

## ФОРМИРОВАНИЕ ОПТИМАЛЬНОГО МАРШРУТА БОЛЬШИХ ГРУПП ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ

Максимов Д. Ю.<sup>1</sup>

(ФГБУН Институт проблем управления  
им. В.А. Трапезникова РАН, Москва)

*В подходе Artificial General Intelligence (Universal AI) интеллект рассматривается как информационный процессор, потребляющий и выдающий информацию, которой и определяется поведение системы. В рамках этого подхода М. Hutter получил способ выбора оптимальной траектории агента в абстрактной среде. Однако этот способ требовал численных оценок вознаграждения при том или ином движении, способ получения которых оставался открытым. Также этот метод не подходит для оценок движения группы агентов. В настоящей работе оценки вознаграждений предоставляются самой средой и предложен способ выбора траектории, который применим для группы агентов. Параллельное выполнение группой интеллектуальных агентов ряда задач, представляется тензорным произведением соответствующих процессов в категории игр Конвея, которая сопоставлена среде и движениям агентов. Оптимальный маршрут группы определяется как игра с наибольшим суммарным выигрышем в этой категории. Выигрыш представлен степенью определенности (видимости) цели агента, т.е. некоторым множеством, а не числом. Доказано, что такое определение выигрыша может быть использовано в категорной конструкции для игр Конвея. В этой категории тензорное произведение является операцией линейной логики. Также линейная логика применяется для выбора целей, которые система может достичь, из всего множества видимых целей. Для этого все множество целей представляется в виде решетки, на которой задана структура линейной логики. Решетка целей в этом случае становится множеством истинностных значений логики. Целям, которые достигаются параллельно, также соответствует тензорное произведение (как и параллельным процессам в среде), но только теперь это произведение элементов решетки целей. Цели выбираются по наибольшей степени истинности того элемента решетки, который соответствует их параллельному достижению. В результате получена формула для оценки наиболее выигрышного маршрута в абстрактной среде для группы агентов.*

**Ключевые слова:** интеллектуальные агенты, выбор маршрута, решетка целей, игровая семантика.

---

<sup>1</sup> Дмитрий Юрьевич Максимов, науч. сотр. (dmtax@inbox.ru).

## 1. Введение

Важными задачами при описании движения групп агентов являются задача целераспределения и задача определения оптимального маршрута. В случае когда эти агенты предполагаются интеллектуальными, они должны самостоятельно решать эти задачи. Данная работа посвящена тому, чтобы представить решение этих задач агентами в абстрактной среде только на основании информации, которая предоставляется средой, и информации о целевом предназначении системы, которая заложена изначально и представлена структурами на множестве целей системы.

В подходе Artificial General Intelligence (AGI) или Universal Artificial Intelligence (UAI) искусственный интеллект представляется как информационный процессор, потребляющий и выдающий информацию, и исследования в этом направлении фокусируются на системах, *действующих* разумно [10, 11, 12, 14, 15, 20, 23, 24, 25]. В рамках AGI предложен способ выбора оптимальной траектории движения в абстрактной среде наиболее интеллектуального агента в смысле некоторой меры интеллектуальности [16]. Эта модель основана на вероятностном моделировании среды, определении следующего движения такой системы на основании предыдущего опыта, числовой оценке вознаграждения позиций системы и максимизации предполагаемого будущего вознаграждения вдоль траектории. Однако способ получения этой числовой оценки отсутствует. Также нет моделей, представляющих поведение групп таких агентов.

В [18] также рассматриваются вероятностные свойства траектории движущегося агента. Агент находит маршрут, близкий к оптимальному, на решетке, состоящей из клеток с разной непроеходимостью. С помощью имитационного моделирования изучается распределение времени выхода агента в конечную точку для случайных ландшафтов разного вида. Для описания движения групп агентов в [3, 4] описывается клеточный автомат, моделирующий движение агентов, и исследуются его особенности в связи с типом ландшафта, по которому перемещаются агенты.

В [5, 6, 7] продемонстрировано, что достаточно только наличия структуры (решеточной или еще и моноида, что определяет линейную логику) на множестве задач [6, 7] или целей [5] системы для того, чтобы система вела себя вполне разумно и даже похоже на поведение, например, колонии муравьев [7]. Вероятностного моделирования среды, в отличие от [16] и [18], не предполагается.

В этой статье предлагается развитие подхода [5, 6, 7] исходя из идеи, что параллельное выполнение процессов достижения разных целей и разными агентами в абстрактной среде можно представить как тензорное произведение в линейной логике. Представление параллельно выполняемых процессов тензорным произведением уже использовалось в [21], но в этой работе процессы достижения в среде системой интеллектуальных агентов своих целей формально описываются как игра с вознаграждениями позиций. Такие игры моделируются в категории игр Конвея, определенной в [22]. Но, в отличие от [22], вознаграждение позиций представляется множествами, описывающими информацию о целях (представляющими степень их определенности). Таким образом, вознаграждения предоставляются средой и являются элементами решетки множеств, а не числами, как в [16]. Доказано, что такое изменение определения вознаграждения соответствует конструкции [22] и не меняет выводов этой работы.

Кроме того, так же как и в [7], структуры решетки и линейной логики на множествах целей агентов и системы используются для определения приоритетов различных параллельных процессов в такой категории игр. Эти структуры задаются исходя из представлений о назначении системы, в отличие от аналогичных структур в категории игр, которые задаются средой. Выбираются параллельные процессы достижения таких целей, переход на достижение которых имеет наибольшую оценку в решетке целей системы. Решетки целей агентов формируются агентами как их всевозможные намерения и могут использоваться для уточнения выбора маршрута в системе в случае, если однозначной оценки выигрыша маршрута получить не удается.

В результате получена формула, которая определяет маршрут группы и выбор ею целей с наибольшим суммарным вознаграждением. Эта формула отражает движение по пути наибольшей определенности или наибольшей видимости целей, вдоль которого средой предоставляется наибольшая информация о совокупности целей. Выбор же этой совокупности из всего множества видимых целей определяется из представлений о назначении группы, которые определяют логическую структуру на множестве ее задач и, следовательно, оценки той или иной совокупности целей.

Работа построена следующим образом: в разделе 2 как можно проще вводится необходимый математический аппарат. Доказываются теоремы о соответствии используемого определения вознаграждений конструкции из [22]. В разделе 3 представлены способы выбора целей и оптимального маршрута группы. Также обсуждаются трудности такого подхода и способы их преодоления. Раздел 4 заключает работу.

