

Зайцев Евгений Игоревич

кандидат технических наук, доцент, доцент кафедры цифровых технологий обработки данных, МИРЭА – Российский технологический университет, Москва. ORCID: 0000-0002-1979-5611, ResearcherID: ABA-4823-2020, SPIN-код: 9662-7658, AuthorID: 57218190023

Электронный адрес: zajcev@mirea.ru

Evgeniy I. Zaytsev

Ph.D. of Engineering Sciences, Docent, Associate Professor at the Department of digital data processing technologies, MIREA – Russian Technological University, Moscow. ORCID: 0000-0002-1979-5611, ResearcherID: ABA-4823-2020, SPIN-code: 9662-7658, AuthorID: 57218190023

E-mail address: zajcev@mirea.ru

Нурматова Елена Вячеславовна

кандидат технических наук, доцент, доцент кафедры разработки программных решений и системного программирования, МИРЭА – Российский технологический университет, Москва. ORCID: 0000-0001-8511-0978, ResearcherID: ABA-4823-2020, SPIN-код: 7036-3661, AuthorID: 57218190023

Электронный адрес: nurmatova@mirea.ru

Elena V. Nurmatova

Ph.D. of Engineering Sciences, Docent, Associate Professor at the Department of software development and system programming, MIREA – Russian Technological University, Moscow. ORCID: 0000-0001-8511-0978, ResearcherID: ABA-4823-2020, SPIN-code: 7036-3661, AuthorID: 57218190023

E-mail address: nurmatova@mirea.ru

АГЕНТНО-ОРИЕНТИРОВАННАЯ ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ: АРХИТЕКТУРА И РАЗРАБОТКА

Аннотация. В статье рассматривается архитектура, анализ, проектирование и реализация агентно-ориентированной интеллектуальной системы поддержки принятия решений. Описываются структура программных агентов, организация их совместной работы, реализуемые программными агентами функции и запросы к распределенной базе знаний, алгоритмы многоагентного обучения с подкреплением. Рассматриваются методы повышения производительности и надежности системы. Описана подсистема управления распределенной базой знаний, способная адаптироваться к меняющимся условиям внешней среды, а также своевременно и адекватно реагировать на ситуации, возникающие при дефиците вычислительных или временных ресурсов. Полученные результаты демонстрируют эффективность использования агентно-ориентированного подхода для создания интеллектуальной системы поддержки принятия решений, способной находить и анализировать решения сложных проблем в условиях неопределенности и динамичной внешней среды.

Ключевые слова: многоагентная система, программные агенты, распределенная база знаний, искусственная нейронная сеть, обучение с подкреплением.

Для цитирования: Зайцев Е.И., Нурматова Е.В. Агентно-ориентированная интеллектуальная система поддержки принятия решений: архитектура и разработка // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ, управление. 2025. № 3. С. 85 – 96. DOI: 10.18137/RNU.V9I87.25.03.P.85

AGENT-BASED INTELLIGENT DECISION SUPPORT SYSTEM: ARCHITECTURE AND DEVELOPMENT

Abstract. The article considers the architecture, analysis, design and implementation of the Agent-based Intelligent Decision Support System. It describes the structure of software agents, the organization of their joint work, the functions implemented by software agents and queries to the distributed knowledge base, and algorithms for multi-agent reinforcement learning. The authors consider methods for increasing the system's performance and reliability, algorithms for optimizing the logical structure of the knowledge base. A subsystem for managing a distributed knowledge base is described, capable of adapting to changing environmental conditions, as well as responding in a timely and adequate manner to situations arising from a shortage of computing or time resources. The obtained results demonstrate the effectiveness of the agent-oriented approach for creating an intelligent decision support system capable of finding and analyzing solutions to complex problems in conditions of uncertainty and a dynamic environment.

Keywords: multi-agent system, software agents, distributed knowledge base, artificial neural network, reinforcement learning.

For citation: Zaytsev E.I., Nurmatova E.V. (2025) Agent-based Intelligent Decision Support System: Architecture and development. *Vestnik of Russian New University. Series: Complex Systems: Models, analysis, management.* No. 3. Pp. 85 – 96. DOI: 10.18137/RNUV9187.25.03.P.85 (In Russian).

Введение

Разработка агентно-ориентированной интеллектуальной системы поддержки принятия решений (далее – АИСППР), состоящей из слабосвязанных интеллектуальных программных модулей, независимо решающих свои частные задачи и объединяющих свои локальные возможности для решения общей проблемы, расширяет сферу автоматизации информационных процессов на плохоформализуемые проблемные области, в которых для принятия обоснованных решений требуется анализировать большие объёмы данных, осуществлять логические выводы и адаптироваться к условиям неопределенности и нестационарности внешней среды.

