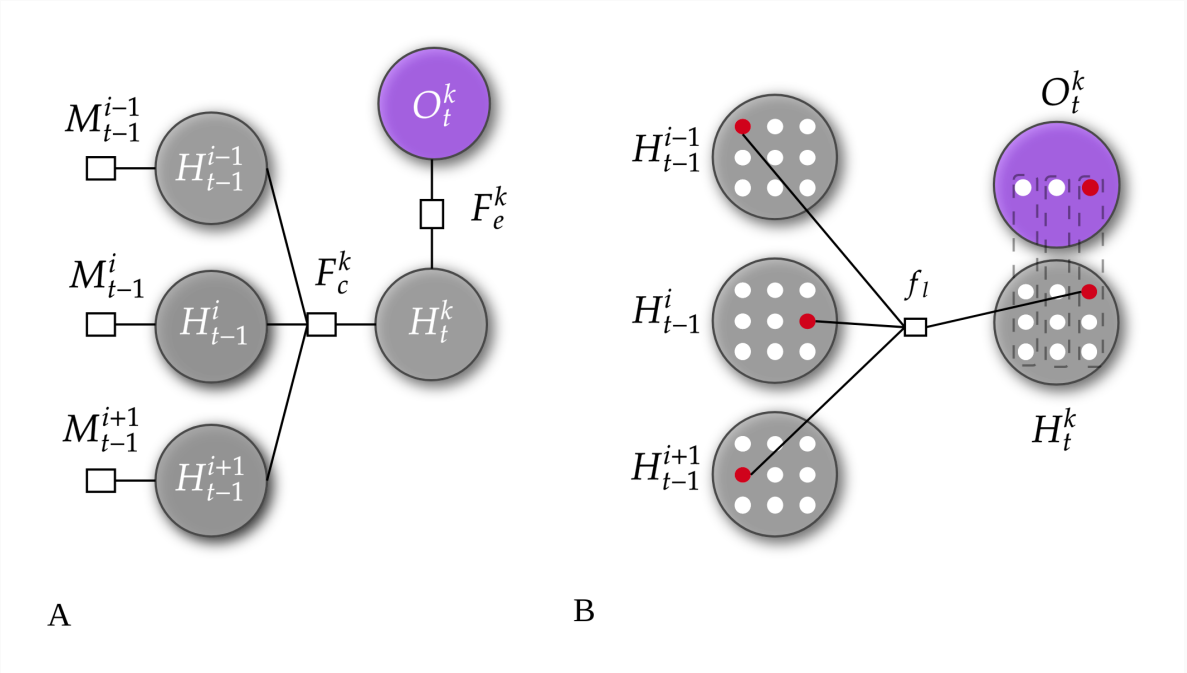
- **Метод для tPC: Minimization of Prediction Error**

$$F^{\mu}(W) = \left\| x^{\mu} - Wf(x^{\mu-1}) \right\|^2 \tag{1}$$

  - Update W:

$$\Delta W \propto -\frac{\partial F^{\mu}(W)}{\partial W} = \varepsilon^{\mu} f(x^{\mu-1})^{\top}$$

Models: Distributed Hebbian Temporal Memory



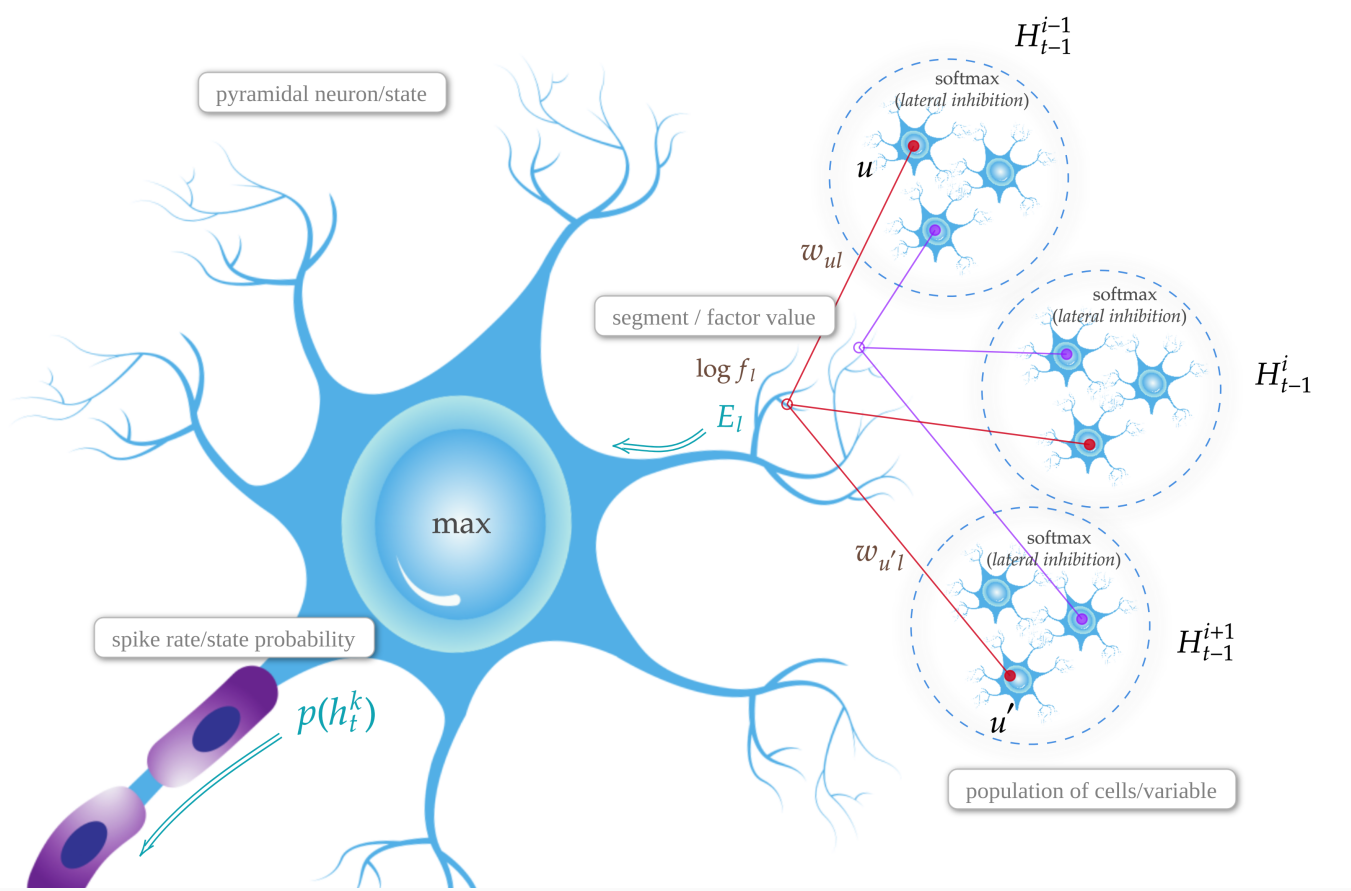
A. Belief propagation:

$$p(h^k_t) \propto \sum_{\{h^i_{t-1}: i \in \Omega_k\}} \prod_{i \in \Omega_k} M^i_{t-1}(h^i_{t-1}) F^k_c(h^k_t, \{h^i_{t-1} : i \in \Omega_k\})$$

$$p(h^k_t \mid o^k_t) \propto p(h^k_t) \mathbb{I}[h^k_t \in \mathrm{col}(o^k_t)]$$

