

УДК 004.896

ОСНОВНЫЕ АСПЕКТЫ ТЕОРИИ ОГРАНИЧЕННОЙ РАЦИОНАЛЬНОСТИ В МНОГОАГЕНТНЫХ СИСТЕМАХ

Стропалов А.С., Федяев О.И.

Донецкий национальный технический университет, Украина

В данной работе рассматривается значимость и роль постулата «ограниченной рациональности» при создании и моделировании интеллектуальных организаций мультиагентного типа.

Введение

В науке «искусственный интеллект» (ИИ) до недавнего времени доминировали формальные (рационалистические) подходы, основанные на классической логике, символьной парадигме и традиционных подходов в инженерии знаний. Сейчас в ИИ начинают развиваться новые направления: семиотические и нейроинтеллектуальные модели, «мягкие вычисления» и «вычислительный интеллект», многоагентные системы (МАС) и искусственные организации, которые позволяют создавать интеллектуальные системы нового поколения. Большие перспективы на этом эволюционном пути развития ИИ связывают с использованием идей и принципов синергетики в ИИ, означающих совместное действие и сотрудничество [1]. Принцип согласования индивидуальных интересов и различных точек зрения, характерный для деятельности коллективов и организаций, начинает всё более активно использоваться специалистами ИИ. Его берут за основу при проектировании взаимодействий искусственных агентов, построении МАС и интеллектуальных организаций.

Установившаяся традиция рационализма в информатике и ИИ впервые была нарушена Г.Саймоном. Им был сформулирован очень важный принцип *ограниченной рациональности* [2.1]. Согласно этому принципу, существуют пределы рациональности лиц, принимающих решения, обусловленные их индивидуальными психофизиологическими ограничениями (по скорости переработки информации, объёму памяти и т.д.). Принятие решений в организациях всегда происходит в условиях неопределённости, риска, т.е. в условиях, когда управляющий принципиально не может охватить все альтернативы и оценить все последствия принимаемых решений. Таким образом, поведение отдельного, изолированного агента принципиально не может достигать абсолютного (или даже высокого) уровня рациональности. Число альтернатив, которые следует рассматривать, обычно столь велико, а информация, необходимая для их оценки, столь обширна, что даже приближение к абсолютной рациональности невозможно. Поэтому управленческие решения в большой организации, понимаемой как сеть агентов, основаны не на полной информации и оптимизации, а на удовлетворении противоречивых критериев различных агентов (узлов сети) в условиях неопределённости.

В этой связи возникает актуальная задача анализа и изучения методологических рекомендаций, вытекающих из теории ограниченной рациональности, которые помогут в разработке интеллектуальных организаций мультиагентного типа. Основной вопрос, который рассматривается в данной работе – поиск способа для наилучшего описания ситуации принятия решений и агента, на который возлагается принятие решений.

1 Рационалистическая парадигма в информатике и ИИ

Эта парадигма до недавнего времени была основополагающей не только в информатике и ИИ, но и в смежных научных дисциплинах: теории управления, теории принятия решения, когнитивной психологии. Верным спутником рационализма выступает принцип редукционизма – сведения реального сложного явления к сильно упрощённой модели. Считается, что решение агента является рациональным, если варианты его выбора, способны дать максимально выгодный результат с учётом всех ограничений.

Предпосылками рациональности решения являются [7]:

- чёткость проблемы;
- ориентированность на достижение цели;
- знание всех вариантов и их последствия известны;
- чёткость преимуществ, т.е. преимущества ясны;
- постоянство преимуществ;
- отсутствие временных или материальных ограничений;
- максимальная отдача, т.е. окончательный выбор максимизирует преимущества данного выбора.

Полностью рациональный агент (англ. rational agent) – это агент, действующий оптимальным образом для достижения наилучшего ожидаемого результата. Данный термин является одним из фундаментальных в экономике, теории принятия решений и ИИ [4]. Понятие «рациональный агент» пришло в ИИ из экономики, и совершило в нём настоящую революцию, объединив разрозненные направления исследований.