Интеграция низкоуровневого (реактивного) восприятия с высокоуровневым логическим мышлением является одной из наиболее актуальных задач в области искусственного интеллекта [1; 2]. При разработке АИСППР основные аспекты интеграции реактивной (нейросетевой) и когнитивной (семиотической) моделей рассматриваются в контексте агентно-ориентированного подхода. Низкоуровневое восприятие прикладных программных агентов реализуется с помощью методов глубокого машинного обучения, а высокоуровневые рассуждения осуществляются с использованием логических и вероятностных представлений и логического вывода.

Подсистема представления и обработки знаний в АИСППР принципиально отличается от традиционных систем, базирующихся на знаниях (Knowledge-Based System – KBS), которые имеют только один интеллектуальный решатель и осуществляют поиск решения на графе редукции задачи с использованием полной и непротиворечивой базы знаний. Для повышения производительности традиционных KBS разработаны параллельные алгоритмы обработки редукционных графов, однако эти алгоритмы могут быть

реализованы только в мультипроцессорных системах с общей памятью (Shared Memory Processors – SMP).

В АИСППР, ориентированной на использование в мультикомпьютерных системах с распределенной памятью (No Remote Memory Access – NORMA), вместо одного интеллектуального решателя задач используется сеть автономных интеллектуальных (прикладных) программных агентов, взаимодействующих друг с другом посредством передачи сообщений. Инкапсулированные в прикладных программных агентах когнитивные структуры данных, методы логического вывода и искусственные нейронные сети позволяют агентам действовать рационально и адаптироваться к внешней среде. Подсистема обработки знаний и алгоритмы многоагентного обучения с подкреплением обеспечивают возможность накапливать опыт через взаимодействие агентов друг с другом и с окружающей средой [3].

В АИСППР для передачи сообщений используется специализированный механизм межпроцессного взаимодействия, реализованный с помощью системных программных агентов, входящих в состав системной библиотеки LibOS. Специализированные программные модули LibOS, ассоциированные с системными программными агентами, обеспечивают более высокую производительность по сравнению с универсальными библиотеками.

В процессе решения задач программные агенты обращаются к микросервисам, которые с целью повышения производительности и надежности АИСППР дублируются на разных вычислительных узлах. Балансировка нагрузки в вычислительном кластере, как и передача сообщений между узлами, осуществляется с использованием системных программных агентов. Повышение производительности АИСППР достигается также за счёт динамической оптимизации логической структуры распределенной базы знаний (далее – РБЗ) и адаптации её условиям нестационарной внешней среды.

Архитектура АИСППР

Агентно-ориентированная интеллектуальная система поддержки принятия решений имеет многоуровневую архитектуру, в которой на самом верхнем уровне прикладные программные агенты решают сложные плохоформализуемые задачи. На базовом уровне системные программные агенты планируют и управляют распределенными вычислительными ресурсами, взаимодействуя с помощью внутрипроцессных и внепроцессных вызовов API (Application Programming Interfaces) с операционной системой, системными службами и микросервисами (см. Рисунок 1).

В процессе обработки знаний прикладные программные агенты выполняют анализ, сравнение, спецификацию объектов и событий проблемной области. При спецификации осуществляется поиск в базе знаний значений атрибутов с использованием как четких, так и нечетких запросов.

Пример нечеткого запроса, включающего две лингвистические переменные (Performance, Price) с кусочно-линейными функциями принадлежности, а также результат его выполнения после обращения к реляционной таблице, содержащей пять записей, показан на Рисунок 2.

Для эффективной обработки большого количества запросов может потребоваться масштабирование системы, то есть добавление новых микросервисов для обработки возросшей нагрузки. Масштабирование повышает не только производительность, но и надежность АИСППР.