## **2. Используемый математический аппарат**

### **2.1. РЕШЕТКИ [1]**

Определение 1. Частично-упорядоченное множество – это множество, на котором определено бинарное отношение  $x \leq y$ , удовлетворяющее для всех  $x, y, z$  следующим условиям:

- $x \leq x$  (рефлексивность);
- если  $x \leq y$  и  $y \leq x$ , то  $x = y$  (антисимметричность);
- если  $x \leq y$  и  $y \leq z$ , то  $x \leq z$  (транзитивность).

Это означает, что, в отличие от линейно-упорядоченного множества, не все элементы сравнимы между собой.

Определение 2. **Верхней гранью** подмножества  $X$  частично-упорядоченного множества  $P$  называется такой элемент  $a \in P$ , который содержит все  $x \in X$ .

Определение 3. **Точная верхняя грань** подмножества  $X$  ( $\sup X$ ) – это такая его верхняя грань, которая содержится в любой другой его верхней грани.

Понятие *точной нижней грани* ( $\inf X$ ) определяется двойственно (т.е. это наибольший элемент  $b \in P$ , содержащийся во всех  $x \in X$ ).

**Определение 4. Решетка** – это частично-упорядоченное множество, имеющее для любых двух элементов их точную верхнюю грань (или **объединение**)  $\vee$  и точную нижнюю грань (или **пересечение**  $\wedge$ ).

**Определение 5. Полурешетка** – это частично-упорядоченное множество, в которой для любых двух элементов определена только одна операция – объединение в верхней полурешетке или пересечение в нижней.

**Определение 6. Полная решетка** – это решетка, в которой любые два подмножества имеют объединение и пересечение.

Это означает, что в непустой полной решетке есть наибольший « $\top$ » и наименьший « $0$ » элементы. На диаграмме решетки (например, как на рис. 1) чем больше элемент (т.е. вершина, узел диаграммы решетки), тем выше он расположен, и сравнимые между собой элементы лежат на одном пути из  $0$  в  $\top$ .

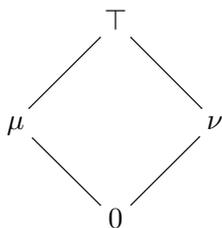


Рис. 1. Пример диаграммы решетки

Объединение элементов на диаграмме – ближайший наибольший элемент для обоих, пересечение – ближайший наименьший для обоих. Заметим, что любая конечная решетка является полной.

**Определение 7. Образующими решетки (ее генераторами)** называются те элементы, из которых путем применения операций объединения и пересечения получаются все остальные элементы.

**Определение 8.** В решетках для любых элементов  $a, b$  может быть определена импликация  $c = a \Rightarrow b$  как наибольший элемент, который пересекается с  $a$  так же, как  $b$ :  $c \wedge a = b \wedge a$ .

**Определение 9.** Решетка, имеющая импликацию, называется **брауэровой решеткой**. В такой решетке импликация  $\neg a = a \Rightarrow 0$  называется псевдодополнением  $a$ .

В брауэровых решетках выполняются законы дистрибутивности для объединения и пересечения. Все конечные дистрибутивные решетки являются брауэровыми.

## 2.2. ЛИНЕЙНАЯ ЛОГИКА

Исторически первая и наиболее простая семантика линейной логики [13] — это семантика фазового пространства. В этом случае на решетке задано умножение элементов  $\cdot$ , которое определяет дополнительные к решеточным логические операции.

**Определение 10. Фазовым пространством** называется пара  $(M, \perp)$ , где  $M$  — коммутативный (мультипликативный) моноид  $(B, \cdot, 1)$ , в котором несущее множество  $B$  является решеткой (решеткой подмножеств моноида), а  $\perp \subset B$  — произвольный элемент решетки (т.е. подмножество моноида), который принимается за *false*.

В линейной логике элемент *false* отличается от  $\mathbf{0}$  (т.е. наименьшего элемента решетки) в общем случае (в отличие от классической или интуиционистской логик). В линейной логике определены следующие мультипликативные (т.е. определяющиеся моноидальным умножением) операции:

**Определение 11. Линейная импликация**  $\multimap$  на элементах решетки  $X, Y \subseteq M$  производится следующим образом:  $X \multimap Y = \{z \mid z \cdot x \in Y, \forall x \in X\}$ .

**Определение 12. Дуальный элемент**  $X^\perp = X \multimap \perp$  — линейное отрицание  $\bar{X}$ .

Определение 13. **Ко-тензорное произведение**, которое связано с **линейной импликацией**  $\multimap$  на элементах решетки  $X, Y \subseteq M$ , производится следующим образом:  $X \wp Y = X^\perp \multimap Y = (X^\perp \cdot Y^\perp)^\perp$ .

Элемент *false* является нейтральным элементом операции  $\wp$ . Линейная логика – ресурсозависима: в ней посылка (ресурс) может быть употреблена только один раз. Поэтому импликация  $X \multimap Y$  интерпретируется как потребление  $X$  и получение  $Y$  [13].

Определение 14. *Определена также мультипликативная конъюнкция  $\otimes$ , или **тензорное произведение**, следующим образом:  $X \otimes Y = (X \multimap Y^\perp)^\perp = (X \cdot Y)^\perp = (X^\perp \wp Y^\perp)^\perp$ .*

$I$  – нейтральный элемент мультипликативной конъюнкции, дуальный  $\perp$ .

Определение 15. *Здесь везде  $X, Y$  – **факты**, т.е. такие подмножества  $M$ , что  $X = X^{\perp\perp}$  или, эквивалентно,  $X = Y^\perp$  для некоторого  $Y \subset M$ .*

Важно отметить, что консеквентом импликации может быть только факт. Истинными (valid) считаются факты, которые включают в себя моноидальную единицу 1. Например,  $I = \perp^\perp = \{1\}^{\perp\perp}$

Определение 16. *На фактах определены также решеточные, дуальные друг другу, операции аддитивной конъюнкции  $\&$  и аддитивной дизъюнкции  $\oplus$  (по сути это **пересечение** и **объединение** элементов решетки):  $X \& Y = X \wedge Y$ ;  $X \oplus Y = (X \vee Y)^{\perp\perp} = (X^\perp \wedge Y^\perp)^\perp$ .*

Их нейтральные элементы –  $\top$  для  $\oplus$  и  $\mathbf{0}$  для  $\&$  – также связаны отношением дуальности:  $\top^\perp = \mathbf{0}$ ;  $\mathbf{0}^\perp = \top$ , где  $\top = M = \emptyset^\perp$ ;  $\mathbf{0} = M^\perp = \emptyset^{\perp\perp}$ . Дуальность операций понимается как в теории множеств:  $\vee^\perp = \wedge$ ,  $\wedge^\perp = \vee$ , т.е., например,  $(a \vee b)^\perp = a^\perp \wedge b^\perp$ , где дуальность означает дополнение.