Поскольку в реальном мире абсолютная (полная) рациональность практически недостижима, ожидается, что агент в ходе принятия решения будет сдерживаться по принципу ограниченной рациональности.

2 Основные принципы теории ограниченной рациональности

Признанным основоположником обобщенной модели экономического поведения, которая получила название теории ограниченной рациональности (bounded rationality) считается нобелевский лауреат, американский экономист, профессор психологии и информатики Г. Саймон [2].

Его теория исходила из следующих наблюдений:

1. Абсолютная рациональность требует полного знания всех принимаемых решений. В реальной действительности это не достижимо.
2. Существуют принципиальные ограничения по возможности прогнозирования последствий принимаемых решений (как будущих событий).
3. Абсолютная рациональность требует выбора из всех возможных вариантов поведения. На практике лишь небольшое число возможных вариантов может быть принято во внимание.

Г. Саймон со своими сотрудниками провёл ряд эмпирических исследований того, как реально протекает процесс принятия решений в фирмах и на их основе разработал нормативные алгоритмы принятия «правильных» решений [3].

Для того чтобы максимизировать полезность или прибыль, экономическому субъекту просто не хватает счётных способностей. Проблема субъекта состоит не столько в том, что у него мало информации, сколько в том, что её слишком много относительно возможностей её обработки. Процесс принятия решений в модели Г. Саймона можно описать двумя главными понятиями – поиска и принятия удовлетворительного варианта.

По мнению Г. Саймона, не может быть всеобъемлющей функции полезности, которая позволила бы сравнить разнородные альтернативы. Эта функция, по его мнению, имеет всего два $\{0, 1\}$ или три $\{-1, 0, 1\}$ значения, где 1 обозначает удовлетворительный вариант, -1 – неудовлетворительный, а 0 – безразличный.

В результате хозяйственный субъект осуществляет поиск вариантов решения задачи до тех пор, пока не будет найдено первое приемлемое решение, т.е. поступает так, как показано на рис.1.

Приемлемость или неприемлемость варианта каждый определяет для себя сам. Г. Саймон характеризует этот процесс с помощью заимствованной из психологической науки категории «уровня притязаний». Концепция уровня притязаний предполагает, что в каждый момент времени у человека есть некоторое представление о том, на что он может (имеет право) рассчитывать. Уровень притязаний не является застывшим, планка всё время сдвигается в зависимости от результатов последнего шага. Если он был успешным, уровень притязаний поднимается вверх – человек ставит себе более высокую цель. В случае неудачи уровень притязаний опускается, поскольку человек начинает более критично