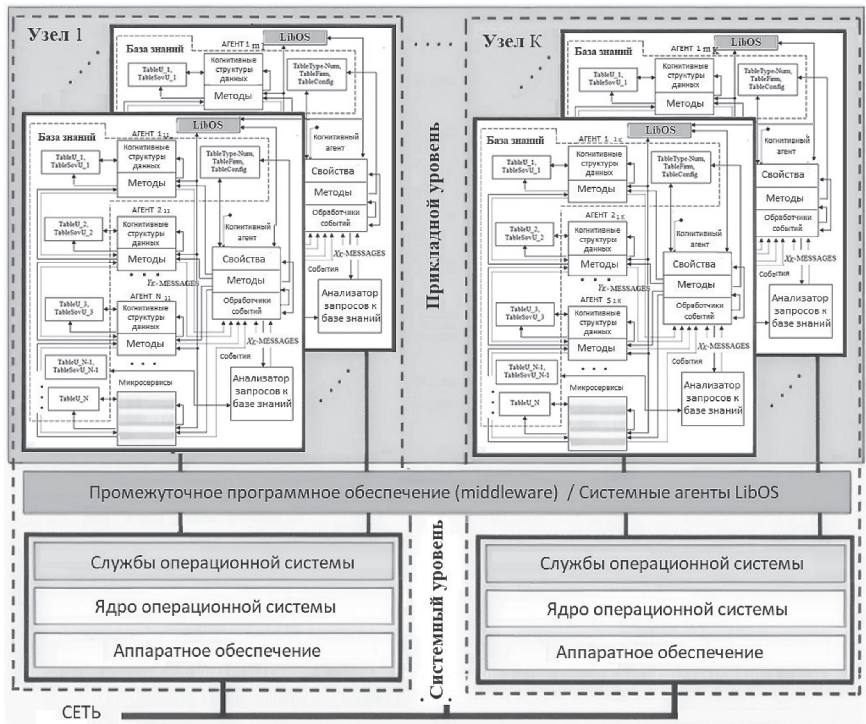


Рисунок 1. Архитектура АИСПП

Источник: здесь и далее рисунки выполнены авторами.

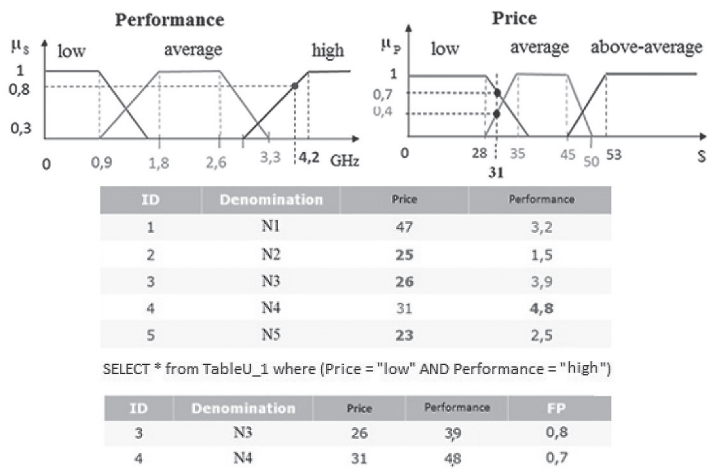


Рисунок 2. Функции принадлежности и результат нечеткого запроса

При взаимодействии с микросервисами прикладные программные агенты отправляют запросы системным программным агентам, которые балансируют нагрузку, перенаправляя эти запросы соответствующим службам. В случае выхода из строя одного из вы-

Агентно-ориентированная интеллектуальная система
поддержки принятия решений: архитектура и разработка

числительных узлов запросы от программных агентов направляются на другие узлы (см. Рисунок 3).

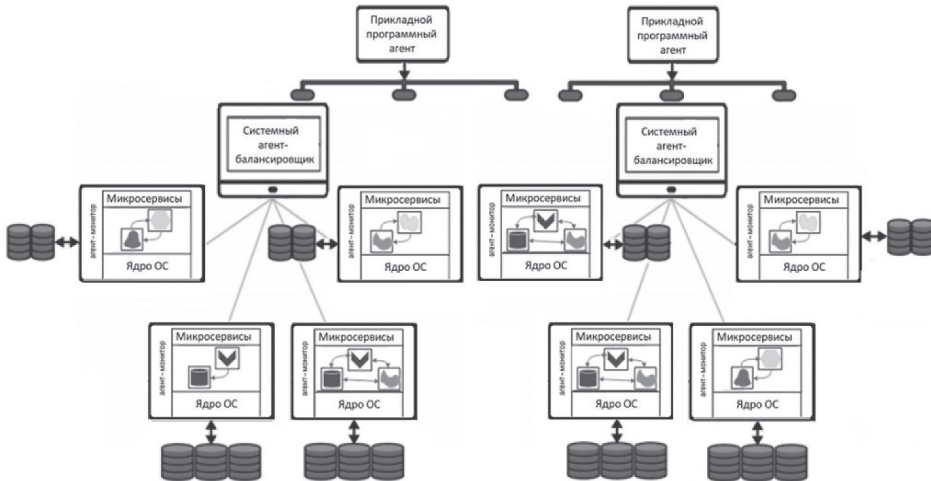


Рисунок 3. Взаимодействие программных агентов с микросервисами

Проектирование программных агентов АИСППР

Для описания функционирования программных агентов используются диаграммы состояний-переходов. Пример диаграммы с пятью состояниями, два из которых являются составными (имеющими вложенные состояния), показан на Рисунке 4.