Нам потребуется также определение **экспоненциала** !. Для этого в фазовом пространстве выделяются два взаимно дуальных множества **открытых**  $Op$  и **замкнутых**  $Cf$  фактов.

Определение 17. *Множество **замкнутых фактов**  $\mathcal{F}$  обладает следующими свойствами:*

- $\mathcal{F}$  замкнуто относительно аддитивной конъюнкции  $\&$ .
- $\mathcal{F}$  конечно замкнуто относительно мультипликативной дизъюнкции  $\mathcal{N}$  (т.е. мультипликативная дизъюнкция конечного числа слагаемых содержится в  $\mathcal{F}$ ).
- $\perp$  — наименьший факт (относительно включения) в  $\mathcal{F}$ .
- для всех  $X \in \mathcal{F}$   $X \mathcal{N} X = X$ .

Определение 18. Множество **открытых фактов**  $\mathcal{G}$ , которое дуально  $\mathcal{F}$ , обладает дуальными свойствами:

- $\mathcal{G}$  замкнуто относительно аддитивной дизъюнкции  $\oplus$ .
- $\mathcal{G}$  конечно замкнуто относительно мультипликативной конъюнкции  $\otimes$ .
- $I$  — наибольший факт (относительно включения) в  $\mathcal{G}$ .
- для всех  $X \in \mathcal{G}$   $X \otimes X = X$ .

Наименьшим открытым фактом является  $\mathbf{0}$ . Соответственно, наибольшим замкнутым —  $\top$ . В такой структуре экспоненциалы определяются так:

- $!X$  — наибольший открытый факт, включенный в  $X$ ;
- $?X$  — наименьший замкнутый факт, содержащий в себе  $X$ .

Легко заметить, что мы имеем две полурешетки — нижнюю  $\mathcal{F}$  и верхнюю  $\mathcal{G}$ .

В этом определении фазового пространства первичным является определение умножения элементов, которое задает все мультипликативные операции. В результате получается линейная логика с определенными свойствами. Но на практике можно исходить из определения классов открытых и замкнутых фактов. Эти классы определяются требованиями задачи. Например, поскольку консеквентом импликации может быть только факт, естественно

потребовать минимальности количества не-фактов. Тогда требование выполнения свойств линейной логики приводит к определению произведений всех элементов решетки [7, 21].

Для этого при выборе элементов, дуальных к не-фактам, следует использовать свойство

$$1) X \subset X^{\perp\perp}.$$

Для определения умножения можно использовать еще следующие требования:

$$2) X \multimap X \geq I;$$

$$3) I \multimap X = X;$$

$$4) X \multimap (B \& C) = (X \multimap B) \& (X \multimap C);$$

$$5) X \multimap B^{\perp} = X \multimap C^{\perp}, \text{ если } B^{\perp} = C^{\perp}.$$

Последнее условие означает, что один и тот же факт может быть дуалом как факта, так и не-факта.

$$6) !X \multimap \mathbf{0} = \neg !X \text{ [19].}$$

Здесь  $\neg !X$  — решеточная операция отрицания, т.е. наибольший элемент решетки, не пересекающийся с  $!X$ .

В результате получается недоопределенная система уравнений для определения произведений элементов решетки, поскольку этих свойств не хватает для однозначного вычисления. Однако это приводит к большей гибкости при выборе варианта поведения (см. разд. 3.3).

### 2.3. ИГРЫ КОНВЕЯ

*Определение 19. Игра (game) Конвея определяется в [22] как корневой граф, в котором вершины  $V$  являются позициями игры, дуги  $E \subset V \times V$  — ее движениями и каждой дуге присвоена полярность  $\pm 1$  в зависимости от того, движение это игрока или оппонента.*

**Определение 20. Траектория (play)** — это путь из корня  $*$  графа. Путь альтернирован, если полярности на соседних ребрах чередуются.

**Определение 21. Стратегия** определяется как непустое множество альтернированных путей четной длины, которые начинаются движением оппонента, замкнуто по префиксу четной длины (т.е. префикс четной длины также принадлежит стратегии) и детерминировано. Детерминированность означает, что два разных пути с общим префиксом, которые различаются двумя движениями, должны совпадать.

**Определение 22. Дуальная игра (play)**  $X^\perp$  определяется как игра, в которой полярности у всех дуг обращены.

**Определение 23. Тензорное произведение**  $X \otimes Y$  двух игр Конвея  $X = (V_X, E_X)$  и  $Y = (V_Y, E_Y)$  определяется как асинхронное произведение графов двух этих игр: позициями произведения являются пары  $(x, y)$ , которые обозначаются  $x \otimes y$ , с корнем  $*_{X \otimes Y} = *_X \otimes *_Y$ , так что  $V_{X \otimes Y} = V_X \times V_Y$ ; движения (ребра)  $E_X \otimes E_Y = \{x \otimes y \rightarrow \begin{cases} z \otimes y, & \text{если } x \rightarrow z \in E_X, \\ x \otimes z, & \text{если } y \rightarrow z \in E_Y; \end{cases}\}$  и полярность движения в  $X \otimes Y$  наследуется из соответствующего движения в  $X$  или  $Y$ .

Тензорное произведение таких игровых процессов интерпретируется как их параллельное выполнение [22]. Можно также определить тензорное произведение стратегий [22], которые являются морфизмами в категории таких игр, но оно нам не потребуется и здесь не приводится.

В категории таких игр моделируется линейная логика (обобщенная). Объектами этой категории являются игры Конвея, а морфизмами  $X \rightarrow Y$  — стратегии в  $X^\perp \wp Y = X \multimap Y$ . Следует заметить, что на графах игр операция  $\wp$  совпадает с  $\otimes$  в этом определении, так что в [22] они даже не различаются. Для стратегий определена их композиция [22], но это определение нам не потребуется и здесь не приводится.

