

Структурирование памяти последовательностей с помощью локальных правил обучения

Дживеликян Е.А.



BioRL | Биологически интерпретируемые агенты

Чем мы занимаемся?

- Решение задач обучения с подкреплением
- С учётом нейрофизиологических ограничений в аналогии мозг–искусственная нейронная сеть ([нейроморфные](#) алгоритмы)

BioRL | Биологически интерпретируемые агенты

Чем мы занимаемся?

- Решение задач обучения с подкреплением
- С учётом нейрофизиологических ограничений в аналогии мозг–искусственная нейронная сеть ([нейроморфные](#) алгоритмы)

Зачем?

- Потенциально новые, более эффективные алгоритмы обучения
- В перспективе, использование [нейроморфных](#) вычислительных устройств
- Когнитивное моделирование

BioRL | Биологически интерпретируемые агенты

Чем мы занимаемся?

- Решение задач обучения с подкреплением
- С учётом нейрофизиологических ограничений в аналогии мозг–искусственная нейронная сеть ([нейроморфные](#) алгоритмы)

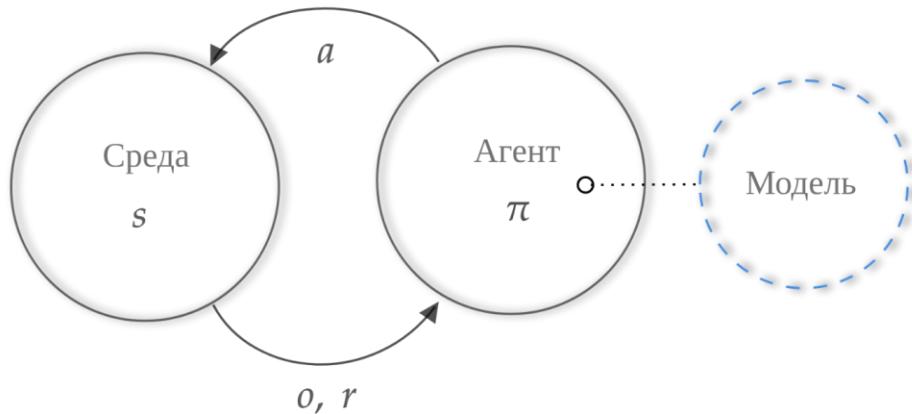
Зачем?

- Потенциально новые, более эффективные алгоритмы обучения
- В перспективе, использование [нейроморфных](#) вычислительных устройств
- Когнитивное моделирование

Глобальная цель:

разработка более [автономных и адаптивных](#) робототехнических устройств

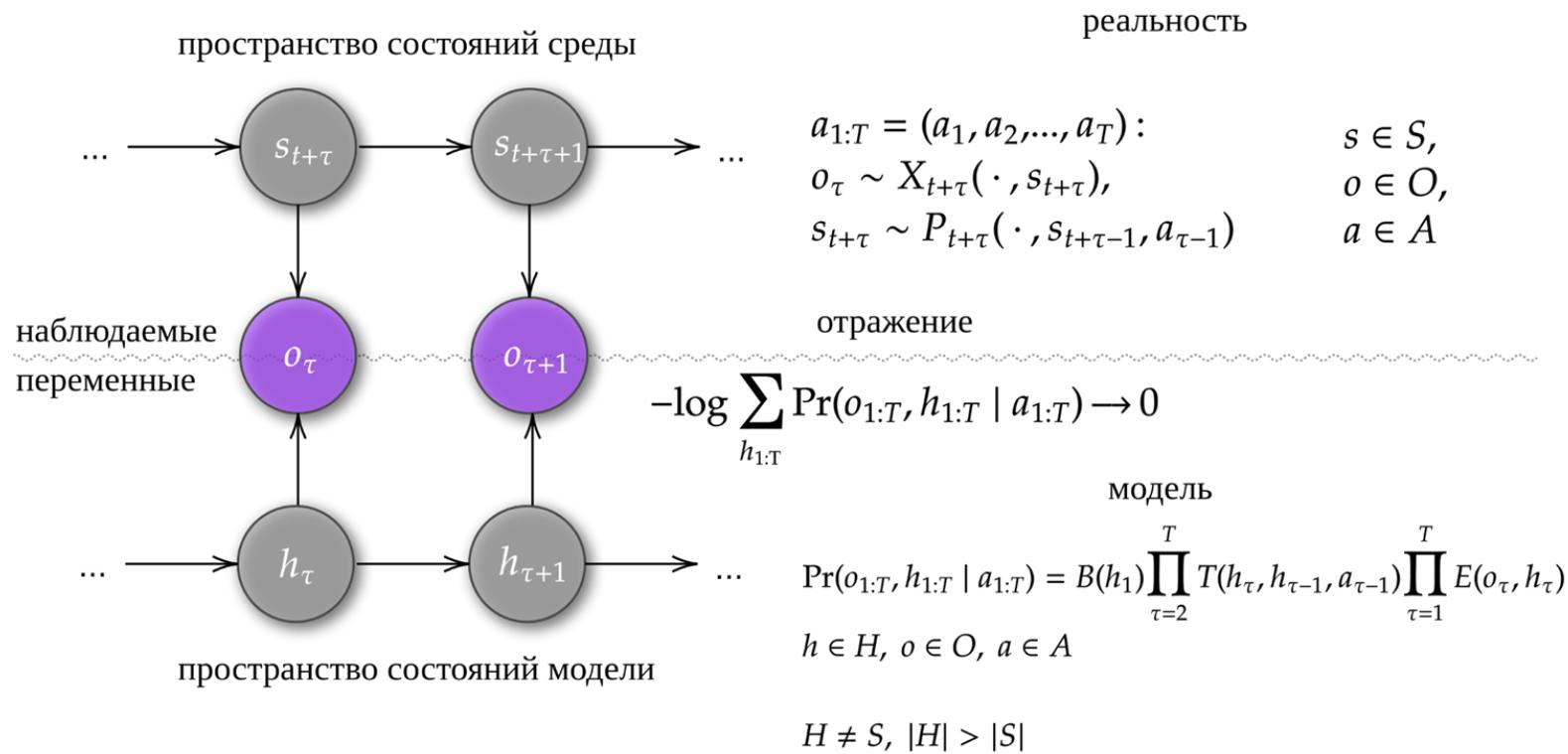
Задача: обучение с подкреплением



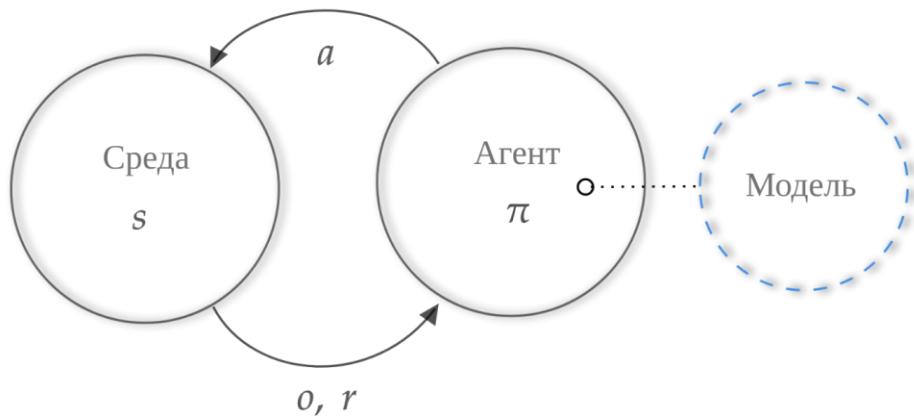
- Агент на основе опыта (**эпизодов** или траекторий) формирует **модель** среды для максимизации вознаграждения в длительной перспективе (lifelong learning, continual learning)

эпизод $((o_1, a_1, r_1), (o_2, a_2, r_2), \dots (o_T, a_T, r_T))$

Модель мира (среды) агента



Модель мира (среды) агента



Свойства модели:

- Эффективность использования опыта (sample efficiency)
- Обобщающая способность (выделение закономерностей | структуры среды)