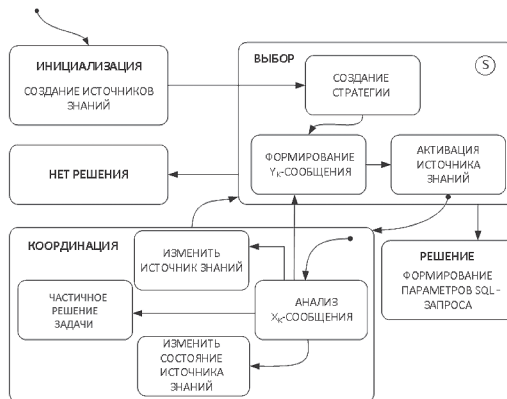


Рисунок 4. Диаграмма состояний-переходов

Из состояния «Инициализация» интеллектуальный агент переходит в составное состояние «Выбор», в котором формируется стратегия и выбирается необходимый источник знаний с учетом связей, установленных на этапе инициализации. Далее агент переходит в составное состояние «Координация», в котором активизируются новые источники знаний. Если решение найдено (состояние «Решение»), полученные в предыдущем состоянии данные используются для формирования новых запросов к базе знаний.

Пример диаграммы взаимодействия программных агентов одного из узлов АИСППР представлен на Рисунке 5.

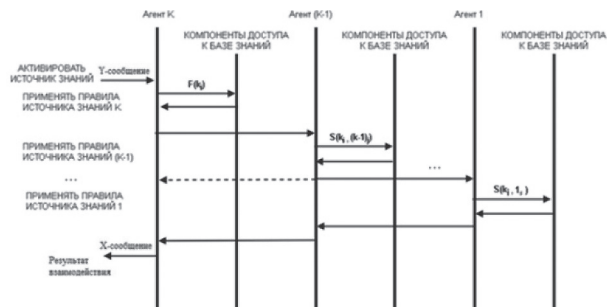


Рисунок 5. Диаграмма взаимодействия

На диаграмме взаимодействия время движется сверху вниз. Компоненты доступа к базе знаний и прикладные программные агенты представлены вертикальными линиями. Сообщения между программными агентами обозначены горизонтальными стрелками. В данном примере каждый агент с номером k взаимодействует только со своим ближайшим соседом, имеющим номер $(k - 1)$. Интеллектуальные агенты этой группы последовательно выполняют свои назначенные задачи в рамках одного процесса (без переключения контекста).

Каждый программный агент самостоятельно принимает решения и обучается, реализуя алгоритм мультиагентного глубокого обучения с подкреплением (Multi-Agent Deep Reinforcement Learning – MDRL), математической моделью которого является децентрализованный частично наблюдаемый марковский процесс принятия решений [4; 5].

В АИСППР реализована децентрализованная модель MDRL, которая позволяет использовать стандартные алгоритмы RL. В децентрализованной модели управление (выполнение) и обучение агентов происходят локально.

Согласно схеме обучения с подкреплением программные агенты наблюдают среду, предпринимают действия, меняющие состояние среды, получают сигналы подкрепления, увеличивающие или уменьшающие вероятность совершения определенного действия при тех же условиях в будущем [6–14].

В настоящее время для обучения агентов АИСППР используется мультиагентный алгоритм оптимизации проксимальной политики (Multi-Agent Proximal Policy Optimization – MAPPO), который является расширением эффективного и относительно простого в реализации алгоритма PPO (Proximal Policy Optimization) [15]. Этот алгоритм хорошо подходит для задач, которые решаются в условиях частичной информации и требуют высокой степени децентрализации и асинхронного управления. Однако, как и другие методы глубокого обучения с подкреплением, он имеет проблемы со стабильностью во время обучения в больших системах. В будущей версии АИСППР предполагается реализовать более эффективный иерархический MDRL, объединяющий алгоритмы Model-free и Model-based, что позволит множеству агентов стабильно обучаться в сложных средах с редкими вознаграждениями.

