

МОДЕЛЬ ДИНАМИКИ КАТЕГОРИЗАЦИОННОГО ОБУЧЕНИЯ У ПТИЦ НА ОСНОВЕ СКРЫТЫХ МАРКОВСКИХ ЦЕПЕЙ

Арцатбанова Е.П. (МГУ), Герштейн Я. А. (МФТИ), Пузик М.А. (МГУ)

Научные руководители – Дживеликян Е.А. (МФТИ), Диффинэ Е.А. (МГУ)

Категоризация – когнитивный процесс, ведущий к способности классифицировать объекты, то есть воспринимать похожие, но не идентичные стимулы, как эквивалентные. Эта ключевая когнитивная способность осуществляется биологическими системами разного уровня и позволяет назначать функциональные атрибуты предметам в окружающем мире. В то же время, назначение атрибутов, или классификация - это базовая задача любой системы воплощённого искусственного интеллекта или искусственного агента. Изучение принципов категоризации у живых систем, с одной стороны, позволяет конструировать более адаптивные и энергоэффективные искусственные системы, используя нейроморфные вычислительные устройства, с другой, позволяет приблизиться к механистическому пониманию когнитивных процессов как результата нейронных взаимодействий (Hassabis et al., 2017).

Птицы особенно удобны для изучения процесса категоризации, т. к., во-первых, как и человек, они имеют хорошо развитую визуальную модальность, что упрощает разработку экспериментов и их интерпретацию. Во-вторых, многие виды птиц, практически с рождения способны взаимодействовать с окружающим миром, проявляя способности к обучению. Поэтому, в качестве модельного животного в данном исследовании рассматриваются цыплята (*gallus domesticus*). Для обучения птиц использовали экспериментальную модель быстрого обучения на «бусиничном полу» (bead floor; Tiunova et al., 1996). Цыплят в возрасте 2 суток помещали в камеру, на полу которой были приклеены бусины разных цветов и рассыпан корм. В сеансе быстрого обучения птицы формировали категорию «несъедобных объектов» – бусин, и относили к этой категории все бусины в камере. На основе видеозаписей поведения птиц в разработанной модели и выявления их выбора на каждом шаге обучения был проведен пошаговый анализ логики принятия решения и учета результатов проб и ошибок во всей последовательности обучения.

Для выявления закономерностей в поведении цыплят использовалась модель скрытых марковских цепей. Можно предположить, что наблюдаемые последовательности клевков бусинок могут быть обусловлены некоторыми скрытыми марковскими состояниями, которые можно выявить с помощью модели скрытых марковских цепей, обученной с использованием алгоритма Баума-Велша. Используя данный подход, можно увидеть, что наблюдаемые последовательности цветов бусинок неудачных клевков действительно имеют дополнительную структуру, помимо просто частотной: сравнивая оригинальные и перемешанные последовательности, мы можем видеть, что распределения предсказаний модели, обученной на оригинальной последовательности, отличаются от частотного распределения перемешанной последовательности. Это позволяет сделать предположение, что в процессе обучения категоризации, цвет бусинки влияет на её отнесение к той или иной категории.

Чтобы проверить данную гипотезу, была также разработана модель другого эксперимента, используя пол с бусинами одного цвета. В каждом отдельном эпизоде обучения птицам предъявляли пол с новым цветом бусин. Результаты так же показали, что формирование категории зависит от последовательности предъявленных цветов. Подобное поведение может быть связано как с врожденными функциональными системами цыплят, так и с их индивидуальным опытом.

Далее планируется обучение искусственного агента на основе модели категоризации, предложенной в Turner et al., 2019, который будет решать те же задачи категоризации, но на виртуальном бусиничном полу. Что позволит исследовать различные методы формирования представлений визуального стимула с помощью искусственных нейронных сетей и сравнивать поведение искусственного агента с поведением цыплят в аналогичных экспериментах. Мы рассчитываем получить более значимые различия моделей, чем были представлены в аналогичной работе (Pandey et al., 2024), поскольку будет исследоваться не только конечная метрика успеха категоризации, но и динамика обучения. На основании этих виртуальных экспериментов можно будет предлагать модели экспериментов с живыми птицами, воспроизводя условия, в которых сравниваемые модели формирования признаков давали бы видимые различия в поведении птиц.

Исследование выполнено при финансовой поддержке гранта Некоммерческого фонда содействия развитию науки и образования ИНТЕЛЛЕКТ, а также при поддержке Междисциплинарной научно-образовательной школы "Мозг, когнитивные системы, искусственный интеллект" МГУ имени М.В. Ломоносова.

Hassabis, Demis, Dharshan Kumaran, Christopher Summerfield, and Matthew Botvinick. 2017. Neuroscience-Inspired Artificial Intelligence. *Neuron*. 95 (2): 245–58.

Tiunova A., Anokhin K., Rose S., Mileusnic R. 1996. Involvement of glutamate receptors, protein kinases, and protein synthesis in memory for visual discrimination in the young chick. *Neurobiology of Learning and Memory*. 65(3), 233–234.

Turner BM. Toward a common representational framework for adaptation. 2019. *Psychol Rev*. 126(5):660-692.

Pandey L, Lee D, Wood SMW, Wood JN. 2024. Parallel development of object recognition in newborn chicks and deep neural networks. *PLoS Comput Biol*. 20(12):e1012600.