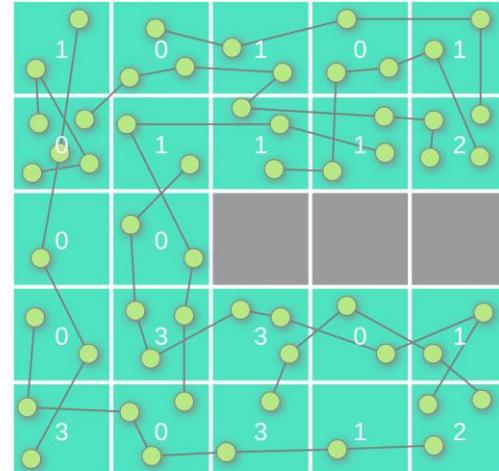
Sample efficiency vs обобщение

Клеточная среда

1	0	1	0	1
0	1	1	1	2
0	0			
0	3	3	0	1
3	0	3	1	2

Sample efficiency vs обобщение

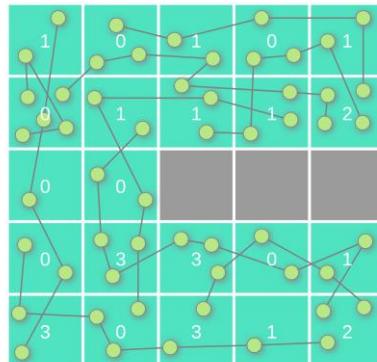
Память траекторий | [эпизодов](#)



скрытое
состояние

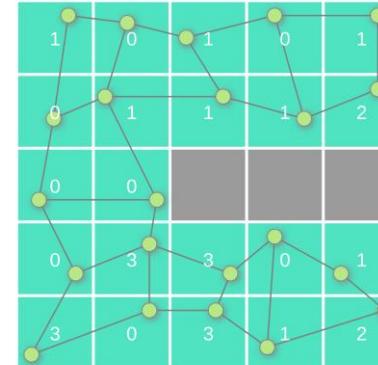
Sample efficiency vs обобщение

Память траекторий | **эпизодов**



скрытое
состояние

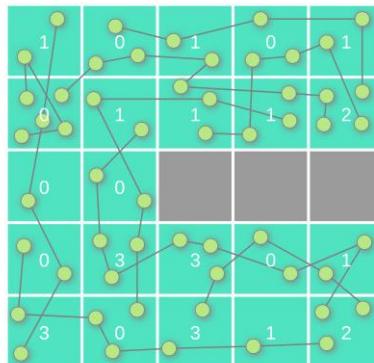
Структурированная
память переходов среды