Агентно-ориентированная интеллектуальная система
поддержки принятия решений: архитектура и разработка

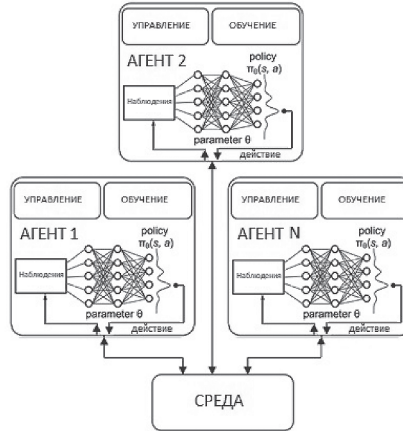


Рисунок 6. Децентрализованная модель MDRL

Оптимизация базы знаний АИСППР

Подсистема управления распределенной базой знаний АИСППР динамически оптимизирует логическую структуру РБЗ и адаптирует её к меняющимся условиям нестационарной внешней среды [16]. Критерии эффективности синтеза оптимальной структуры приведены в Таблице. В качестве основного критерия используется минимум общего времени последовательного выполнения множества транзакций. Задача оптимизации по указанному критерию относится к классу NP-трудных задач, решение которой в данном случае осуществляется с помощью метода ветвей и границ.

Математическая постановка задачи имеет следующий вид:

$$\min_{\{a_{it}, b_{tr}, b_{jm}\}} \sum_{k=1}^{K_0} \sum_{p=1}^{P_0} \sum_{kp}^{Q} \left\{ \sum_{t=1}^T v_{st}^n \left[\sum_{m=1}^M (1 - Z_{srpr}^m) (t^{acc} + t_{rm}^{transf}) + \sum_{r=1}^R Z_{sr} (t^{task} + t_{rm}^{ser} + t_{rm}^{transf}) \right] \right\},$$

где t^{task} – средняя длительность формирования одного запроса в составе транзакции;

t_{rm}^{ser} – средняя длительность выбора маршрута, установления логических соединений между узлом-клиентом (r) и узлом-сервером (m);

t_{rm}^{transf} – средняя длительность передачи раздела данных (логической записи) с узла клиента (r) на узел сервера (m) по оптимальному пути;

t^{acc} – среднее время доступа к РБЗ и поиска записей;

$\frac{Q}{kp}$ – частота использования запроса (p) транзакцией (k);

$\frac{Q}{kp}$ – бинарный признак использования запроса (p) транзакцией (k).

Критерии эффективности задачи синтеза

Переменная	Значения
a_{it}	Определяет распределение логической записи i РБЗ в раздел t : $a_{it} = 1$, если запись попала в раздел; $a_{it} = 0$ в противном случае
b_{tr}	Определяет размещение раздела t на сервере узла r распределённой системы
v_{st}^n	Определяет обновление транзакцией s , состоящей из n запросов, раздела t $v_{st}^n = 1$, если раздел обновлен
Z_{st}	Показатель P_{sr} определяет множество узлов r серверов РБЗ, фиксируемых транзакцией s : $P_{sr} = \sum_{t=1}^T \sum_{i=1}^I v_{st}^n a_{it} b_{tr}$ $Z_{st} = 1$, если $P_{sr} \geq 1$
Z_{st}^m	Показатель P_{sm} определяет множество узлов m серверов РБЗ, содержащих данные по транзакции s раздела j : $P_{sm} = \sum_{j=1}^J w_{sj}^n b_{jm}$ $Z_{st} = 1$, если $P_{st} \geq 1$;

Источник: таблица составлена авторами.

На Рисунке 7 показаны основные этапы синтеза оптимальной логической структуры РБЗ.

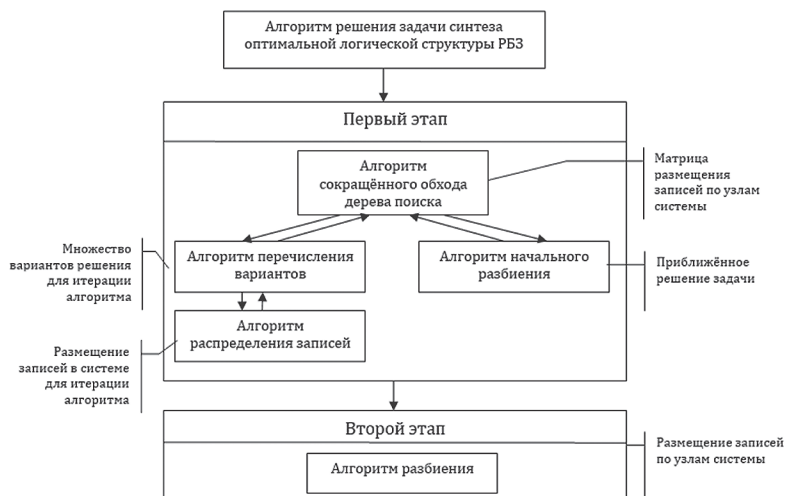


Рисунок 7. Этапы синтеза логической структуры РБЗ

В результате синтеза оптимальной логической структуры РБЗ определяется состав разделов РБЗ, формализованный в виде матрицы a_{ir} ; структура размещения разделов на вычислительных узлах АИСППР, формализованная в виде матрицы b_{ir} ; граф распределения РБЗ на вычислительных узлах b_{im} , содержащих данные по транзакциям конкретного

