

УДК 004.81

Метапознание и прогнозирование последствий в когнитивных архитектурах

Гапонов Владислав Родионович^(*)

gaponovvr@student.bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Россия

Аннотация. Рассмотрены современные подходы к развитию когнитивных архитектур с учетом механизмов метапознания и прогнозирования последствий. Показано, что добавление в архитектуру метакогнитивного уровня обеспечивает саморегуляцию и адаптивность систем искусственного интеллекта. Проведен сравнительный анализ зарубежных и российских исследований, а также выявлены ключевые ограничения существующих решений. Предложены принципы построения когнитивных архитектур, которые ориентированы на рефлекссию.

Ключевые слова: когнитивные архитектуры, метапознание, рефлексия, прогнозирование, искусственный интеллект

Metacognition and predicting consequences in cognitive architectures

Gaponov Vladislav Rodionovich^(*)

gaponovvr@student.bmstu.ru

BMSTU, Moscow, Russia

Abstract. Modern approaches to the development of cognitive architectures are considered, taking into account the mechanisms of metacognition and prediction of consequences. It is shown that the addition of a metacognitive layer to the architecture ensures the self-regulation and adaptability of artificial intelligence systems. A comparative analysis of foreign and Russian studies has been conducted, and key limitations of existing solutions have been identified. The principles of building cognitive architectures that are focused on reflection are proposed.

Keywords: cognitive architectures, metacognition, reflection, forecasting, artificial intelligence

Введение. Когнитивные архитектуры применяются для моделирования памяти, рассуждений и обучения, однако остаются ограниченными при работе в динамических и неопределенных условиях. В классических системах, таких как ACT-R и SOAR, отсутствует явный уровень рефлексии, что снижает их адаптивность в динамически изменяемой среде [1].

Зарубежные исследования показывают, что расширение когнитивных моделей за счет добавления метауровня позволяет системам не только выполнять различные задачи, но и анализировать собственные действия. Российские работы также подчеркивают значимость метакогнитивных процессов человека, демонстрируя рост эффективности при использовании стратегий самоконтроля [2, 3]. Следовательно, при разработке когнитивных архитектур важно делать их работу похожей на работу человеческого мозга.

Методы и материалы; результаты. Аналитический обзор показал, что символические архитектуры обеспечивают логическую строгость, но ограничены в обучении. Гибридные системы, такие как CLARION, демонстрируют адаптивность, однако их внутренние процессы непрозрачны, и система напоминает черный ящик [1]. Архитектуры с метакогнитивным уровнем, такие как MIDCA и MCL, включают механизмы мониторинга и коррекции, но функции прогнозирования последствий в них развиты слабо.

В российских исследованиях показано, что метакогнитивный контроль повышает устойчивость когнитивных процессов в условиях неопределенности [3]. В различных экспериментах выявлено, что развитие метакогнитивных способностей положительно влияет на глубину усвоения знаний и стратегическое планирование [2]. Соответственно, если применить эти методы при разработке когнитивных архитектур, то эффективность таких систем должна повыситься. А их структура будет похожа на структуру человеческого мозга. В таблице представлен сравнительный анализ когнитивных архитектур.

Сравнительный анализ когнитивных архитектур

Архитектура	Поддержка метапознания	Прогнозирование последствий	Прозрачность решений
SOAR	Неявная (через правила)	Отсутствует	Высокая
ACT-R	Частичная, ограниченная	Отсутствует	Средняя
CLARION	Частично реализована	Ограниченное	Низкая
MIDCA	Явная (метацикл)	Частично реализовано	Средняя
MCL	Явная (метаконтроль)	Частично реализовано	Средняя

Заключение. Интеграция механизмов метапознания и прогнозирования последствий в когнитивные архитектуры является важным направлением развития искусственного интеллекта. Такие системы способны не только адаптироваться к изменяющимся условиям, но и обеспечивать саморегуляцию своей работы. Рефлексия в когнитивных архитектурах открывает новые возможности для применения таких систем в образовании, интеллектуальных помощниках и критически важных системах поддержки принятия решений.

Список источников

- [1] Sun R. Cognition and Multi-Agent Interaction. From Cognitive Modeling to Social Simulation. New York, Cambridge University Press, 2005, 450 p.
- [2] Кроколева С.С. Роль метакогнитивных способностей в эффективности усвоения учебного материала с применением чата-GPT. *Цифровая гуманитаристика и технологии в образовании (DHTE-2024)*. V Междунар. науч.-практ. конф.: сб. ст. Москва, Московский гос. псих.-пед. ун-т, 2024, с. 387–403.
- [3] Морошкина Н.В., Зверев И.В., Нездоймышапко Л.А., Тихонов Р.В. Метакогнитивный мониторинг и контроль в ситуации распределенного познания. *Вестник Санкт-Петербургского университета. Психология*, 2023, т. 13, № 3, с. 324–346. <https://doi.org/10.21638/spbu16.2023.303>