Игра Конвея с вознаграждением [22] — это игра с дополнительным весом в каждой вершине  $\{1, 1/2, 0\}$ , в зависимости от

выигрышности позиции: позиция считается выигрышной, если ее вес 1 или  $1/2$ . Для тензорного произведения игр эти веса подчиняются булевским правилам конъюнкции и импликации. Таким образом, игра  $X \otimes Y$  Конвея с вознаграждением определяется как игра  $X \otimes Y$  с вознаграждениями  $k_{X \otimes Y}(x \otimes y) = k_X(x) \wedge k_Y(y)$  и игра  $X \multimap Y$  имеет вознаграждения  $k_{X \multimap Y}(x \multimap y) = k_X(x) \Rightarrow k_Y(y)$ .

*Определение 24.* Стратегия является **выигрышной**, если каждая траектория (play) заканчивается в выигрышной позиции.

В категории игр Конвея с вознаграждениями морфизмами являются выигрышные стратегии.

Докажем, что категорная конструкция не меняется, если числовые веса заменяются множествами, которые образуют брауэрову решетку, булевская конъюнкция — решеточной, а веса в линейной импликации выражается через решеточные операции. Чем большее множество связано с позицией, тем она выигрышной. Выигрышными считаются все позиции с весами, отличными от 0 (пустого множества). Предполагается, что существует некоторое универсальное множество, содержащее все остальные, которое является наибольшим элементом  $\top$  решетки. В таком случае оценки выигрыша для тензорного произведения и импликации представляются следующим образом:  $k_{X \otimes Y}(x \otimes y) = k_X(x) \otimes k_Y(y) \equiv k_X(x) \wedge k_Y(y)$ ,  $k_{X \multimap Y}(x \multimap y) = k_X(x) \multimap k_Y(y) \equiv k_{X^\perp}(x) \vee k_Y(y) = \neg k_X(x) \vee k_Y(y)$ , где  $\neg k$  — псевдодополнение множества  $k$ . Таким образом, в дуальной игре  $X^\perp$  с вознаграждениями выигрыш позиции определяется как  $\neg k$ , где  $k$  — выигрыш в игре  $X$ .

В получающейся брауэровой решетке  $M$  предполагается, что не существует такого элемента  $k \neq 0$ , что  $k$  получается как пересечение всех элементов  $x \neq 0$  решетки  $M$ :  $k = \bigwedge_{x \in M} x$ . Это означает, что  $\neg k \neq 0$ ,  $\forall k \neq \top$ . В разделе 3.2 выигрыш  $k = \top$  никогда не возможен, поэтому примем, что  $\neg k \neq 0$ ,  $\forall k \in M$ .

В этом случае так определенные выигрышные стратегии образуют композиции, как требуется для морфизмов:

**Предложение 1.** Стратегия  $\rho \circ \sigma : X \multimap Z$  является выигрышной, когда две стратегии  $\sigma : X \multimap Y$  и  $\rho : Y \multimap Z$  – выигрышные.

**Доказательство.** Стратегии в категории игр Конвея подчиняются правилу композиции [22]. Поэтому достаточно проверить условие выигрышности. Но это условие выполняется тривиально в силу  $k_{X \multimap Y} = \neg k_X \vee k_Y$  и  $\neg k \neq 0, \forall k \in M$ .

Докажем также, что введенное определение для выигрышей линейной импликации законно:

**Предложение 2.** Категория игр Конвея с вознаграждениями, в которой выигрыши позиций принимают значения в брауэровой решетке, является симметричной моноидально-замкнутой.

**Доказательство.** Категория игр Конвея является симметричной моноидально-замкнутой [22]. Поэтому достаточно проверить условие  $(k_X \otimes k_Y) \multimap k_Z = k_X \multimap (k_Y \multimap k_Z)$ . Но это условие выполняется в силу выполнимости в брауэровых решетках формулы  $\neg(k_X \wedge k_Y) \vee k_Z = \neg k_X \vee (\neg k_Y \vee k_Z)$ .

Таким образом, для введенного определения морфизмов как выигрышных стратегий с выигрышем, оценивающимся множеством, категорная конструкция [22] не меняется.

### 3. Выбор маршрута

#### 3.1. ВЫБОР ЦЕЛЕЙ

Отождествим цели системы с процессами или задачами их достижения и далее будем использовать эти понятия как синонимы. Примем, что наиболее желательное поведение системы состоит в достижении всех целей, а неприемлемое поведение состоит в бездействии. В таком случае ценности различных вариантов поведения системы, т.е. ценности разных наборов возможных процессов, образуют решетку целей  $M$ . Чем выше в диаграмме решетки расположен набор выполняемых процессов, тем такое поведение более ценно, однако не все элементы решетки сравнимы между собой (рис. 2).

В этой решетке пересечения означают подпроцессы, кото-

рые входят в различные задачи. Например, элемент  $c_1$  обозначает движение роботов-дворников при выносе  $E$  и поиске  $X_1$  мусора в [7]. Объединения генераторов соответствуют объединению различных процессов, которые не выполняются параллельно.

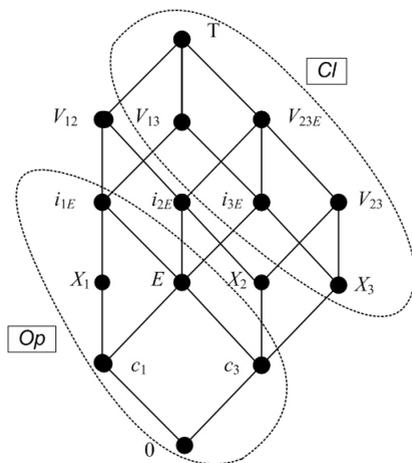


Рис. 2. Пример решетки задач системы. Элементы  $X_i$  и  $E$  являются генераторами, т.е. отдельными процессами

Определим на решетке моноидальное умножение элементов и классы открытых  $Op$  и замкнутых  $Cl$  фактов (на рис. 2 обведены линиями), т.е. структуру линейной логики в соответствии с правилами 1–6 раздела 2.2. В этом случае параллельно выполняемым процессам достижения разных целей будет соответствовать тензорное произведение соответствующих элементов решетки. Приоритеты различных параллельно выполняемых процессов и, соответственно, выбор наиболее предпочтительного варианта определяются из сравнения оценок соответствующих тензорных произведений (произведение элементов решетки – это тоже элемент решетки, и чем выше он лежит на диаграмме решетки, тем он более ценен).