раздела. Оптимизация РБЗ повышает скорость доступа к знаниям АИСППР, обеспечивает возможность масштабирования системы и подключения новых источников знаний.

Подсистема обработки нечетких запросов

Автоматизация обработки фрагментов РБЗ, имеющих неопределённый характер, позволяет моделировать сложные причинные взаимосвязи и увеличивает диапазон выборки по запросу. В условиях неопределенности применение четких SQL-запросов к базе знаний во многих случаях не позволяет получить необходимую информацию, поэтому программные агенты используют нечёткие запросы.

На Рисунке 8 показана подсистема обработки запросов с нечёткими условиями, которая включает следующие модули:

- модуль добавления данных в таблицу лингвистических переменных (далее – ЛП), их терм-множеств;
- модуль формирования параметров нечеткого запроса: блоки выбора ЛП, терм-множеств, нечетких модификаторов, ослабляющих/усиливающих диапазон значений терм-множеств;
- модуль настройки параметров нечёткого запроса: блок добавления сформированного нечеткого запроса в текстовое поле со всеми выбранными параметрами;
- модуль обработки и отображения результатов запроса с нечетким условием.

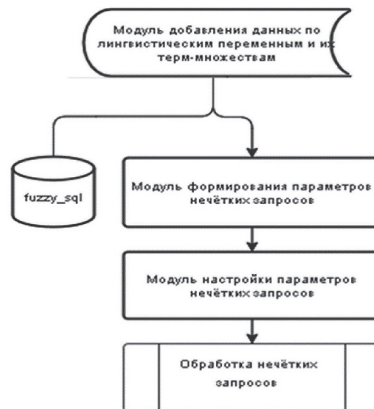


Рисунок 8. Подсистема обработки нечётких запросов

Инструмент нечетких запросов позволяет согласовать формальные критерии и неформальные требования к объектам базы знаний. В таком случае данные, не удовлетворяющие какому-то одному критерию, могут быть выбраны, если они имеют хорошие показатели по другим критериям. Такими критериями являются длительности выполнения запроса, декомпозиции запроса на подзапросы, обслуживания запроса на сервере, фиксации и передачи сообщения о результатах завершения на сервер узла администратора, снятия блокировок с записей.

Заключение

В статье рассмотрены архитектура и разработка агентно-ориентированной интеллектуальной системы поддержки принятия решений, включающей системных и прикладных про-

граммных агентов. Прикладные программные агенты решают сложные плохоформализуемые задачи и помогают конечному пользователю выбрать наилучшее решение. Прикладные программные агенты способны рационально действовать в неопределенной динамической среде, использовать для решения задач неполные и нечеткие знания, обучаться, формируя эффективную стратегию действий посредством глубокого обучения с подкреплением. Системные программные агенты повышают производительность и надежность АИСППР, реализуя эффективные специализированные методы управления информационными и вычислительными ресурсами. Повысить эффективность АИСППР позволяют также рассмотренные в статье алгоритмы синтеза оптимальной логической структуры распределенной базы знаний.