B. Monte-Carlo обучение: $M^i_{t-1} : i \in \Omega_k$ и $p(h^k_t \mid o^k_t)$

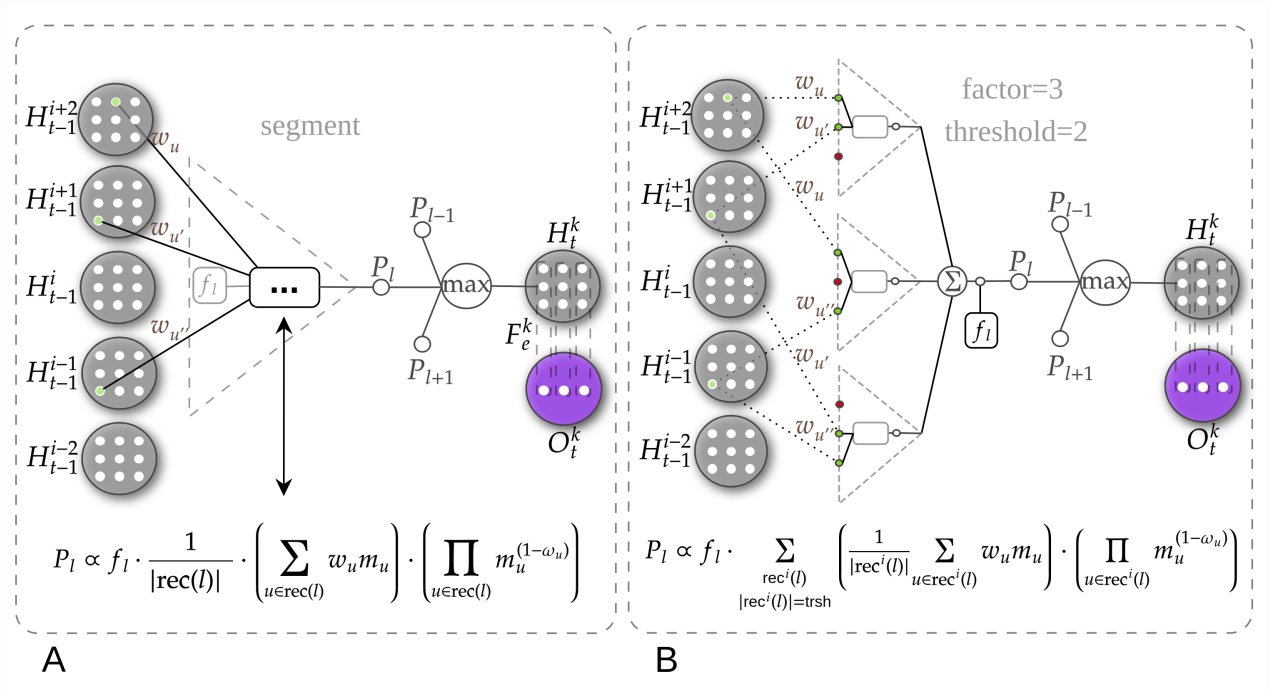
Models: Distributed Hebbian Temporal Memory



Фактор-функция $f_l \leftrightarrow$ дендритный сегмент пирамидного нейрона.

Процесс обучения заключается в выращивании дендритных сегментов и коррекции весов синапсов.

DHTM improvements: soft segment & threshold adaptability

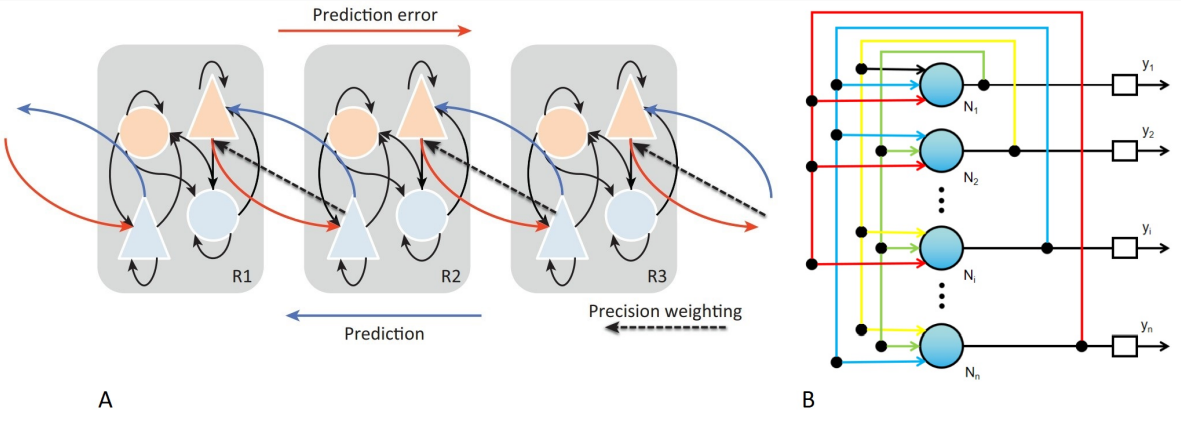


(A) Original DHTM. (B) DHTM with soft segments[1].

- Меньшее значение фактора ⇒ более высокий порог
  - Значение фактора ↔ частота активации

[1] Prokhorenko A. et al. "Soft Adaptive Segments for Bio-Inspired Temporal Memory". In: *Hybrid Artificial Intelligent Systems*. 2025.

Models: Temporal Predictive Coding

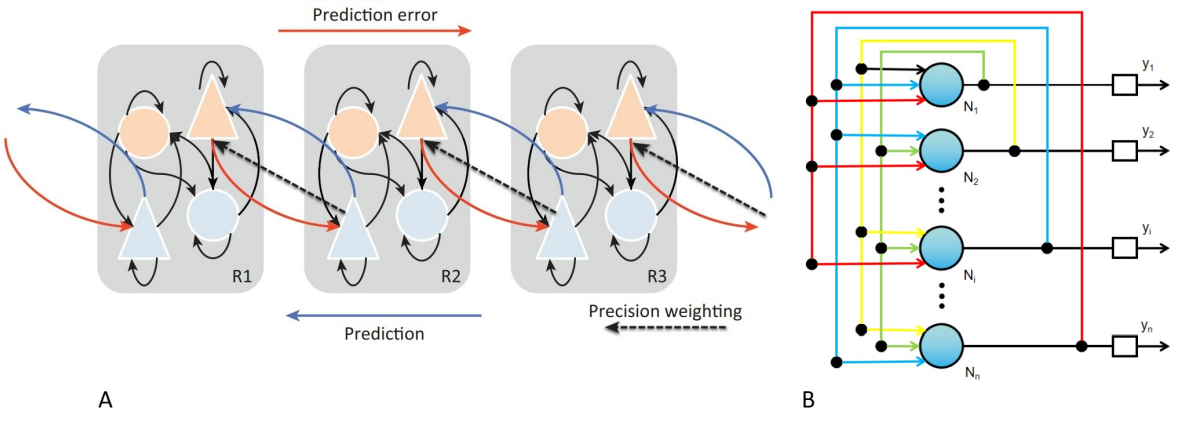


(A) Hierarchical predictive coding. (B) Hopfield network.