Тогда линейную импликацию  $X \multimap Y$  на этой решетке можно рассматривать как переход от выполнения одного набора

процессов  $X$  к другому  $Y$ . Значения таких импликаций можно сравнивать так же, как тензорные произведения и так же выбирать наиболее предпочтительный вариант перехода в случае необходимости нового выбора целей. Эта импликация выражается через умножение элементов и вычисление дуальных так:  $X \multimap Y = (X \cdot Y^\perp)^\perp = (X \otimes Y^\perp)^\perp$  (раздел 2.2). Поскольку умножения элементов вычисляются неоднозначно (соответствующая система линейных уравнений недоопределена), то в общем случае такие импликации и тензорные произведения только на основании этой системы уравнений могут иметь несколько значений, что может быть использовано для устранения неопределенности при сравнении несравнимых значений в решетке (раздел 3.3).

Исходно предполагается, что система  $l$  агентов движется в некоторой абстрактной среде, первоначально не видя целей. Эта среда никак не формализована, а просто представляет для агентов источник информации (которую агенты оценивают сами, по своим моделям) о возможности движения и о целях. Поэтому  $l$  процессов движения в среде  $a_j$  должны быть добавлены в образующие решетки целей  $M$ . Поскольку цели в среде присутствуют, но не видны, это означает, что среда видна до некоторого горизонта, причем горизонт может меняться в зависимости от направления. После обнаружения  $n$  целей  $b_i$  система должна выбрать  $k$  из них в соответствии со своей решеткой целей и своими возможностями. Именно, выбираются те  $k$  целей, достижение которых возможно и для которых переход от свободного движения к их достижению имеет наибольшую оценку в решетке целей  $M$ :  $a_1 \otimes \dots \otimes a_l \multimap b_1 \otimes \dots \otimes b_k = (a_1 \otimes \dots \otimes a_l \otimes (b_1 \otimes \dots \otimes b_k)^\perp)^\perp = \max$ . Это максимальная оценка импликации, представляющей переход от параллельного выполнения одних процессов к параллельному выполнению других. Эта оценка должна быть наибольшей среди всех возможных параллельных процессов достижения целей из  $n$  обнаруженных.

### 3.2. ОПРЕДЕЛЕНИЕ ВЫИГРЫША ПОЗИЦИЙ

Рассмотрим процесс взаимодействия системы со средой как игру Конвея. В этой игре среда – это игрок, который предоставля-

ет системе (оппоненту) информацию об объектах среды. Поэтому стратегии являются выигрышными для среды и выигранная игра заканчивается в позиции, в которой системе предоставляется наиболее полная информация. Система движется в среде из одной позиции в другую, используя на каждом шаге эту информацию, для того чтобы достичь свои цели. Эти цели являются некоторыми объектами среды и, с другой стороны, являются элементами решетки целей системы. Для системы из  $l$  агентов получается  $l$  параллельных процессов (корневых графов) достижения целей системы и, соответственно, полная игра будет их тензорным произведением. Таким образом, изначально система находится в среде в корневой позиции  $* = *_1 \otimes \dots \otimes *_l$  полной игры  $l$  агентов  $A = A_1 \otimes \dots \otimes A_l$ , имея решетку общих целей  $M$ . В силу того что среда видна только до горизонта, все пути в этой игре конечны и решетки намерений отдельных агентов  $M_i, i = 1, \dots, l$ , не определены до тех пор, пока нет видимых целей (предполагается, что агенты заранее не специализированы на определенные цели).

Игра  $A_i$  представляет все возможные свободные (не имеющие конкретной цели) движения агента  $i$  в среде. Реальная траектория выбирается из требования максимального общего выигрыша вдоль предполагаемого пути. При этом движения агента в среде оцениваются в соответствии с критерием оптимальности свободного движения  $a_i$  в игре  $A_i$  или процесса  $b_j$  в игре  $B_j$  достижения цели  $b_j$ . Здесь различаются в обозначениях цель  $b_j$  (процесс ее достижения) как элемент решетки целей и игра  $B_j$  (конкретная реализация процесса достижения цели, конкретный путь в среде).

В случае процесса свободного движения  $a_i$  в игре  $A_i$  таким критерием оптимальности может быть наибольшая степень свободы будущих движений, или наибольшая степень видимости из позиций предполагаемого пути, или и т.д., лишь бы эта информация описывалась множеством, которое является образом среды. В случае процесса достижения цели  $b_j$  положим, что чем лучше цель видна, т.е. чем больше о ней информации предоставляется средой, тем больше выигрыш. Таким образом, при движении

системы в среде информация, предоставляемая системе средой в позиции  $p_i$  о цели  $b_j$  или о среде  $a$  (т.е. образ цели или среды), является выигрышем в этой позиции и описывается множеством  $k(p_i, b_j)$  или  $k(p_i, a)$ . В случае если система имеет в памяти некий образ цели или образ среды (при движении по определенному маршруту в среде), это множество может соответствовать степени пересечения образа из памяти с полученным от среды. Чем больше это пересечение, чем более отчетливо идентифицируется цель, тем выигрыш больше. Однако в силу того, что невозможно иметь в памяти полный, всесторонний образ цели, невозможно получить и полное совпадение образа в памяти с образом, полученным от среды. Поэтому выигрыш не может быть равен максимальному элементу решетки множеств, описывающей цель.

Предположим, что система обнаружила  $n$  целей  $b_1 \dots b_n$ . Тогда выигрышная стратегия игры  $A' = A_1 \otimes \dots \otimes A_l \multimap B_1 \otimes \dots \otimes B_k$  определяет переход к новой игре  $A'$   $l$  параллельных процессов достижения  $k$  целей, которые могут быть достигнуты с наибольшей оценкой, из обнаруженных  $n$  в следующем смысле.

Будем выбирать путь из требования максимизации выигрыша вдоль пути внутри горизонта видимости:

$$\begin{aligned}
 (1) \quad & k_{A_1 \otimes \dots \otimes A_l \multimap B_1 \otimes \dots \otimes B_k}^{play} (a_1 \otimes \dots \otimes a_l \multimap b_1 \otimes \dots \otimes b_k) \equiv \\
 & \equiv k_{(A_1 \otimes \dots \otimes A_l)^\perp \wp B_1 \otimes \dots \otimes B_k}^{play} ((a_1 \otimes \dots \otimes a_l)^\perp \wp b_1 \otimes \dots \otimes b_k) \equiv \\
 & \equiv k_{(A_1 \otimes \dots \otimes A_l)^\perp}^{play} ((a_1 \otimes \dots \otimes a_l)^\perp) \vee k_{B_1 \otimes \dots \otimes B_k}^{play} (b_1 \otimes \dots \otimes b_k) \equiv \\
 & \equiv \neg k_{(A_1 \otimes \dots \otimes A_l)}^{play} (a_1 \otimes \dots \otimes a_l) \vee k_{B_1 \otimes \dots \otimes B_k}^{play} (b_1 \otimes \dots \otimes b_k) = \\
 & = \max_{plays} \left[ \bigcup_{play} (\neg k_{A_1}(a_1) \wedge \dots \wedge \neg k_{A_l}(a_l)) \bigcup_{play} (k_{B_1}(b_1) \wedge \dots \wedge k_{B_k}(b_k)) \right].
 \end{aligned}$$