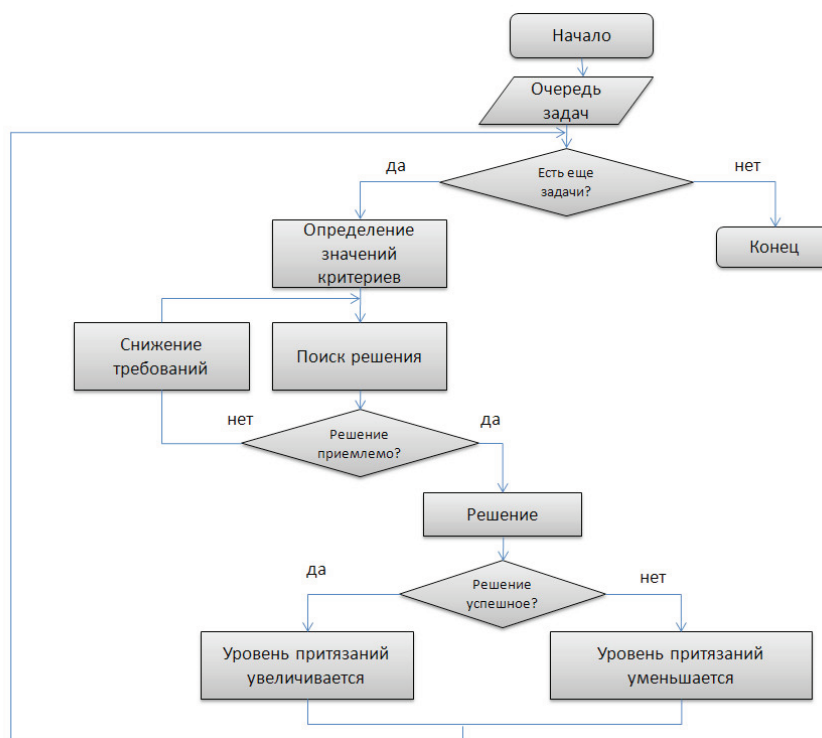


Рисунок 1. Алгоритм работы модели, построенной на концепции ограниченной рациональности

относиться к своим способностям. Вариант считается удовлетворительным, если он позволяет человеку преодолеть планку – уровень притязаний.

Можно заметить, что выбор удовлетворительного варианта требует от экономического субъекта гораздо меньшей информированности и счётного искусства. Ему уже не надо иметь точную информацию об исходе данного варианта и сравнивать его с исходами альтернативных вариантов в рамках общей функции полезности. Достаточно интуитивного представления о том, что данный вариант выше или ниже приемлемого уровня. При этом сравнивать варианты между собой вообще нет необходимости [3].

Концепция ограниченной рациональности является единственной в экономической теории формальной моделью человеческого поведения, альтернативной максимизации полезности и прибыли, хотя её применение на практике требует сложных формул и вычислений. В настоящее время модель «ограниченной рациональности» успешно применяется в нормативных рекомендациях и в компьютерных программах.

3 Использование теории ограниченной рациональности в многоагентных системах

Концепция ограниченной рациональности нашла широкое применение в теории многоагентных систем. МАС состоит из нескольких взаимодействующих агентов, которые могут иметь разную архитектуру. В зависимости от типа агентов МАС могут быть использованы для решения многих проблем, которые сложно или невозможно решить с помощью одного агента или монолитной системы. На рис. 2 показана одна из архитектур интеллектуального агента. Она включает 4 компонента. Наиболее существенное различие наблюдается между обучающим компонентом, который отвечает за внесение усовершенствований, и производительным компонентом, который обеспечивает выбор внешних действий [5].

Производительный компонент получает воспринимаемую из внешней среды информацию и принимает решение об ответных действиях. Обучающий компонент использует информацию обратной связи от критика с оценкой того, как действует агент, и определяет, каким образом должен быть модифицирован производительный компонент для того, чтобы он успешнее действовал в будущем.