скрытое
состояние =
состояние
среды

Sample efficiency vs обобщение

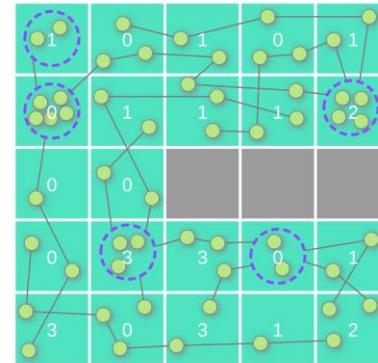
Память траекторий | эпизодов



скрытое
состояние

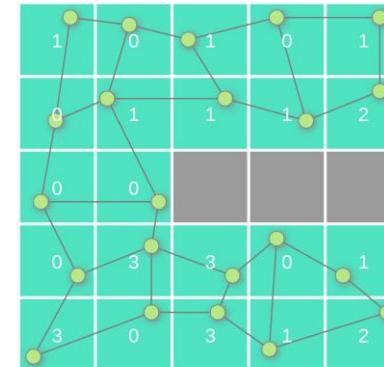
Структуризаци

1



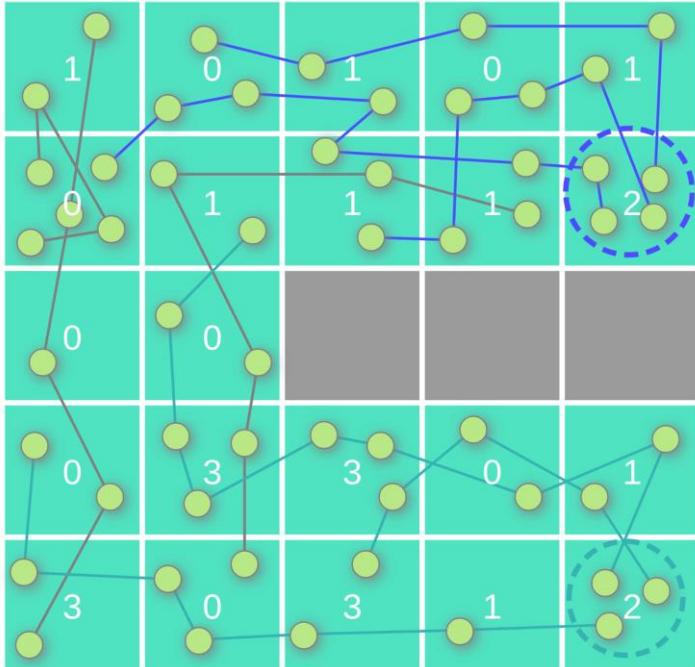
 скрытое
состоение

Структурированная память переходов среды



 скрытое
состояние = состояние
среды

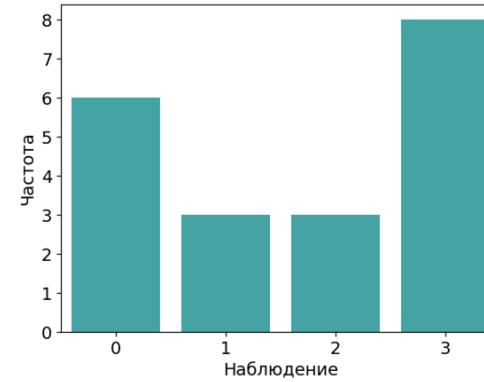
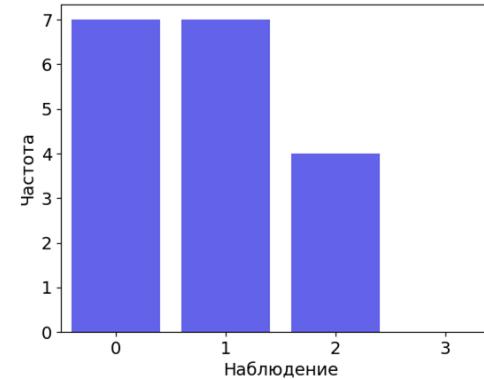
Successor Features | Распределение будущих наблюдений



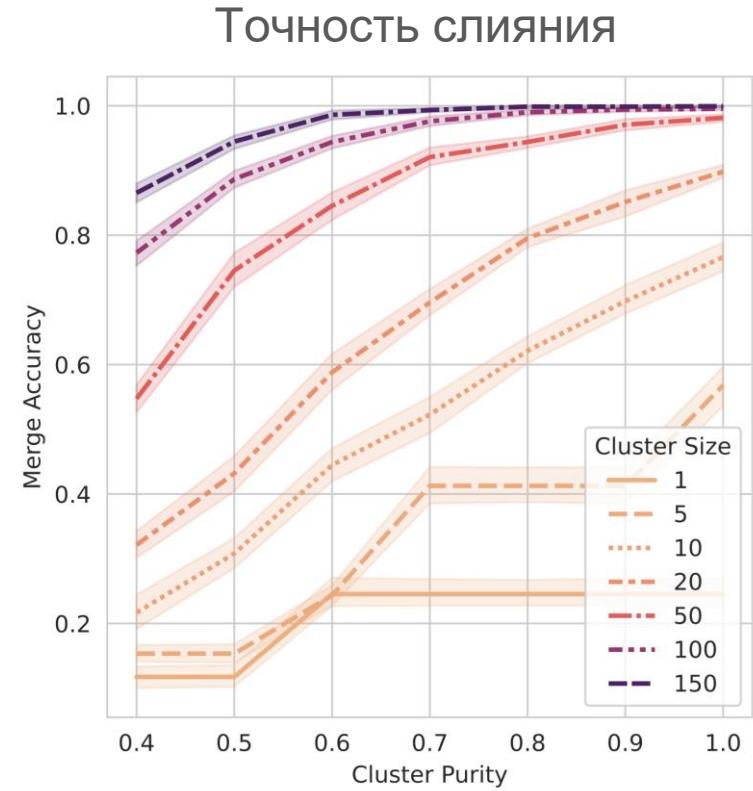
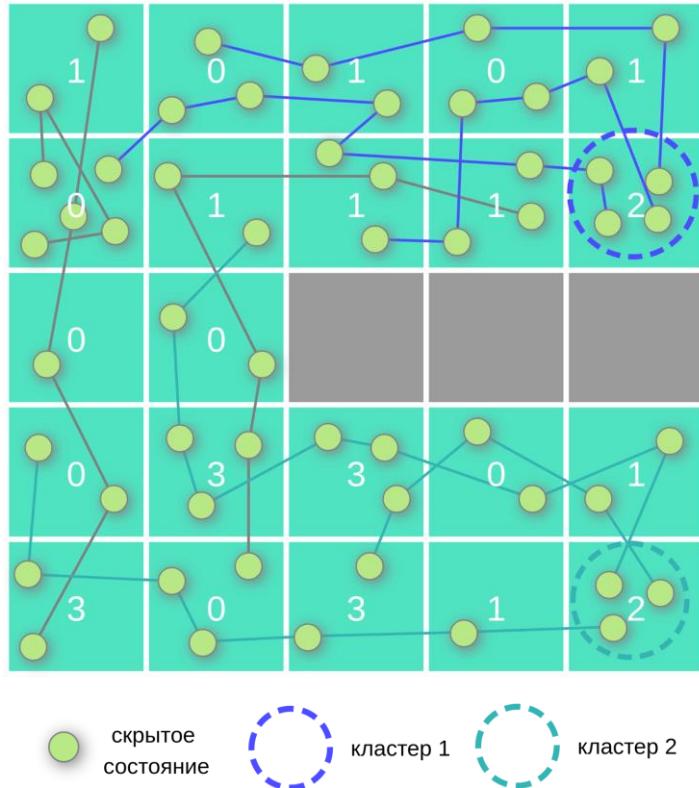
скрытое
состояние