Здесь максимум берется среди всех возможных путей. Выигрыш  $k_{A_1 \otimes \dots \otimes A_l \multimap B_1 \otimes \dots \otimes B_k}^{play} (a_1 \otimes \dots \otimes a_l \multimap b_1 \otimes \dots \otimes b_k)$  максимизируется в игре  $A'$  и соответствует конкретному пути в среде, конкретной реализации процесса  $a_1 \otimes \dots \otimes a_l \multimap b_1 \otimes \dots \otimes b_k$  достижения  $k$  целей. При этом сами параллельные процессы достижения  $k$  целей выбираются из требования наибольшей оценки в решетке

ке целей для этих процессов<sup>2</sup>. Эта формула аналогична формуле из [16] в том смысле, что максимизируется выигрыш вдоль планируемого пути, но сам путь и оценки выигрыша вычисляются иначе. Кроме того, в [16] нет способа выбора целей, к которым планируется путь.

Смысл формулы (1) в том, что, в соответствии с семантикой линейной импликации, система переходит от выполнения процессов свободного движения  $A_i$  к процессам достижения целей  $B_j$ . Выигрыши вычисляются для стратегии (т.е. пути) игры  $A^\perp \wp B$  (с упрощением обозначений). При этом максимизируется информация о целях  $k_{B_j}(b_j)$  и минимизируется информация о свободном движении  $k_{A_j}(a_j)$  (поскольку максимизируются псевдодополнения  $\neg k_{A_j}(a_j)$ ). Такие  $k$  и  $\neg k$  можно рассматривать как аргументы и контраргументы для соответствующего движения в соответствии с идеями ДСМ-метода правдоподобного вывода [2, 8, 9]: чем лучше цель видна (т.е. чем больше о ней информации), тем сильнее аргументы для перехода на ее достижение. Также чем сильнее контраргументы против свободного движения (т.е. чем меньше информации о возможности такого движения), тем сильнее аргументы для перехода от свободного движения к достижению какой-либо цели.

### 3.3. ТРУДНОСТИ ВЫБОРА

Первое замечание касается стабильности выбора целей в процессе их достижения. Поскольку только что обнаруженные цели видны плохо (так как степень видимости с расстоянием уменьшается), то идентифицировать их при обнаружении можно только с какой-то вероятностью. Поэтому выбор параллельных процессов достижения приоритетных целей из всех обнаруженных может меняться в процессе движения. Так же как и намерения отдельных агентов.

---

<sup>2</sup> Следует подчеркнуть, что операции  $\otimes$  и  $\rightarrow$  для игр  $A, B$  и для целей  $a, b$  разные. Для целей они определяются структурой линейной логики на решетке целей, — это внутренняя характеристика системы. А для игр структура линейной логики задается средой.

Второе замечание касается трудностей вычисления маршрута для больших групп. Поскольку выбор основан на переборе вариантов всех возможных путей, то с ростом числа агентов растет и объем вычислений. В случае близкого горизонта видимости, когда длина планируемых путей мала (например, у муравьев или роботов-дворников [7]), увеличение числа агентов не сильно скажется на времени вычисления, но с увеличением дальности видимости сложность вычислений может стать серьезной проблемой.

Третье замечание относится к однозначности определения маршрута и выбора целей. Здесь основное достоинство метода — решеточная структура выигрышей и множества целей — становится недостатком: в решетке не все элементы сравнимы между собой. Поэтому в случае необходимости выбора между несравнимыми значениями нужны дополнительные меры. Для этого можно воспользоваться большей степенью видимости какой-то цели по сравнению с другими (пример 1) при выборе целей и разной ценностью целей в решетках намерений отдельных агентов при выборе маршрута(пример 2).

**Пример 1.** Предположим, что решетка целей группы роботов-дворников имеет вид рис. 2 и два робота находятся в состоянии поиска мусора  $X_1$ . Это означает, что в системе выполняется процесс  $X_1 \otimes X_1$ . При обнаружении двух куч мусора, в одной из которых есть объект, требующий распила (задача  $X_3$ ), возможны следующие переходы к выполнению других процессов:

- 1)  $X_1 \otimes X_1 \rightarrow E \otimes E = (X_1 \otimes X_1 \otimes (E \otimes E)^\perp)^\perp$ ;
- 2)  $X_1 \otimes X_1 \rightarrow E \otimes X_3 = (X_1 \otimes X_1 \otimes (E \otimes X_3)^\perp)^\perp$ ;
- 3)  $X_1 \otimes X_1 \rightarrow X_3 \otimes X_3 = (X_1 \otimes X_1 \otimes (X_3 \otimes X_3)^\perp)^\perp$ .

В варианте 1 оба робота переключаются на вынос мусора. В варианте 2 один робот выносит мусор, а другой пилит то, что нельзя вынести. В варианте 3 оба робота переключаются на распил. В [7] вычислены все используемые здесь произведения, откуда следуют следующие возможные значения импликаций, которые оценивают варианты переходов:

- 1)  $X_1 \otimes X_1 \rightarrow E \otimes E = c_3$ ;
- 2)  $X_1 \otimes X_1 \rightarrow E \otimes X_3 = V_{23}$ ;
- 3)  $X_1 \otimes X_1 \rightarrow X_3 \otimes X_3 = i_{1E}$ .

Оказывается, что в первом варианте оценка меньше, чем в двух других, а оценки вариантов 2 и 3 несравнимы. Но предположим, что мусор, который не требует распила, виден лучше, или он ближе, или есть еще какое-нибудь преимущество. Тогда вариант 2 предпочтительнее для перехода, поскольку он включает задачу выноса  $E$  не требующего распила мусора наряду с распилом  $X_3$ .

•

**Пример 2.** В случае несравнимости выигрышей при выборе маршрута можно использовать метод из [5]. Пусть имеется игры трех агентов  $A_i$  и в позициях  $p_1$  обнаружены цели  $b_1, b_2$  (рис. 3).