Predictive Coding:

- Иерархическая структура.
- Запомненные шаблоны — аттракторы, аналогично сетям Хопфилда.
- Продолжительное запоминание и восстановление с помощью байесовских обновлений.

Models: Temporal Predictive Coding



(A) Hierarchical predictive coding. (B) Hopfield network.

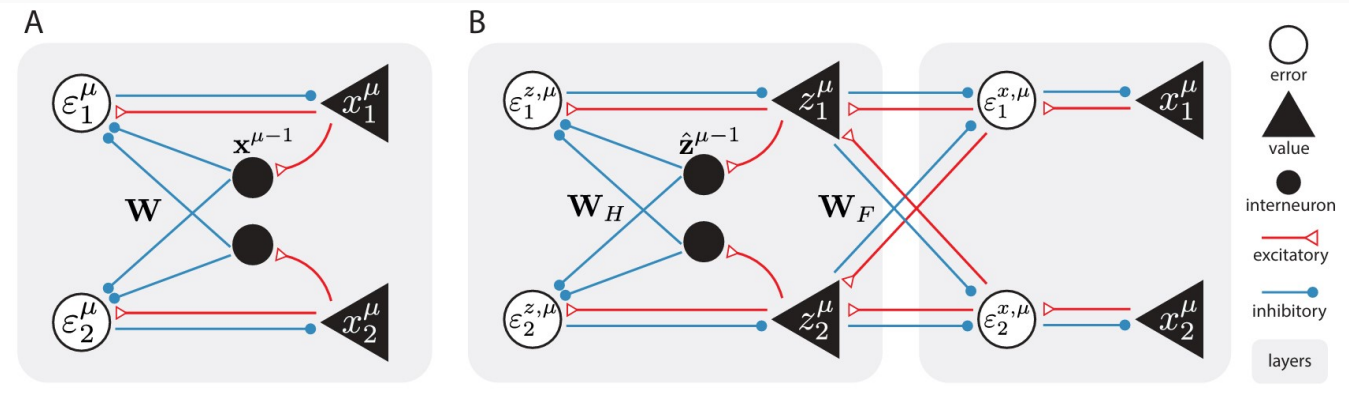
Modern Continuous AHN:

$$W_{AHN} = \sum_{\mu=1}^{P} x^{\mu+1}(x^{\mu})^{\top} \tag{2}$$

$$R_{AHN}(q) = \operatorname{sgn}\left(W_{AHN} q\right) \tag{3}$$

$$R_{MCAHN}(q, \beta) = \sum_{\mu=1}^{P} x^{\mu+1} \operatorname{softmax}\left(\beta (x^{\mu})^{\top} q\right) \tag{4}$$

Models: Temporal Predictive Coding



A: single-layer tPC. B: 2-layer tPC

Single-layer tPC:

- Loss & update:

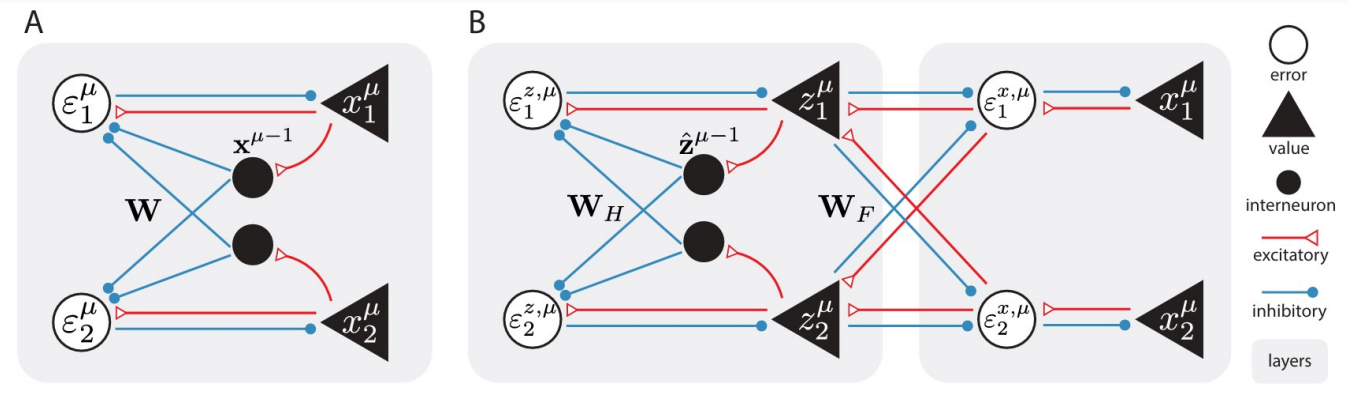
$$F^{\mu}(W) = \|x^{\mu} - Wf(x^{\mu-1})\|^2 \tag{5}$$

$$\Delta W \propto \varepsilon^{\mu} f(x^{\mu-1})^{\top}, \quad \varepsilon^{\mu} = x^{\mu} - Wf(x^{\mu-1}) \tag{6}$$

- Recall:

$$F^{\mu}(\hat{x}^{\mu}) = \|\hat{x}^{\mu} - Wf(q)\|^2, \quad \dot{\hat{x}}^{\mu} \propto -\varepsilon^{\mu} \tag{7}$$

Models: Temporal Predictive Coding



A: single-layer tPC. B: 2-layer tPC

2-layer tPC:

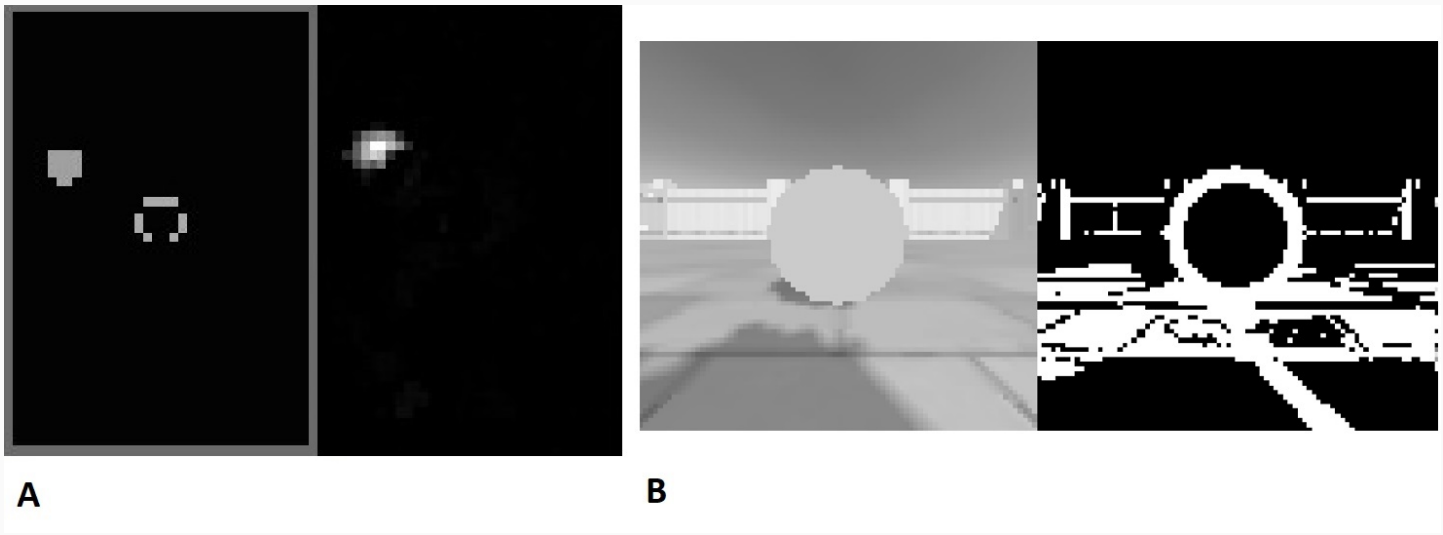
$$F^{\mu} = \|z^{\mu} - W_H f(\hat{z}^{\mu-1})\|^2 + \|x^{\mu} - W_F f(z^{\mu})\|^2 \tag{8}$$

$$\Delta W_H \propto \varepsilon_z f(\hat{z}^{\mu-1})^{\top}, \quad \Delta W_F \propto \varepsilon_x f(z^{\mu})^{\top} \tag{9}$$

Recall:

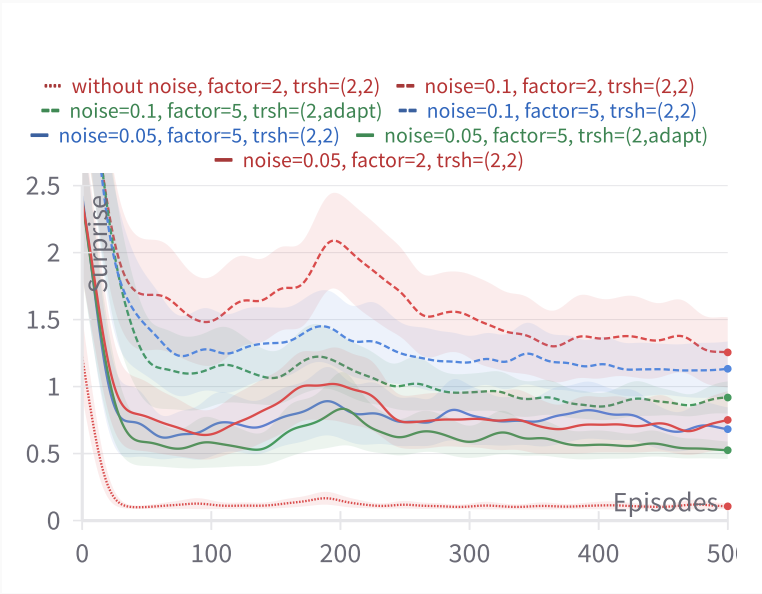
$$\dot{\hat{x}}^{\mu} \propto -\varepsilon_x, \quad \varepsilon_x = \hat{x}^{\mu} - W_F f(z^{\mu}) \tag{10}$$