клuster 1

кластер 2

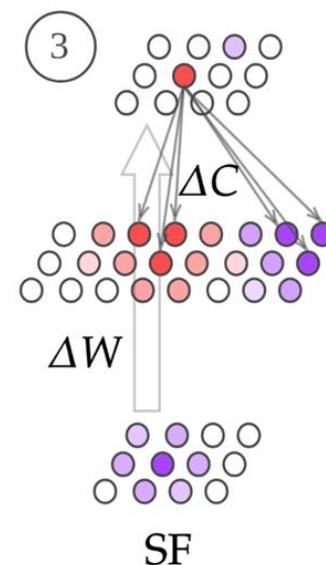
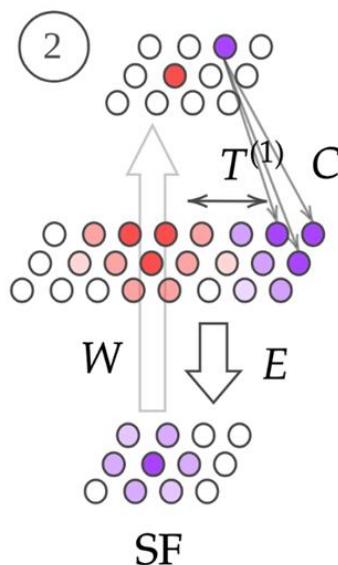
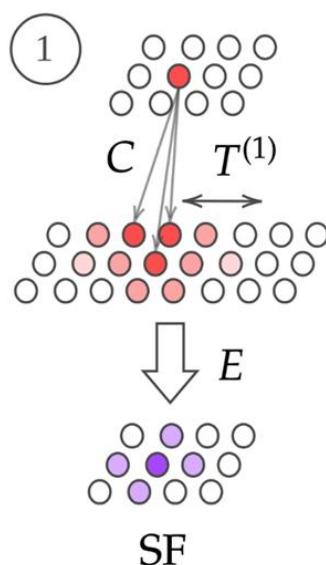
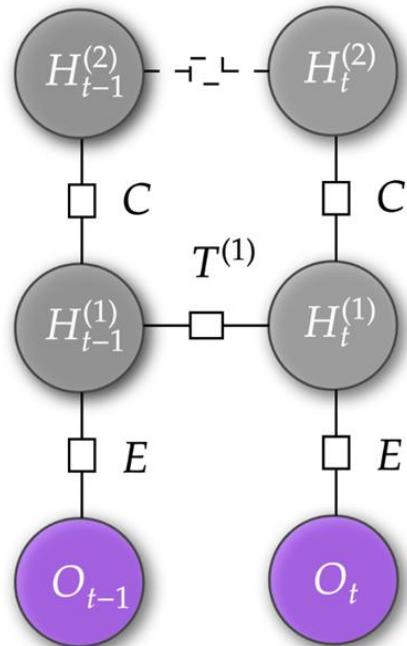