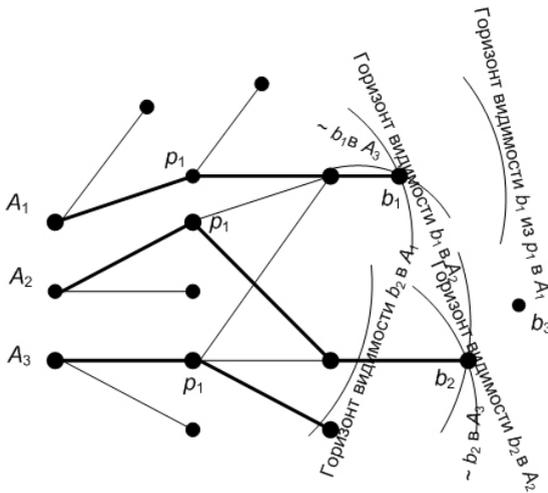


Рис. 3. Пример игры (изображены только движения системы)

Предположим также, что выигрыши в позициях  $p_1$  удовлетворяют следующим условиям:

- (2)  $k_{A_1}(p_1, b_1) > k_{A_2}(p_1, b_1) \bowtie k_{A_3}(p_1, b_1)$ ,
- (3)  $k_{A_3}(p_1, b_2) \bowtie k_{A_2}(p_1, b_2) > k_{A_1}(p_1, b_2)$ ,

где знак  $\bowtie$  означает несравнимость. Ясно, что цель  $b_1$  будет достигаться агентом-1, поскольку он видит ее лучше других. Пусть для двух оставшихся агентов решетки желаний имеют вид рис. 4.

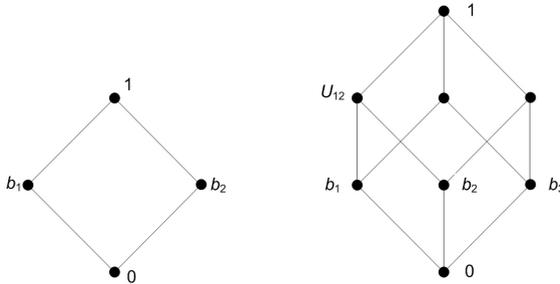


Рис. 4. а) решетка желаний агента-2, б) решетка желаний агента-3

На этом рисунке  $b_i$  – это желания агента, цели, которые агент хотел бы достичь;  $U_{12}$  – возможное намерение, в которое входят два желания, т.е.  $U_{12}$  – это те желания, которые система способна осуществить. Таким образом, агент-2 может при возможности достичь целей  $b_1, b_2$ , а агент-3 имеет еще дополнительную цель  $b_3$ . Тогда, в соответствии с [5], ценность намерения агента можно оценить формулой:

$$(4) \quad V_{intention} = n_{desires} / N_{desires\ total},$$

где  $V_{intention}$  – ценность намерения,  $n_{desires}$  – количество желаний в намерении,  $N_{desires\ total}$  – общее количество желаний агента. Эта формула оценивает намерение по двум параметрам: количеству желаний, входящих в намерение и близости намерения к наиболее желательному, т.е. к наибольшему элементу решетки, который объединяет все цели. Таким способом можно сравнивать ценности вершин в разных решетках. Например, ценность варианта  $U_{12}$ :  $V_{U_{12}} = 2/3$ . Поскольку  $V_{b_2} = 1/2$  в решетке агента-2, а  $V_{b_2} = 1/3$  в решетке агента-3, то следует выбрать агента-2 на достижение цели  $b_2$ , поскольку этот выбор более ценен. Таким образом, агент-3 остается свободным и может искать цель  $b_3$ . Результирующая игра изображена на рис. 3 жирными линиями. •

Наконец последний тонкий момент связан с тем, что цель может быть видна, но пути в среде к ней может не быть. Это означает, что может не быть монотонной последовательности выигрышей  $k_A(p_i, b_j)$ , которые относятся к достижению цели  $b_j$ , в позициях  $p_i$  движения  $A$ .

#### **4. Заключение**

В работе представлен подход к решению группой интеллектуальных агентов двух задач: задачи целераспределения и задачи выбора оптимального маршрута в среде. Решение этих задач основано только на информации, которую агенты получают от среды, и информации о целевом назначении системы агентов, которая заложена в систему изначально и представлена структурой линейной логики на множестве целей системы.

Эта структура получается из общих требований выполнения свойств линейной логики на решетке целей системы и представляет внутреннюю логику, присущую системе. Решетка же целей получается из соотношения между собой задач, которые призвана решать система (ее целей), и представления о наиболее активном поведении системы, когда достигаются все цели, как о наиболее ценном, а о бездействии — как о наименее ценном. В результате процессы параллельного достижения некоторых целей представляются как тензорное произведение в линейной логике на решетке целей, и можно выбирать те процессы, переход на выполнение которых наиболее ценен. Такой переход представляется линейной импликацией, значение которой оценивается наибольшим возможным элементом решетки целей. Так происходит выбор достижимых целей из всех обнаруженных.

Интересно, что маршрут выбирается тоже используя структуру линейной логики, но только эта структура задается средой: все возможные маршруты можно представить как пути в категории игр, в которой моделируется эта линейная логика. Опять переход от одной игры к другой представляется линейной импликацией и выбирается такая результирующая игра, суммарный выигрыш вдоль которой наибольший. Выигрыш, в свою очередь, определяется образами среды или цели в позициях игры. Это то, что агент видит на планируемом пути и эта информация представляется средой.

Получается, что в предложенном способе определения целей и маршрута в рамках подхода UAI, т.е. при представлении

агента как информационного процессора, линейная логика играет важную роль. До конца причина этого явления не понятна, но, возможно, дело в том, что в линейной логике посылка в выводе может использоваться только раз, что соответствует переходу от одного процесса, который завершается, к другому, который начинается. Это как раз соответствует внутренней логике процессов, происходящих в такой системе.

### Литература

1. БИРКГОФ Г. *Теория решеток*. – М.: Изд-во «Наука», – 1984. – 567 с.
2. ВИНОГРАДОВ Д.В. *Еще один вариант логики аргументации* // Научно-техническая информация. – Сер. 2. – №5. – М.: ВИНТИ, 2006. – С. 1–4.
3. КУЗНЕЦОВ А.В. *Упрощенная модель боевых действий на основе клеточного автомата* // Известия РАН. Теория и системы управления. – 2017. – Т. 56, № 3. – С. 59–71.
4. КУЗНЕЦОВ А.В. *Организация строя агентов с помощью клеточного автомата* // Управление большими системами. – 2017. – №70. – С. 136–170.
5. ЛЕГОВИЧ Ю.С., МАКСИМОВ Д.Ю. *Выбор исполнителя в группе интеллектуальных агентов* // Управление большими системами. – 2015. – №56. – С. 78–94.
6. МАКСИМОВ Д.Ю. *Реконфигурирование системной иерархии методами многозначной логики* // Автоматика и телемеханика. – 2016. – №3. – С. 123–136.
7. МАКСИМОВ Д.Ю., ЛЕГОВИЧ Ю.С., РЫВКИН С.Е. *Влияние структуры системных задач на поведение системы* // Автоматика и телемеханика. – 2017. – №4. – С. 135–148.
8. ФИНН В.К. *Об одном варианте логики аргументации* // Научно-техническая информация. – Сер. 2. – №5-6. – М.: ВИНТИ, 1996. – С. 3–19.