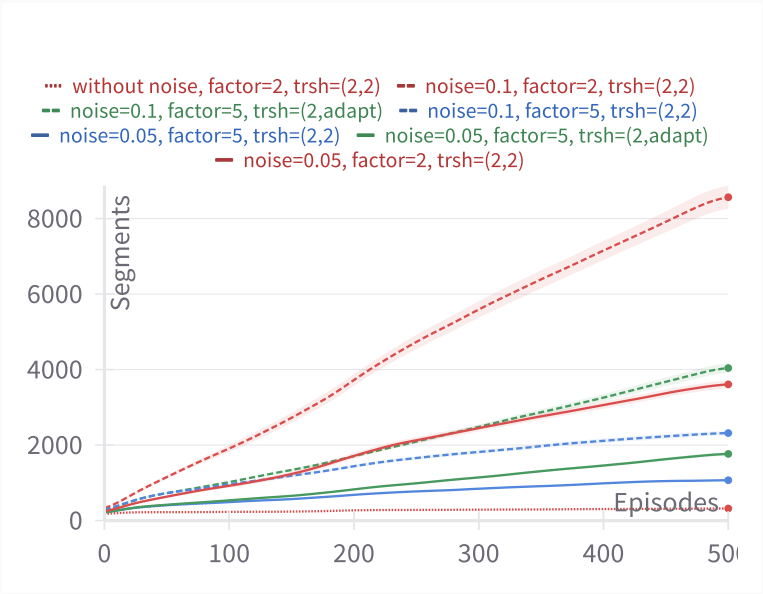
DHTM: Experiments setup

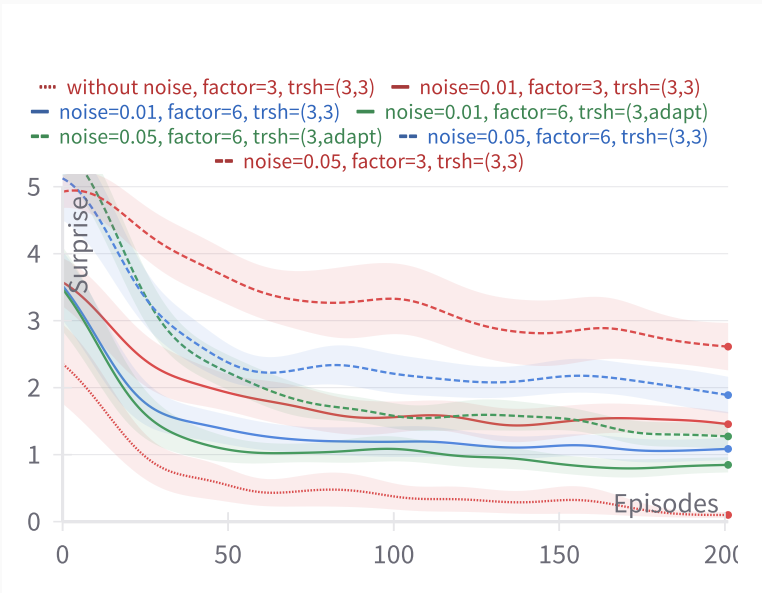


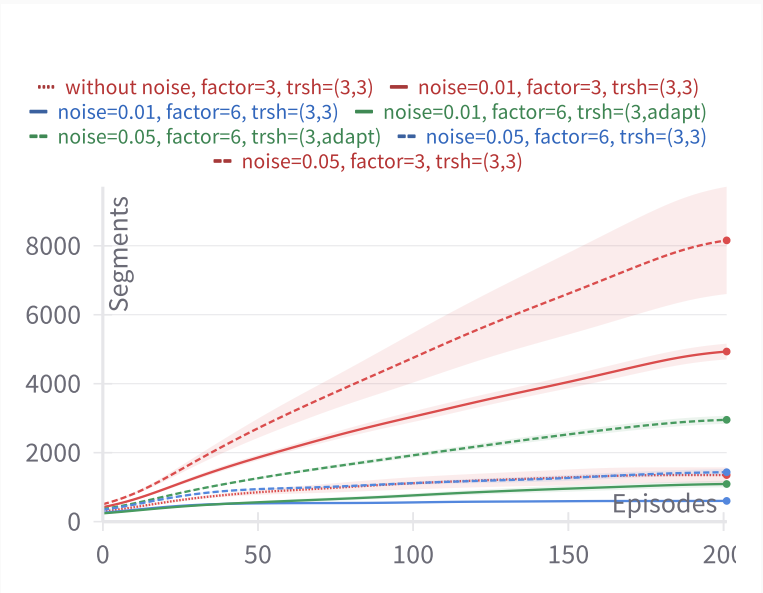
Примеры исходных данных raw (слева) и их предварительно обработанной версии (справа) из сред Pinball(A) и AnimalAI(B).

DHTM Results: noisy observations in Pinball & AnimalAI



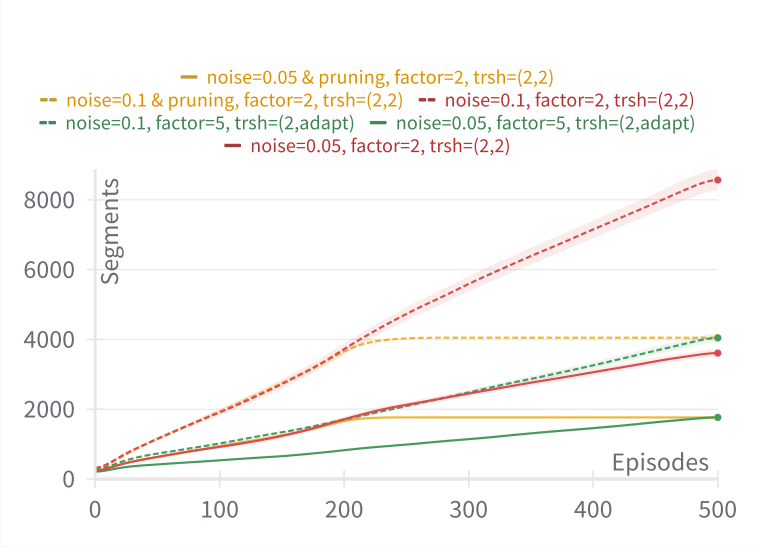


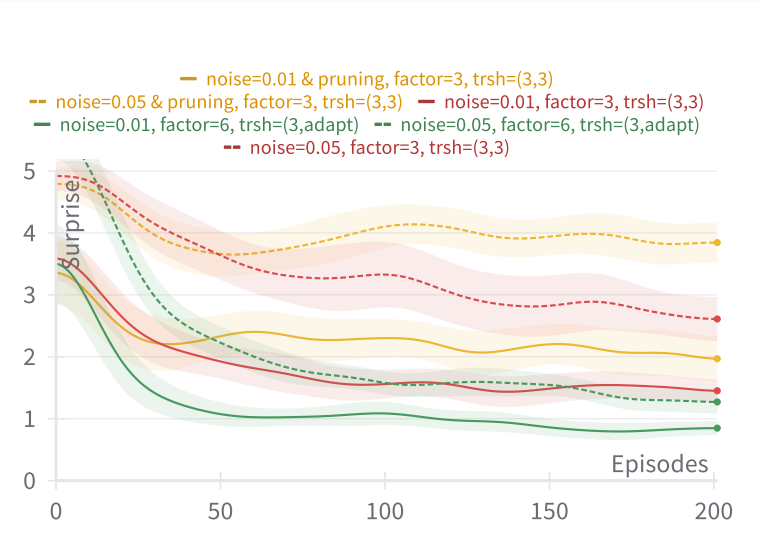


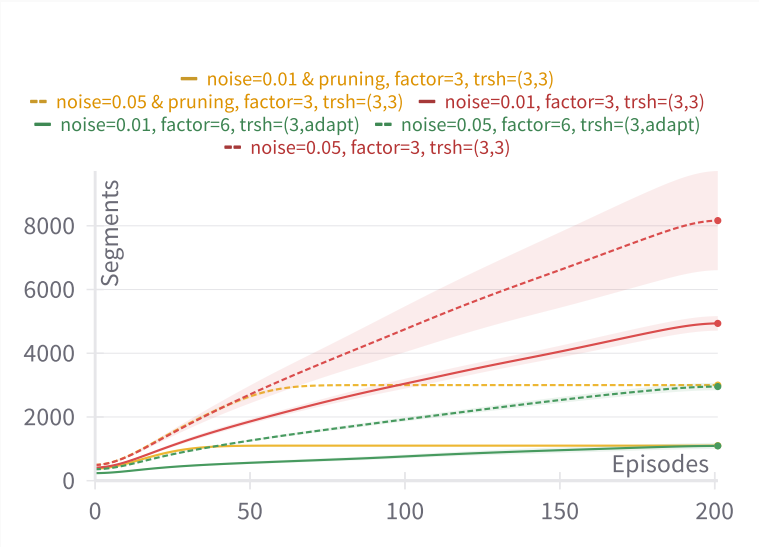


DHTM Results: vs Segments Pruning

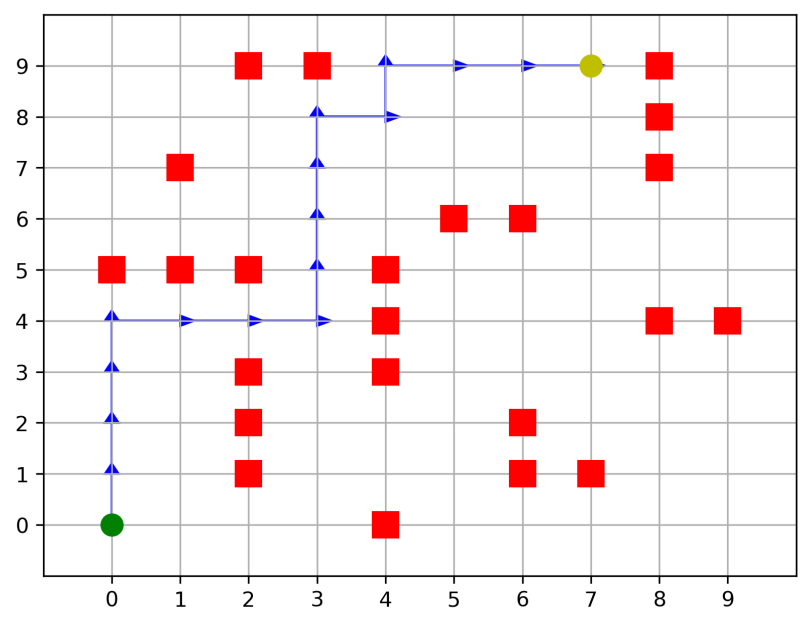




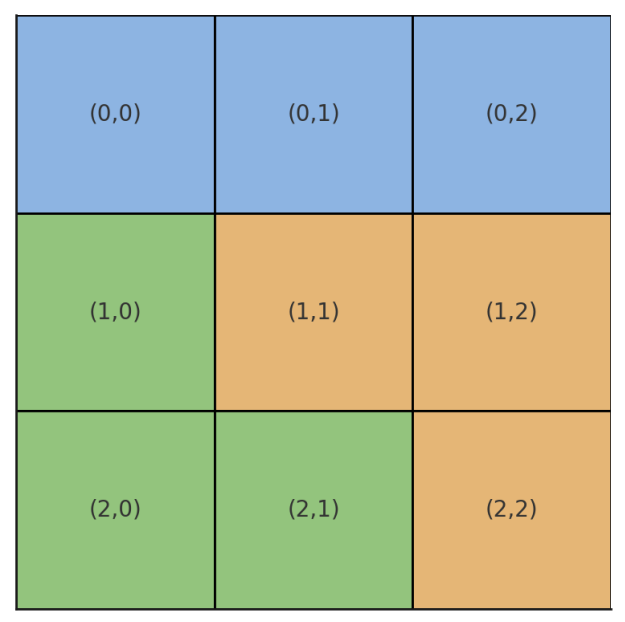




tPC: Experiments setup

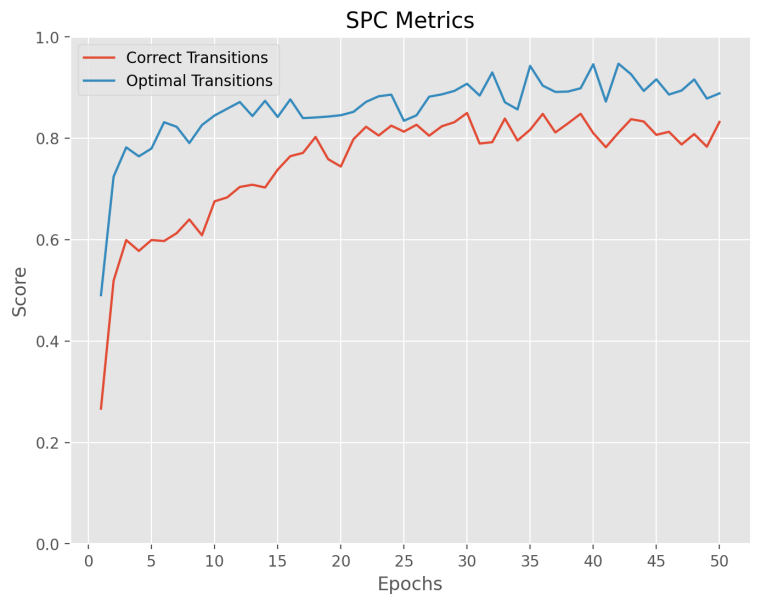


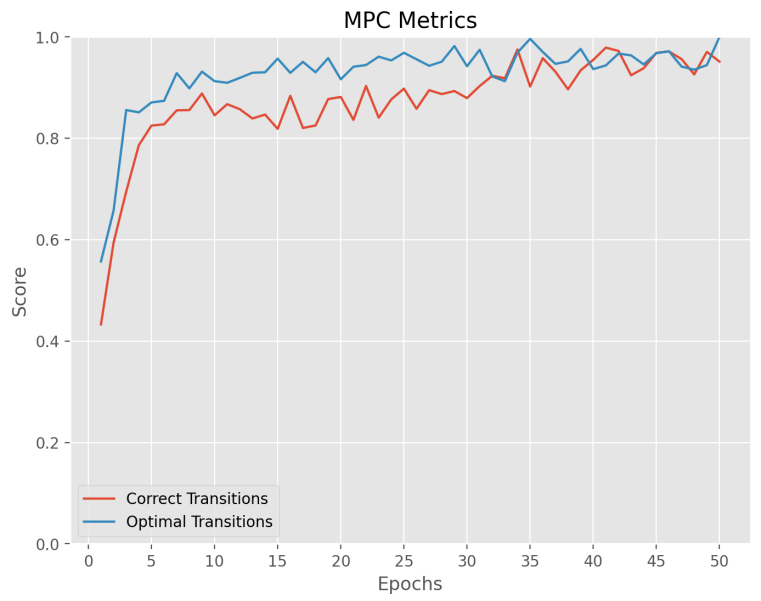
Grid-world среда. Random start/end points.
Val — отложенный набор точек, не фигурирующий в train.

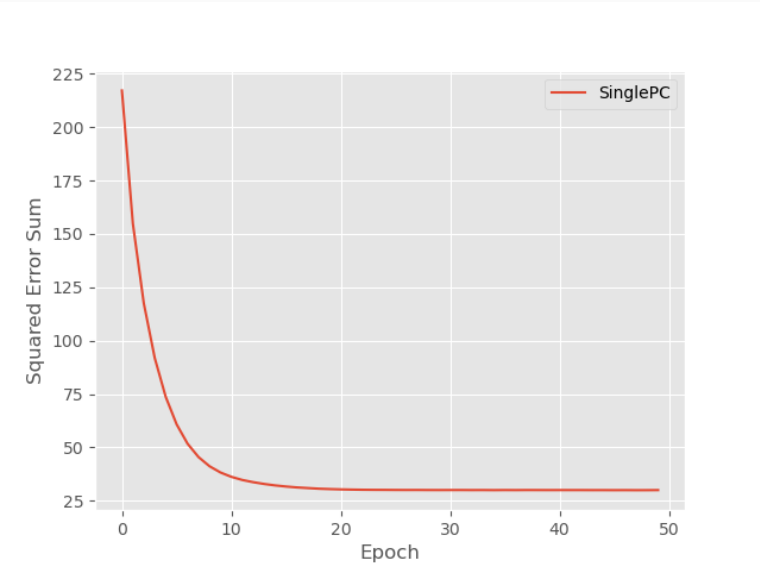


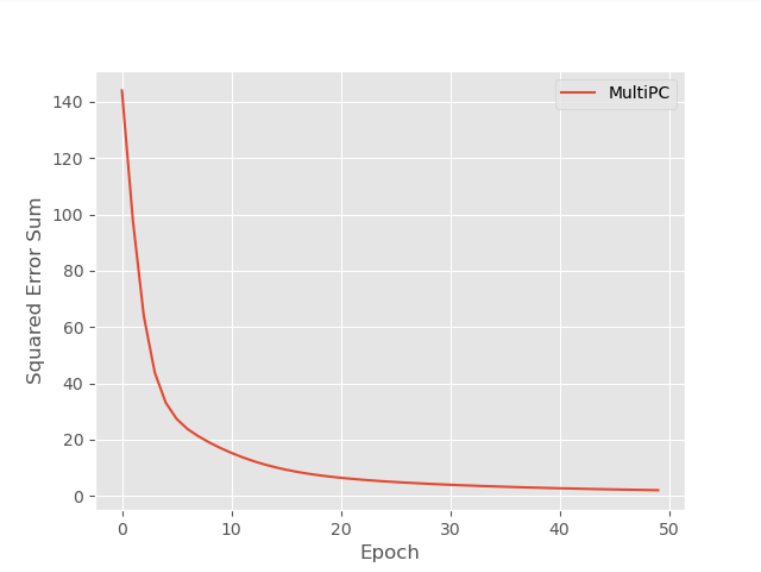
Grid-world среда с кодированием позиций.

tPC Results: training





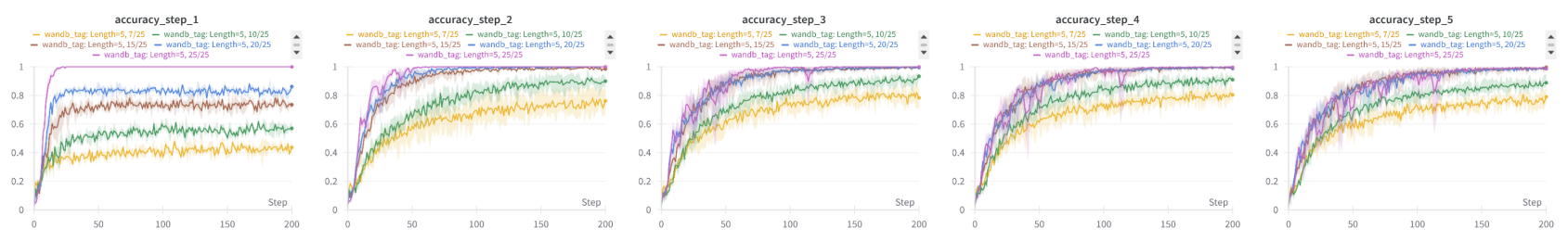




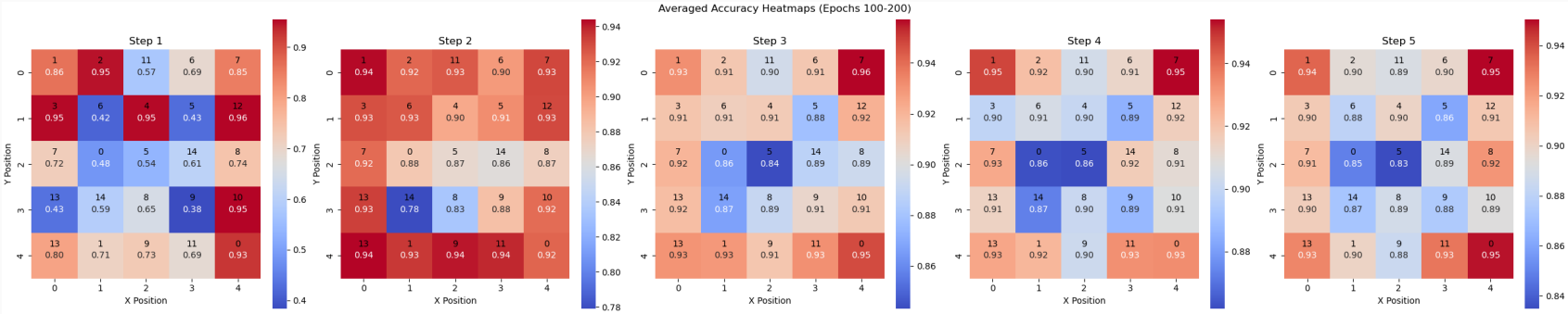
tPC Results: eval metrics

Mode	Metric	MCAHN	Single-tPC	Multi-tPC
online	correct transitions	0.693	0.712	0.861
	optimal transitions	0.712	0.756	0.879
offline	correct transitions	-	0.325	0.575
	optimal transitions	-	0.381	0.591

tPC Results: POMDP training

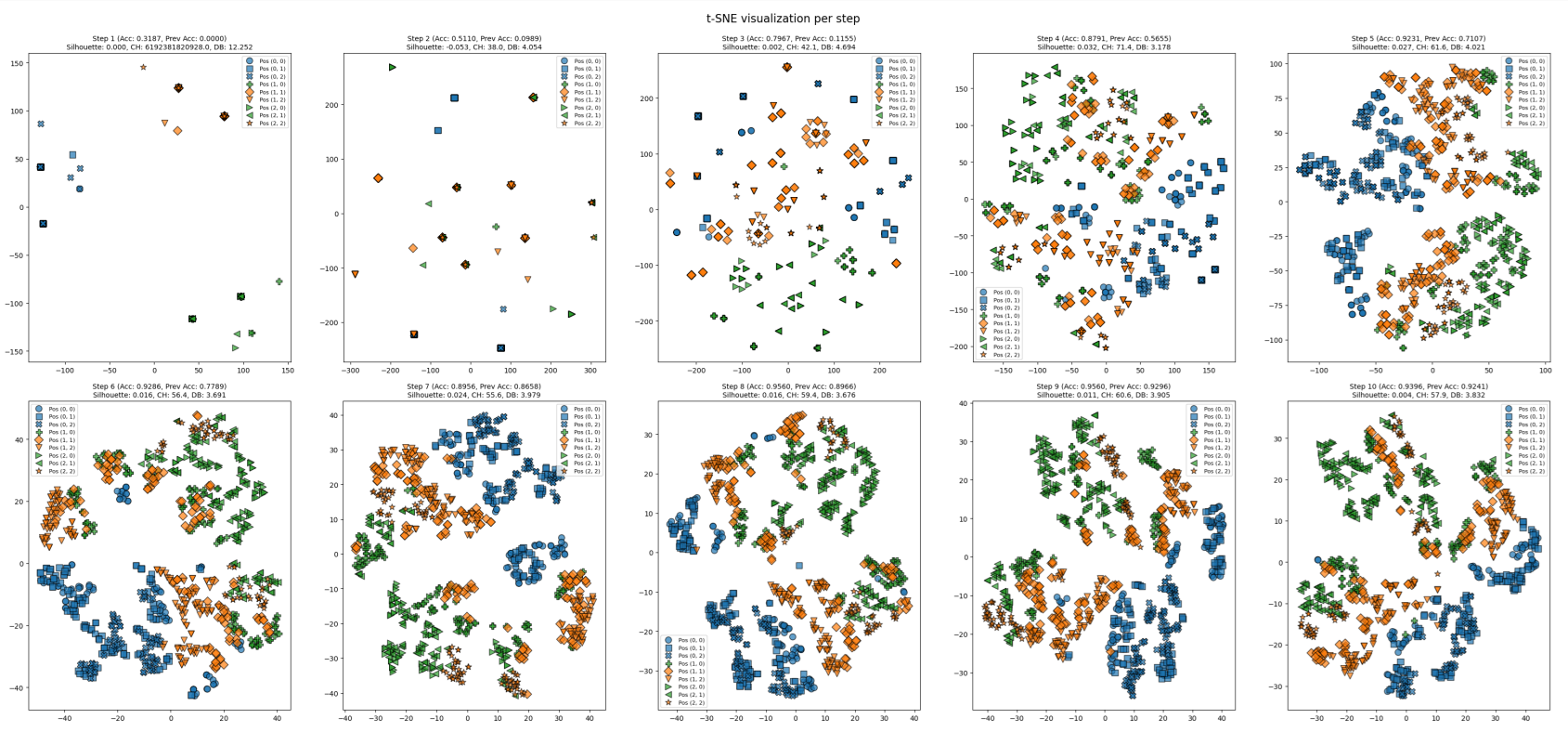


Accuracy per steps



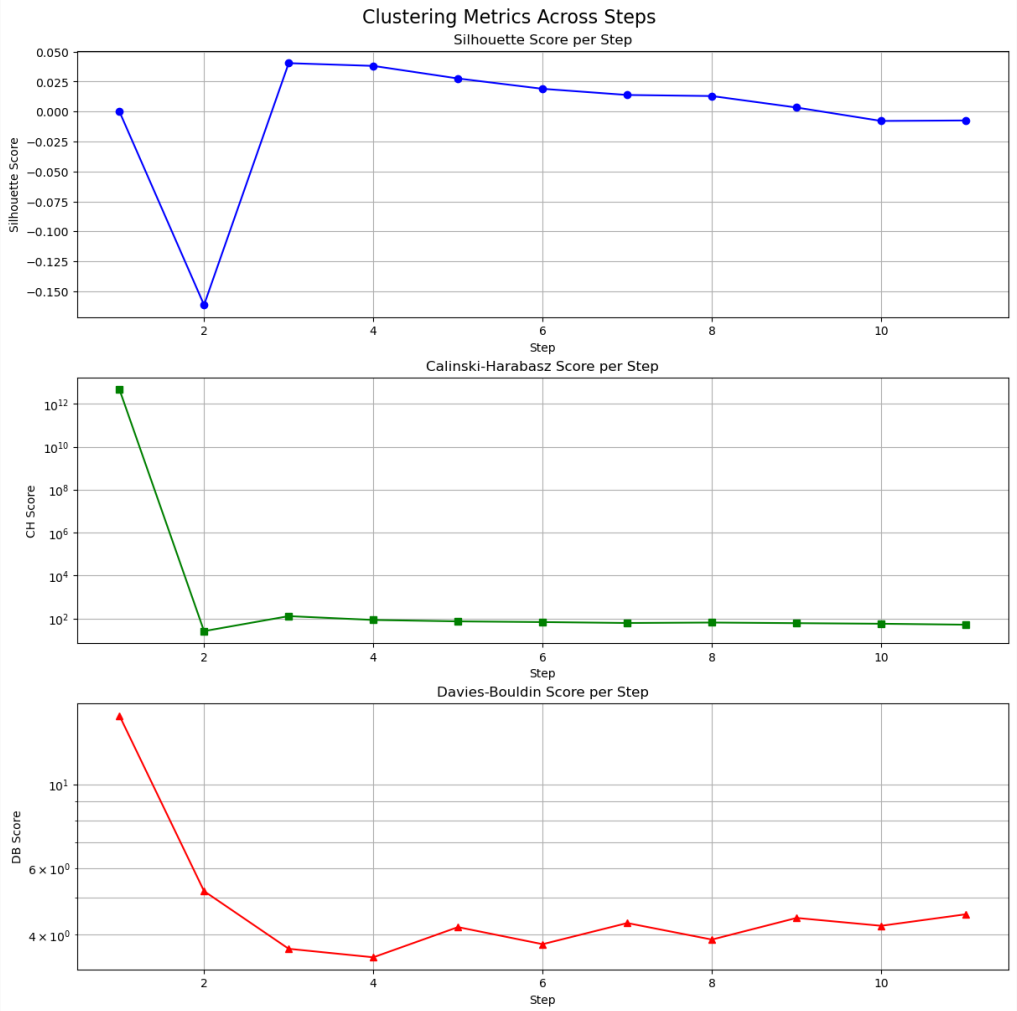
Heatmaps on 15/25

tPC Results: hidden layer in POMDP

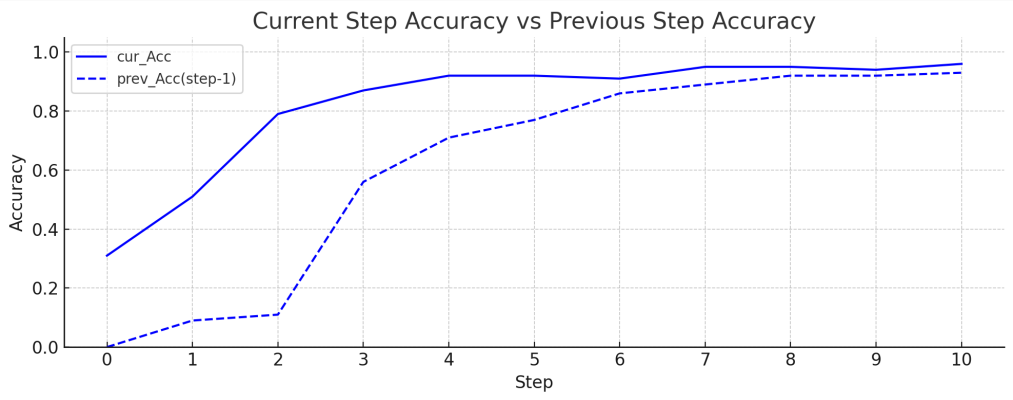


t-SNE per steps

tPC Results: clustering analysis in POMDP



tPC Results: clustering analysis in POMDP



Conclusion

- Улучшения DHTM для повышения шумоустойчивости и эффективности за счет адаптивных сегментов.
- Исследования tPC и HN моделей в задачах, требующих формирования представления о среде.
- Исследование tPC в POMDP средах, анализ скрытого слоя в таких средах.
- В дальнейшем предполагается:
  - Дополнительные эксперименты с tPC в качестве модели мира в RL
  - Расширение tPC в сторону многослойности
  - Публикация результатов по части tPC