Слияние кластеров с похожим SF-представлением



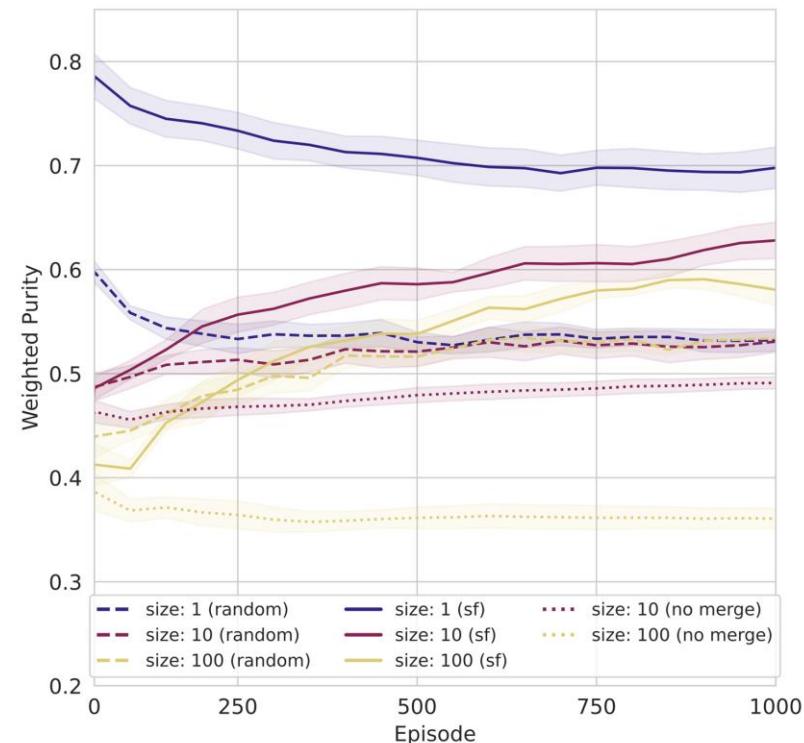
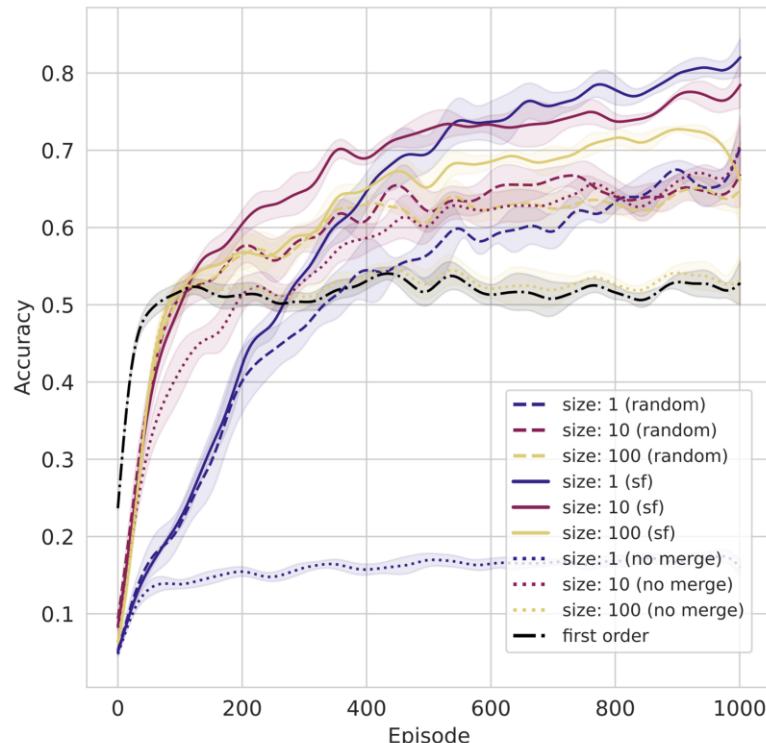
Двухуровневая модель памяти: нейронная реализация