9. ANSHAKOV O.N, FINN V.K., VINOGRADOV D.V. *Logical Means for Plausible Reasoning of JSM-type* / Многочленные логики и их применение: Т.2: Логика в системах искусственного интеллекта; [под ред. В.К. Финна]. – М.: Изд-во ЛКИ, 2008. – С. 226–236.
10. BOSTROM N. *Superintelligence: Paths, Dangers, Strategies*. – New York: Oxford University Press, 2014.
11. EVERITT T., HUTTER M. *Universal artificial intelligence* // In: “Foundations of Trusted Autonomy. Studies in Systems, Decision and Control” / Eds.: Abbass H., Scholz J., Reid D. – Springer, Cham, 2018. – Vol. 117. – P. 15–46.
12. FALLENSTEIN B., SOARES N. *Problems of Self-reference in Self-improving Space-Time Embedded Intelligence* // Int. Conf. on Artificial General Intelligence. – 2014. – P. 21–32.
13. GIRAR J.-Y. *Linear logic* // Theoretical Computer Science. – 1987. – No. 50. – P. 1–102.
14. GOERTZEL B. *Artificial general intelligence: concept, state of the art, and future prospects* // Journal of Artificial General Intelligence. – 2014. – Vol. 5, No. 1. – P. 1–48.
15. HEIN A.M., CONDAT H. *Can Machines Design? An Artificial General Intelligence Approach* // ArXiv.org, 2018. – URL: <https://arxiv.org/pdf/1806.02091.pdf> (дата обращения: 16.10.2018).
16. HUTTER M. *One Decade of Universal Artificial Intelligence* // Theoretical Foundations of Artificial General Intelligence. – 2012. – P. 66–88.
17. KUZNETSOV A.V. *Model of the motion of agents with memory based on the cellular automaton* // Int. Journal of Parallel, Emergent and Distributed Systems. – 2018. – Vol. 33, No. 3. – P. 290–306.
18. KUZNETSOV A., SHISHKINA E., SITNIK S. *Probabilistic Properties of Near-Optimal Trajectories of an Agent Moving over a Lattice* // Journal of Optimization Theory and Applications. – 2018. – <https://doi.org/10.1007/s10957-018-1374-6>.

19. LAFONT Y. *Linear Logic Pages* – URL: <http://iml.univ-mrs.fr/lafont/pub/llpages.pdf> (дата обращения: 16.10.2018).
20. LEIKE J., HUTTER M. *Bad Universal Priors and Notions of Optimality* // JMLR: Workshop and Conf. Proc. – 2015. – Vol. 40. – P. 1–16.
21. MAXIMOV D., LEGOVICH Y., GONCHARENKO V. *A Way to Facilitate Decision Making in a Mixed Group of Manned and Unmanned Aerial Vehicles* // ArXiv.org, – 2018. – URL: <https://arxiv.org/abs/1809.10441> (дата обращения: 16.10.2018).
22. MELLIES P.-A., TABAREAU N. *Resource modalities in tensor logic* // Ann. Pure. Appl. Logic. – 2010. – No. 5. – P. 632–653.
23. SOARES N. *Formalizing Two Problems of Realistic World-Models* // Intelligence.org. – 2015. – URL: <https://intelligence.org/files/RealisticWorldModels.pdf> (дата обращения: 16.10.2018).
24. SOARES N., FALLENSTEIN B. *Questions of Reasoning Under Logical Uncertainty* // Intelligence.org. – 2015. – URL: <https://intelligence.org/files/QuestionsLogicalUncertainty.pdf>. (дата обращения: 16.10.2018).
25. SOARES N., FALLENSTEIN B. *Toward Idealized Decision Theory* // ArXiv.org, – 2015. – URL: <https://arxiv.org/abs/1507.01986>. (дата обращения: 16.10.2018).

## AN OPTIMAL ITINERARY GENERATION OF LARGE INTELLECTUAL AGENT GROUPS

**Dmitry Maximov**, Institute of Control Sciences of RAS, Moscow, researcher (dmmax@inbox.ru).

*Abstract: Intelligence is represented in the Artificial General Intelligence (or Universal AI) approach as an information processor which consumes and gives out information. This information completely determines the system behavior. M. Hutter has obtained an optimal agent trajectory selecting in an abstract environment within this approach. However, the method demands number assessments of the reward in such a movement, and the way to obtain these assessments was not defined. The method does not also fit to estimate an agent group movement. In this paper, reward assessments are represented by the environment. The way to choose a trajectory for an agent group is also suggested. A group of intelligent agents which fulfill a set of tasks in parallel is represented by the tensor multiplication of corresponding processes in the Conway game category. The category corresponds to the environment and agents' moves. An optimal itinerary of the group is defined as a play with maximal total reward in the category. The reward is represented as a degree of certainty (visibility) of an agent goal, i.e., by some set, not a number. It is proved that such a reward determination may be used in the categorical construction of Conway games. The tensor product in the category is the operation of linear logic. Such logic is also used in selecting goals which the system can achieve of all visible ones. In the purpose, the whole goal set is represented as a lattice with the linear logic structure. The goal lattice is the truth value set of the logic in this case. The tensor product of the lattice elements corresponds to goals which are achieved in parallel. The system goals are chosen by the highest truth value of the lattice element which corresponds to they parallel achieving processes. Such a construction results in the formula of the optimal itinerary estimation in an abstract environment for an agent group.*

Keywords: intelligent agents; itinerary choice; goal lattice; game semantics.

УДК 004.8, 007.5

ББК 22.18, 39.17

DOI: 10.25728/ubs.2019.78.3

*Статья представлена к публикации членом редакционной коллегии Я.И. Квинто.*

*Поступила в редакцию 15.11.2018.*

*Дата опубликования 31.03.2019.*