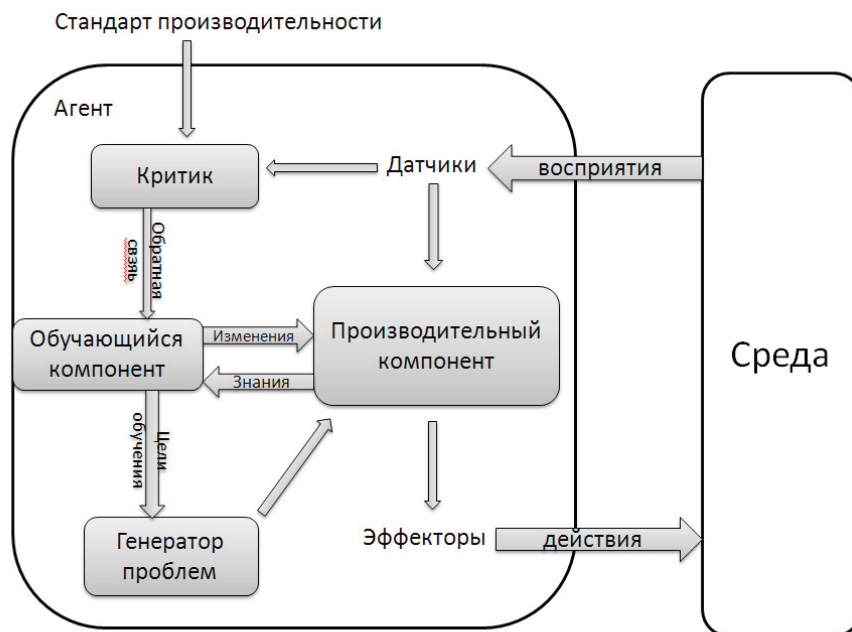


Рисунок 2. Архитектура интеллектуального агента

Критик сообщает обучающему компоненту, насколько хорошо действует агент с учётом установленного стандарта производительности. Критик необходим, поскольку сами результаты восприятия не дают никаких указаний на то, успешно ли действует агент.

Последним компонентом агента является генератор проблем. Его задача состоит в том, чтобы предлагать действия, которые должны привести к получению нового и информативного опыта [6].

Выводы

На данном этапе научных исследований авторы основное внимание уделили анализу концепции ограниченной рациональности в поведении людей, рассматриваемой через призму искусственного интеллекта.

По мнению специалистов в этой области классическая теория поведения человека, основанная на принципе полной рациональности при принятии человеком решений, далека от реальности. К факторам, ограничивающим рациональность агентов, относятся:

- недостаточная мотивация;
- локальный характер располагаемой информации (ограничения на ресурсы);
- принципиальная неполнота и неточность информации, получаемых от других агентов;
- случайные воздействия среды.

Любой агент в организации всегда находится во взаимодействии с другими агентами, ведёт переговоры, пытается создать коалиции и т. п.; поэтому его восприятие субъективно и отрывочно, а знания, необходимые для решения задачи, нечётки, неточны и ограничены [1]. Поэтому сейчас в области МАС доминирует тезис о целесообразности в таких условиях использовать коллективную выработку решений агентами в ходе их совместной деятельности. На основе этого тезиса построена новая поведенческая теория организаций. Она основана на признании ограниченности чистого рационализма и привлечении идей относительной рациональности в поведении людей, интеллектуализации и биологизации организаций. Поэтому, если МАС должна работать в сложной, трансформируемой, динамической среде, характеризующейся высоким уровнем неопределённости, то её необходимо строить на основе следующих новых принципах:

1. Вместо жёсткой «древовидной» иерархии целесообразно использовать «плоские» сетевые структуры, в которых агенты рассматриваются не как шестерёнки, а как узлы сети,

осознающие цели МАС и реализующие направленные взаимодействия с другими узлами [1].

2. Предоставлять интеллектуальным агентам не узкую, а более широкую специализацию, обеспечивающую решение крупных задач.
3. Переходить от единоначалия к кооперации и координации действий.

Такой подход к построению многоагентных моделей неоднородных динамических систем с распределённым интеллектом позволит более реалистично смоделировать работу всей системы, а также различные аспекты поведения людей-участников таких систем.

Литература

- [1] Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика.- М.: Эдиториал УРСС, 2002. – 352 с.
- [2] Саймон Г. Науки об искусственном: Пер. с англ. – М.: Мир, 1972.
- [3] История экономических учений/Под ред. В. Автономова, О.Ананьина, Н. Макашева: Учеб. пособие. – М: ИНФРА-М, 2002. - 784 с. – (Серия «Высшее образование»).
- [4] Рациональный агент. Материал из Википедии – свободной энциклопедии. Электронный ресурс. Режим доступа: http://ru.wikipedia.org/wiki/Рациональный_агент
- [5] Многоагентная система. Материал из Википедии – свободной энциклопедии. Электронный ресурс. Режим доступа: http://ru.wikipedia.org/wiki/Многоагентная_система
- [6] Интеллектуальные агенты. Материал из Википедии – свободной энциклопедии. Электронный ресурс. Режим доступа: <http://www.itfru.ru/index.php/intellectual-agents>
- [7] Принятие решений. Материал из Википедии – свободной энциклопедии. Электронный ресурс. Режим доступа: http://club-energy.ru/8_2.php