Литература

1. Рассел С., Норвиг П. Искусственный интеллект: современный подход. Т. 2. Знания и рассуждения в условиях неопределенности : пер. с англ. СПб. : Диалектика – Вильямс, 2021. 475 с. ISBN 978-5-907365-26-1.
2. Рассел С., Норвиг П. Искусственный интеллект: современный подход. Т. 3. Обучение, восприятие и действие : пер. с англ. СПб. : Диалектика, 2022. 640 с. ISBN 978-5-907365-27-8.
3. Zaytsev E.I., Khalabiya R.F., Stepanova I.V., Bunina L.V. Multi-Agent System of Knowledge Representation and Processing // Proceedings of the Fourth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’19), Springer, 2020. P. 131–141. DOI: 10.1007/978-3-030-50097-9_14. EDN YXQAZQ.
4. Самтон Р.С., Барто Э.Дж. Обучение с подкреплением / пер с англ. А.А. Слинкина. М. : ДМК Пресс, 2020. 552 с. ISBN 978-5-97060-097-9.
5. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. Reinforcement Learning : An Introduction. 2nd edition. Bradford Books. 552 p. ISBN 0262039249.
6. Barreto A., Munos R., Schaul T., Silver D. Successor features for transfer in reinforcement learning // Advances in Neural Information Processing Systems : 31st Annual Conference on Neural Information Processing Systems (NIPS 2017). Montreal : Curran Associates, 2017. Vol. 30. Pp. 4055–4065.
7. Nguyen T.T., Nguyen N.D., Nahavandi S. Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications // IEEE Transactions on Cybernetics. 2020. Vol. 50. No. 9. P. 3826–3839. DOI: 10.1109/TCYB.2020.2977374. EDN OKVIPT.
8. Zhang K., Yang Z., Başar T. Multi-agent reinforcement learning : A selective overview of theories and algorithms // Vamvoudakis K.G., Wan Y., Lewis F.L., Cansever D. (Eds) Handbook of Reinforcement Learning and Control. Series: Studies in Systems, Decision and Control. 2019. Vol. 325. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-60990-0_12
9. Jurgenson T., Avner O., Groshev E., Tamar A. Sub-goal trees a framework for goal-based reinforcement learning // Proceedings of the 37th International Conference on Machine Learning, PMLR. 2020. Vol. 119. P. 5020–5030. URL: <https://proceedings.mlr.press/v119/jurgenson20a.html> (дата обращения: 17.03.2025).
10. Zhao R., Sun X., Tresp V. Maximum entropy-regularized multi-goal reinforcement learning // Proceedings of the 36th International Conference on Machine Learning, PMLR. 2019. Vol. 97. P. 7553–7562. URL: <https://proceedings.mlr.press/v97/zhao19d.html> (дата обращения: 17.03.2025).
11. Foerster J., Farquhar G., Afouras T., Nardelli N., Whiteson S. Counterfactual Multi-Agent Policy Gradients // Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence. 2018. Vol. 32. No. 1. DOI: <https://doi.org/10.1609/aaai.v32i1.11794>

12. Vasilev B., Gupta T., Peng B., Whiteson S. Semi-On-Policy Training for Sample Efficient Multi-Agent Policy Gradients // arXiv. 2021. DOI: <https://doi.org/10.48550/arXiv.2104.13446>
13. Da Silva F.L., Nishida C.E.H., Roijers D.M., Costa A.H.R. Coordination of Electric Vehicle Charging through Multiagent Reinforcement Learning // IEEE Transactions on Smart Grid. 2020. Vol. 11. No. 3. P. 2347–2356. DOI: 10.1109/TSG.2019.2952331
14. Da Silva F.L., Reali Costa A.H. A survey on transfer learning for multiagent reinforcement learning systems // Journal of Artificial Intelligence Research. 2019. Vol. 64. DOI: <https://doi.org/10.1613/jair.1.11396>
15. Yu C., Velu A., Vinitzky E., Gao J., Wang Y., Bayen A., Wu Y. The Surprising Effectiveness of PPO in Cooperative, Multi-Agent Games // Advances in Neural Information Processing Systems 35 (NeurIPS 2022). URL: https://proceedings.nips.cc/paper_files/paper/2022/hash/9c1535a02f0ce079433344e14d910597-Abstract-Datasets_and_Benchmarks.html (Дата обращения: 17.03.2025).
16. Nurmатова Е.В., Гусев В.В., Котлиар В.В. Analysis of the features of the optimal logical structure of distributed databases // CEUR Workshop Proceedings : GRID 2018 – Selected Papers of the 8th International Conference “Distributed Computing and Grid-Technologies in Science and Education”, Dubna, 10–14 September 2018. Vol. 2267. Dubna : No Publ., 2018. P. 579–584. EDN NCITMG.

References

1. Russel S., Norvig P. (2003) *Artificial Intelligence: A Modern Approach*. 3rd edition. Pearson. 1152 p. ISBN 0136042597. (Russian edition: transl. by A.V. Sleptsov. Vol. 2. Knowledge and Reasoning under Uncertainty. St. Petersburg : Dialektika – Williams Publ., 2009. 475 p.).
2. Russel S., Norvig P. (2022) *Artificial Intelligence: A Modern Approach*. 3rd edition. Pearson. 1152 p. ISBN 0136042597. (Russian edition: transl. by A.V. Sleptsov. Vol. 3. Learning, Perception, and Action. St. Petersburg : Dialektika – Williams Publ., 2009. 640 p.).
3. Zaytsev E. I., Khalabiya R.F., Stepanova I.V., Bunina L.V. (2020) Multi-Agent System of Knowledge Representation and Processing. In: Proceedings of the Fourth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI'19), Springer. Pp. 131–141. DOI: 10.1007/978-3-030-50097-9_14
4. Sutton R.S., Barto E.J. (2018) *Reinforcement Learning : An Introduction*. 2nd edition. Bradford Books. 552 p. ISBN 0262039249. (Russian edition: transl. by A.A. Slinkin. Moscow : DMC Press, 2020. 552 p.).
5. Sutton R.S., Barto A.G. Reinforcement Learning: An Introduction. Second Edition. MIT Press, Cambridge, MA, 2018. – 552 p.
6. Barreto A., Munos R., Schaul T. Silver D. (2017) Successor features for transfer in reinforcement learning. In: *Advances in Neural Information Processing Systems : 31st Annual Conference on Neural Information Processing Systems (NIPS 2017)*. Montreal : Curran Associates, 2017. Vol. 30. Pp. 4055–4065.
7. Nguyen T.T., Nguyen N.D., Nahavandi S. (2020) Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications. In: *IEEE Transactions on Cybernetics*. Vol. 50. No. 9. Pp. 3826–3839. DOI: 10.1109/TCYB.2020.2977374
8. Zhang K., Yang Z., Başar T. (2019) Multi-agent reinforcement learning: A selective overview of theories and algorithms. In: Vamvoudakis K.G., Wan Y., Lewis F.L., Cansever D. (Eds) *Handbook of Reinforcement Learning and Control. Series: Studies in Systems, Decision and Control*. Vol. 325. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-60990-0_12
9. Jurgenson, T., Avner, O., Groshev, E., and Tamar, A. Sub-goal trees a framework for goal-based reinforcement learning. In: *Proceedings of the 37th International Conference on Machine Learning*. 2020.

Vol. 119. Pp. 5020–5030. URL: <https://proceedings.mlr.press/v119/jurgenson20a.html> (accessed 17.03.2025).

10. Zhao R., Sun X., Tresp V. (2019) Maximum entropy-regularized multi-goal reinforcement learning. In: *Proceedings of the 36th International Conference on Machine Learning, PMLR*. Vol. 97. Pp. 7553–7562. URL: <https://proceedings.mlr.press/v97/zhao19d.html> (accessed 17.03.2025).

11. Foerster J., Farquhar G., Afouras T., Nardelli N., Whiteson S. (2018) Counterfactual Multi-Agent Policy Gradients. In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. Vol. 32. No. 1. DOI: <https://doi.org/10.1609/aaai.v32i1.11794>

12. Vasilev B., Gupta T., Peng B., Whiteson S. (2021) Semi-On-Policy Training for Sample Efficient Multi-Agent Policy Gradients. *arXiv*. DOI: <https://doi.org/10.48550/arXiv.2104.13446>

13. Da Silva F.L., Nishida C.E.H., Roijers D.M., Costa A.H.R. (2020) Coordination of Electric Vehicle Charging through Multiagent Reinforcement Learning. In: *IEEE Transactions on Smart Grid*. 2020. Vol. 11. No. 3. P. 2347–2356. DOI: 10.1109/TSG.2019.2952331

14. Da Silva F.L., Reali Costa A.H. (2019) A survey on transfer learning for multiagent reinforcement learning systems. *Journal of Artificial Intelligence Research*. Vol. 64. DOI: <https://doi.org/10.1613/jair.1.11396>

15. Yu C., Velu A., Vinitsky E., Gao J., Wang Y., Bayen A., Wu Y. (2022) The Surprising Effectiveness of PPO in Cooperative, Multi-Agent Games. In: *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*. URL: https://proceedings.nips.cc/paper_files/paper/2022/hash/9c1535a02f0ce079433344e14d910597-Abstract-Datasets_and_Benchmarks.html (accessed 17.03.2025).

16. Nurmatova E.V., Gusev V.V., Kotliar V.V. (2018) Analysis of the features of the optimal logical structure of distributed databases. In: *CEUR Workshop Proceedings : GRID 2018 – Selected Papers of the 8th International Conference “Distributed Computing and Grid-Technologies in Science and Education”, Dubna, 10–14 September 2018*. Vol. 2267. Dubna : No Publ. Pp. 579–584.

Поступила в редакцию: 24.06.2025

Received: 24.06.2025

Поступила после рецензирования: 05.08.2025

Revised: 05.08.2025

Принята к публикации: 20.08.2025

Accepted: 20.08.2025