$$T^{(2)} \propto C^T (T^{(1)})^T C$$



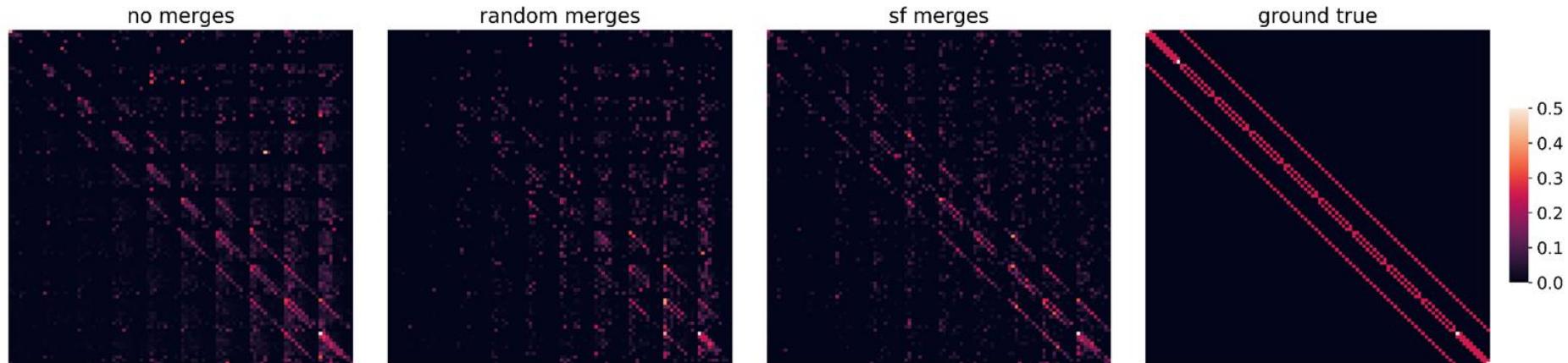
Результаты экспериментов

Предсказание наблюдений в клеточной среде 10x10 с 10-ю цветами

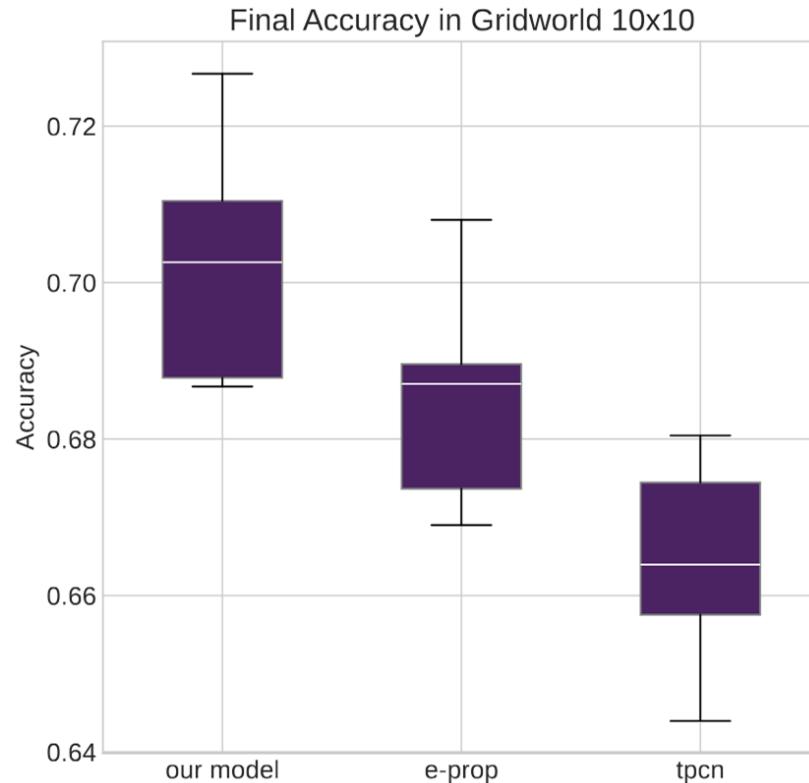


Результаты экспериментов

Восстановленные матрицы переходов



Сравнение с другими моделями



Модель	Число эпизодов	Точность
Наша модель	1000	0.7
E-Prop RNN	5000	0.68
tPCN	2000	0.67

Заключение

Постепенная структуризация памяти **эпизодов** может быть более эффективной (sample efficient), чем структуризация изначально **случайной** матрицы перехода (как в обычных RNN).

Преимущества модели

- Можно реализовать обучение на **локальных правилах**
- Высокая эффективность использования опыта
- Структурированное представление

BioRL | Team

- [Александр Панов](#), д.ф.-м.н., руководитель команды, директор лаборатории Cognitive AI Systems в AIRI
- [Пётр Кудеров](#), к.ф.-м.н., научный сотрудник AIRI
- [Евгений Дживеликян](#) (докладчик), научный сотрудник AIRI
- [Никита Байнаев-Мангилев](#), аспирант Центра когнитивного моделирования МФТИ

Вопросы сюда ⇒



Препринт по теме:

