

(19)



**Евразийское
патентное
ведомство**

(11) **047062**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента
2024.05.29

(51) Int. Cl. **B25J 9/00** (2006.01)
B25J 9/16 (2006.01)

(21) Номер заявки
202393212

(22) Дата подачи заявки
2023.12.13

(54) **СИСТЕМА ФОРМИРОВАНИЯ КОМПЬЮТЕРНО-РЕАЛИЗУЕМОГО МЕТОДА
УПРАВЛЕНИЯ ФУНКЦИОНИРОВАНИЕМ РОБОТА И СПОСОБ ЕЁ РАБОТЫ**

(43) **2024.05.24**

(56) US-A1-20220305646
JP-A-2020192676
DE-A1-2020111271
US-A1-20220297709

(96) **2023000208 (RU) 2023.12.13**

(71)(73) Заявитель и патентовладелец:
**АВТОНОМНАЯ
НЕКОММЕРЧЕСКАЯ
ОБРАЗОВАТЕЛЬНАЯ
ОРГАНИЗАЦИЯ ВЫСШЕГО
ОБРАЗОВАНИЯ "СКОЛКОВСКИЙ
ИНСТИТУТ НАУКИ И
ТЕХНОЛОГИЙ" (RU)**

(72) Изобретатель:
**Чертков Андрей Владимирович,
Рыжак Глеб Владимирович,
Оселедец Иван Валерьевич (RU)**

(74) Представитель:
Ананьева А.Ю. (RU)

(57) Изобретение относится к автоматическому управлению процессами и оптимальному контролю. Технический результат, достигаемый изобретением, заключается в повышении точности формируемого метода. Технический результат достигается благодаря способу формирования компьютерно-реализуемого метода управления роботом, в котором устанавливается задача функционирования робота; случайно задаются ожидания на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров; итеративно выполняется уточнение ожиданий на максимум степени выполнения задачи; определяется набор значений параметров метода, при котором степень выполнения задачи максимальна. При уточнении ожиданий случайно с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум степени выполнения задачи, предлагается заданное количество предлагаемых наборов значений параметров; определяются степени выполнения задачи для предложенных наборов; на их основе отбирается заданное количество наборов значений параметров, степень достижения цели при которых наибольшая; ожидания на максимум уточняются с использованием метода градиентного подъёма таким образом, чтобы значения ожиданий на максимум для отобранных наборов значений параметров повысились.

B1

047062

047062

B1

Область техники, к которой относится изобретение

Изобретение относится к области автоматического управления процессами и оптимального контроля. Более конкретно, изобретение относится к системе формирования компьютерно-реализуемого метода управления функционированием робота и способом её работы.

Уровень техники

Задачи функционирования роботизированных средств, таких, например, как манипуляторы, погрузчики, беспилотные транспортные средства и других часто включают в себя задачи по разработке компьютерно-реализуемых методов (иногда называемых политиками) по управлению функционированием робота в некоторой заданной среде. Целевым результатом применения таких методов может быть приведение робота в некоторое требуемое положение или выполнение им какой-либо задачи по манипулированию материальными объектами. Формирование таких методов является нетривиальной задачей, для решения которой используются методы машинного обучения типа RL обучения агентов (Reinforcement Learning of an agents), в которых в качестве агентов представляются конструктивные элементы и приводы роботов.

Известна система и способ для формирования компьютерно-реализуемого метода управления функционированием в заданной среде робота, выполненного с возможностью управления посредством сообщения ему известного множества управляющих сигналов от системы управления. При этом управляющие сигналы формируются на основании метода управления, определяемого известным набором параметров, принимающих значения в известном множестве. При этом подбор параметров метода, при которых степень выполнения задачи максимальна, выполняется на основании степени выполнения роботом поставленной задачи. Для определения искомых значений параметров в известной системе используется такой метод, как TTOpt, известный первоначально из публикации [Sozykin, Konstantin, et al. "TTOpt: A maximum volume quantized tensor train-based optimization and its application to reinforcement learning." *Advances in Neural Information Processing Systems* 35 (2022): 26052-26065.]. В основе этого метода лежит парадигма ТТ-разложения и идея вычисления подматриц максимального объема (то есть подматриц, имеющих максимальный по модулю определитель) с использованием метода maxvol [Goreinov, Sergei A., et al. "How to find a good submatrix." *Matrix Methods: Theory, Algorithms And Applications: Dedicated to the Memory of Gene Golub*. 2010. 247-256.] для последовательных разверток (развертка - это матрица, получаемая при объединении размерностей тензора в две группы, соответствующие строкам и столбцам) оптимизируемого тензора. В методе используется эмпирический факт близости максимального по модулю элемента подматрицы максимального объема к максимальному по модулю элементу оптимизируемого тензора. На основе этого наблюдения в рамках данного метода последовательно запрашиваются кандидаты на оптимум из специально выбранных последовательных разверток тензора, соответствующих ядрам ТТ-разложения. В случае задачи минимизации, для нахождения минимального элемента осуществляется динамическое отображение посредством задаваемой пользователем функции запрашиваемых элементов тензора, которое преобразует минимальные значения в максимальные.

Недостатком известного метода является невысокая точность формируемого метода управления роботом. Данный метод был принят в качестве ближайшего аналога раскрытого далее изобретения.

Раскрытие изобретения

Задача, стоявшая перед разработчиками раскрытого изобретения, заключалась в создании системы и способа формирования компьютерно-реализуемого метода управления функционированием в заданной среде робота, который бы обеспечивал более высокую точность формируемого метода по сравнению с известным тензорным методом. При этом необходимо было сохранить сравнимую скорость работы.

В первом аспекте изобретения поставленная задача была решена путём создания системы формирования компьютерно-реализуемого метода управления функционированием в заданной среде робота, включающего в себя один или несколько приводов, каждый из которых выполнен с возможностью управляться посредством сообщения ему известного множества управляющих сигналов с помощью системы управления.

При этом система включает в себя блок управления, выполненный с возможностью формирования и сообщения приводам управляющих сигналов на основании метода управления, определяемого набором параметров, каждый из которых может принимать значения в известном множестве, а также выполненный с возможностью считывания наборов значений параметров с носителя значений параметров метода; и носитель значений параметров метода, выполненный с возможностью приёма, записи предлагаемых наборов значений параметров метода, а также с возможностями записи, хранения и предоставления для считывания блоком управления набора значений параметров, при котором степень выполнения задачи максимальна.

При этом система выполнена:

с возможностью установки задачи функционирования робота;

с возможностью задания количества предлагаемых наборов значений параметров, количества отбираемых наборов значений параметров, причём количество предлагаемых наборов значений параметров меньше количества возможных наборов значений параметров метода и больше количества отбираемых наборов значений параметров; а также

с возможностями задания критерия останова, задания значений ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству параметров;

При этом система формирования метода включает в себя:

носитель значений ожиданий на максимум степени выполнения задачи, выполненный с возможностью хранения значений ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров, причём параметрами разложения тензорного поезда являются элементы трёхмерных массивов;

блок предложения наборов значений параметров, выполненный с возможностью до выполнения заданного критерия останова итеративно предлагать заданное количество предлагаемых наборов значений параметров случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум, а также с возможностью последовательной передачи предлагаемых наборов значений параметров на носитель значений параметров метода;

блок определения степени выполнения задачи, выполненный с возможностью определения степени выполнения задачи для предложенных наборов значений параметров, а также с возможностью передачи определённых значений степени выполнения задачи для каждого из предложенных наборов значений параметров блоку сравнения значений степени выполнения задачи;

блок сравнения, выполненный с возможностью отбора из предлагаемых наборов значений параметров заданного количества отбираемых наборов значений параметров, от которых степень выполнения задачи наибольшая, а также с возможностью передачи данных об отобранных наборах значений параметров блоку уточнения значений ожиданий;

блок уточнения значений ожиданий, выполненный с возможностью уточнения параметров разложения тензорного поезда значений ожиданий на максимум с использованием метода градиентного подъёма таким образом, чтобы значения ожиданий на максимум для отобранных значений параметров повысились.

Во втором аспекте настоящего изобретения поставленная задача была решена за счёт создания компьютерно-реализуемого способа формирования компьютерно-реализуемого метода управления функционированием в заданной среде робота, включающего в себя один или несколько приводов, каждый из которых выполнен с возможностью управляться посредством сообщения ему известного множества управляющих сигналов с помощью системы управления, включающей в себя:

блок управления, выполненный с возможностью формирования и сообщения приводам управляющих сигналов на основании метода управления, определяемого набором параметров, каждый из которых может принимать значения в известном множестве, а также выполненный с возможностью считывания наборов значений параметров с носителя значений параметров метода; и

носитель значений параметров метода, выполненный с возможностью приёма, записи предлагаемых наборов значений параметров метода, а также с возможностями записи, хранения и предоставления для считывания блоком управления набора значений параметров, при котором степень выполнения задачи максимальна.

При этом способ включает в себя этапы, на которых устанавливается задача функционирования робота;

случайным образом задаются значения ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров, причём параметрами разложения тензорного поезда являются элементы трёхмерных массивов;

задаётся количество предлагаемых наборов значений параметров и количество отбираемых наборов значений параметров, причём количество предлагаемых наборов значений параметров меньше количества возможных наборов значений параметров и больше количества отбираемых наборов значений параметров;

задаётся критерий останова;

до выполнения заданного критерия останова итеративно выполняется этап уточнения значений ожиданий на максимум степени выполнения задачи;

определяется и передаётся на носитель значений параметров метода набор значений параметров метода, при котором степень выполнения задачи максимальна, и на его основе формируется метод управления роботом.

При этом этап уточнения значений ожиданий на максимум степени выполнения задачи включает в себя следующие действия

случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум степени выполнения задачи, предлагается заданное количество предлагаемых наборов значений параметров;

определяются степени выполнения задачи для предложенных наборов значений параметров;

на основе определённых степеней выполнения задачи из предложенных наборов значений параметров отбирается заданное количество отбираемых наборов значений параметров, степень достижения це-

ли при которых наибольшая;

параметры разложения тензорного произведения функции ожиданий на максимум уточняются с использованием метода градиентного подъёма таким образом, чтобы

значения ожиданий на максимум для отобранных наборов значений параметров повысились.

Технический результат, достигаемый раскрытым изобретением в обоих его аспектах, заключается в повышении точности формируемого метода. Под точностью метода стоит понимать полноту, с которой робот под управлением сформированного метода управления роботом выполняет поставленную перед ним задачу, определяющий также степень выполнения задачи.

Далее в настоящем описании раскрытое изобретение будет описано более детально со ссылкой на конкретные варианты осуществления и прилагаемые фигуры чертежей.

Краткое описание чертежей

Фиг. 1 - функциональная схема системы в соответствии с раскрытым изобретением.

Фиг. 2 - общая блок-схема способа в соответствии с раскрытым изобретением.

Фиг. 3 - блок-схема итерации уточнения значений ожиданий на максимум.

Фиг. 4 - относительная ошибка формирования метода, при различных количествах K предлагаемых наборов значений параметров и количествах k отбираемых наборов значений параметров. Зелёная линия с квадратами - график при $K=50$. Синяя линия со звёздами - график при $K=100$. Оранжевая линия с ромбами - график при $K=150$. Салатовая линия с кругами - график при $K=200$. Фиолетовая линия с пятиугольниками - график при $K=250$.

Фиг. 5 - результаты сравнения раскрытого изобретения с аналогами.

Варианты осуществления изобретения

Изобретение относится к системе и компьютерно-реализуемому способу для формирования компьютерно-реализуемого метода управления функционированием в заданной среде 11 робота 10, включающего в себя один или несколько приводов, каждый из которых выполнен с возможностью управляться посредством сообщения ему известного множества управляющих сигналов с помощью системы 20 управления. К таким роботам может быть отнесены промышленные манипуляторы, беспилотные автомобили, дроны и т.д. При этом система управления включает в себя блок 21 управления и носитель 22 значений параметров метода.

При этом блок 21 управления в соответствии с раскрытым изобретением выполнен с возможностью формирования и сообщения приводам управляющих сигналов на основании метода управления, определяемого набором параметров, каждый из которых может принимать значения в известном множестве, а также выполненный с возможностью считывания наборов значений параметров с носителя 22 значений параметров метода. В качестве такого блока управления может быть использован контроллер, включающий в себя компьютер, выполненный с возможностью выполнения компьютерно-реализуемого метода управления роботом 10.

В соответствии с раскрытым изобретением носитель 22 значений параметров метода, выполнен с возможностью приёма, записи предлагаемых наборов значений параметров метода, а также с возможностями записи, хранения и предоставления для считывания блоком 21 управления набора значений параметров, при котором степень выполнения задачи максимальна. В качестве носителя 22 значений параметров метода может быть использован любой носитель или группа носителей информации, предпочтительно являющийся энергонезависимым. Например, носитель значений параметров метода может включать в себя твердотельный SSD (Solid-State Drive) накопитель.

Способ в соответствии с раскрытым изобретением включает в себя этап, на котором устанавливается 100 задача функционирования робота 10. Задачей, устанавливаемой для робота 10, может быть, например, приведение некоторой части среды 11 в требуемое положение или в перемещении робота 10 в среде 11 в заданное положение. Степень выполнения задачи может оцениваться различными способами. Например, с помощью некоторой функции награды, зависящей от результата управления роботом 10, например, от его конечного или промежуточного положения в среде 11. При этом, предпочтительным является такой метод, при котором степень выполнения установленной задачи максимальна.

В предпочтительных вариантах осуществления метод управления может характеризоваться параметрами, являющимися дискретными. В вариантах осуществления, когда параметры не являются дискретными, способ может включать в себя их дискретизацию.

При этом степень выполнения роботом 10 задачи может характеризоваться численной характеристикой степени выполнения, которую можно представить как некоторую функцию награды от набора значений параметров компьютерно-реализуемого метода. Значение степени выполнения задачи может определяться путём запуска метода управления роботом и последующей фиксации параметров степени выполнения задачи с помощью датчиков, камер и/или других чувствительных регистрирующих устройств.

Задача нахождения максимума степени выполнения задачи может сводиться к нахождению набора значений параметров x_{\max} , для которых значение степени выполнения задачи (выраженной, например, в виде функции $f(x)$ награды) является максимальным:

$$\mathbf{x}_{max} = \max_{\mathbf{x}} \mathbf{f}(\mathbf{x})$$

Однако в других вариантах осуществления это множество может быть любым другим. При этом количество возможных дискретных значений каждой из параметров n_i известно: N_1, N_2, \dots, N_d .

В рамках компьютерно-реализуемого способа степень выполнения задачи может быть представлена в виде параметрической функции $\mathbf{f}(\mathbf{x})$ награды от заданного количества; d дискретных параметров

$$\mathbf{x} = [n_1, n_2, \dots, n_d]$$

которые могут принимать значения из известного множества. Для простоты можно представить, что это подмножество натуральных чисел

$$n_i \in \{1, 2, \dots, N_i\}$$

Однако в других вариантах осуществления это множество может быть любым другим. При этом количество возможных дискретных значений каждой из параметров n_i известно: N_1, N_2, \dots, N_d .

Эта функция награды (степени выполнения задачи) может быть представлена в виде неявно заданного d -мерного тензора:

$$\mathcal{Y} \in \mathbb{R}^{N_1 \times N_2 \times \dots \times N_d}$$

$$\mathcal{Y}[n_1, n_2, \dots, n_d] = \mathbf{f}(\mathbf{x})$$

В настоящем описании под тензором следует понимать многомерную таблицу или многомерный массив.

В предпочтительном варианте осуществления изобретения обеспечение значений степени выполнения задачи осуществляется по мере выполнения способа путём определения её значений для предлагаемых наборов значений параметров. При этом данный способ предполагает определение значений степени выполнения задачи только для части возможных наборов значений параметров.

В соответствии с раскрытым изобретением способ включает в себя этап, на котором случайным образом задаются 200 значения ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров, причём параметрами разложения тензорного поезда являются элементы трёхмерных массивов (Tensor-Train Decomposition, Oseledets I., SIAM Journal on Scientific Computing, (2011), 2295-2317, 33(5), 10.1137/090752286).

Говорится, что d -мерный тензор

$$\mathcal{P} \in \mathbb{R}^{N_1 \times N_2 \times \dots \times N_d}$$

задан в форме разложения тензорного поезда, когда его элементы представлены следующей формулой:

$$\mathcal{P}[n_1, n_2, \dots, n_d] = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \dots \sum_{r_{d-1}=1}^{R_{d-1}} \mathcal{G}_1[1, n_1, r_1] \mathcal{G}_2[r_1, n_2, r_2] \dots \mathcal{G}_d[r_{d-1}, n_d, 1]$$

Где (n_1, n_2, \dots, n_d) - это мультииндекс ($n_i=1, 2, \dots, N_i$ для $i=1, 2, \dots, d$), целые числа R_0, R_1, \dots, R_d , с условием, что $R_0=R_d=1$, именуется рангами тензорного поезда (или ТТ-ранг), а трёхмерные тензоры

$$\mathcal{G}_i \in \mathbb{R}^{R_{i-1} \times N_i \times R_i} \quad (i = 1, 2, \dots, d)$$

ируются ядрами тензорного поезда.

Разложение тензорного поезда (или ТТ-разложение) позволяет представить тензор или дискретизированную многомерную функцию в компактной и наглядной малопараметрической форме, которая линейна по размерности d , т.е. имеет менее

$$d \cdot \max_{i=1, \dots, d} (N_i R_i^2) \sim d \cdot \bar{N} \cdot \bar{R}^2$$

параметров, где

$$\bar{N}$$

и

$$\bar{R}$$

- это эффективный ("средний") размер моды и ТТ-ранг соответственно. Под размером моды здесь понимается количество возможных значений параметров.

В ходе выполнения предложенного метода оценивается вероятность максимума степени выполнения задачи при значениях параметров, соответствующих элементам тензора. Это выполняется путём итеративного уточнения 500 значений ожиданий на максимум в виде тензорного поезда, форма которого соответствует форме задания функции, характеризующей степень выполнения задачи. Иными словами, значения ожиданий на максимум задаются в виде разложения тензорного поезда вида

$$\mathcal{P}_\theta \in \mathbb{R}^{N_1 \times N_2 \times \dots \times N_d}$$

В некоторых вариантах осуществления изобретения параметры тензорного поезда значений ожида-

ний на максимум могут задаваться так, чтобы значения ожиданий на максимум для всех наборов значений параметров были равновероятными, случайными или иным образом. С учётом того, что способ в соответствии с раскрытым изобретением может выполняться многократно для каждой конкретной задачи, предпочтительно задавать значения ожиданий на максимум случайным образом; это позволит разными путями находить параметры, при которых степень выполнения задачи максимальна, что повышает вероятность корректного их нахождения. При этом, как следует из существа значений ожиданий на максимум (значение вероятности), её элементы не могут быть отрицательными. Для функционирования способа не требуется выполнять нормировку значений ожиданий на максимум (приведение вероятностей к значениям, которые в сумме составляют 1).

Чтобы задать неотрицательный ТТ-тензор заданной размерности N_1, N_2, \dots, N_d с постоянным рангом тензорного произведения R , достаточно создать d ТТ-ядер

$$\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_d$$

(трёхмерных тензоров) со случайными элементами из равномерного распределения на интервале (0, 1).

Способ в соответствии с раскрытым изобретением также включает в себя этап, на котором задаётся 300 количество K предлагаемых наборов значений параметров и количество k отбираемых наборов значений параметров. Причём количество предлагаемых наборов значений параметров меньше количества возможных наборов значений параметров и больше количества отбираемых наборов значений параметров.

Конкретные значения количества K предлагаемых наборов значений параметров и количества k отбираемых наборов значений параметров могут выбираться для каждой конкретной задачи. В ходе испытаний раскрытого метода для более простой задачи оптимизации аналитической функции были исследованы значения K от 50 до 250 и значения k от 5 до 25, как это представлено на фиг. 4. Сравнение относительной ошибки выполнения способа показало небольшую разницу в эффективности на данном диапазоне, поэтому ниже приведённые испытания эффективности раскрытого способа проводились при $K=100$ и $k=10$. В некоторых вариантах осуществления изобретения значения количества K предлагаемых наборов значений параметров и количества k отбираемых наборов значений параметров может по мере выполнения способа изменяться. Например, через некоторое количество выполненных итераций уточнения значений ожиданий на максимум количество K предлагаемых наборов значений параметров и/или количество k отбираемых наборов значений параметров могут быть уменьшены.

Способ в соответствии с раскрытым изобретением включает в себя этап, на котором задаётся 400 критерий останова. Критерий останова необходим для обеспечения конечности выполнения итераций уточнения 500 ожиданий на максимум. Критерий останова может включать в себя максимальное количество выполняемых итераций, сходимость метода (например, когда предлагаемые наборы значений параметров повторяются в течение нескольких итераций этапа уточнения 500 ожиданий на максимум), а также другие условия и их комбинации.

Способ в соответствии с раскрытым изобретением включает в себя этап, на котором до выполнения заданного критерия останова итеративно выполняется этап уточнения 500 ожиданий на максимум. Этап уточнения 500 ожиданий на максимум представляет собой последовательность циклично выполняемых шагов, включающих в себя по меньшей мере шаги, на которых случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум, предлагается 501 заданное количество K предлагаемых наборов значений параметров; определяются 502 значения степени выполнения задачи для предложенных K наборов значений параметров; на основе определённых значений степени выполнения задачи из предложенных наборов значений параметров отбирается 503 заданное количество k отбираемых значений параметров, значения степени выполнения задачи при которых являются наибольшими; параметры разложения тензорного произведения значений ожиданий на максимум уточняются 504 с использованием метода градиентного подъёма таким образом, чтобы значения ожиданий на максимум для отобранных k наборов значений параметров повысились. Вместе с тем, при применении метода градиентного подъёма в силу особенностей метода также повышаются значения ожиданий на максимум для наборов значений параметров, близких к отобранным.

Способ в соответствии с раскрытым изобретением включает в себя этап, на котором определяется 600 и передаётся на носитель значений параметров метода набор значений параметров метода, при котором степень выполнения задачи максимальна, и на его основе формируется метод управления роботом.

Способ в некоторых вариантах осуществления включает в себя этап определения максимума степени выполнения задачи. В некоторых вариантах осуществления этап определения максимума может выполняться после завершения этапа уточнения 500 ожиданий на максимум и включать в себя определение значения степени выполнения задачи на основании набора значений параметров, соответствующего максимальному значению ожиданий на максимум. В других вариантах осуществления этап определения максимума может выполняться в ходе этапа уточнения 500 ожиданий на максимум путём фиксации наибольших значений из определённых на шаге определения 502 степени выполнения задачи для предложенных наборов значений параметров.

В соответствии с раскрытым изобретением этап уточнения 500 ожиданий на максимум включает в себя шаг, на котором случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум, предлагается 501 заданное количество K предлагаемых наборов значений параметров.

$$\mathcal{P}_\theta: \mathcal{X}_K = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K\}$$

Предложение 501 заданного количества K предлагаемых наборов значений параметров случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум, в предпочтительном варианте осуществления изобретения включает в себя процедуру семплирования из тензора в формате тензорного поезда. Так, чтобы сгенерировать мульти-индекс \mathbf{x} (или набор значений параметров) с

вероятностью, пропорциональной соответствующему значению

$$p = \mathcal{P}[\mathbf{x}]$$

ТТ-тензора

$$\mathcal{P}$$

может быть использован подход, известный из источника [Sergey Dolgov et al. "Approximation and sampling of multivariate probability distributions in the tensor train decomposition". In: Statistics and Computing 30 (2020), стр. 603-625.]. Этот подход основан на последовательном вычислении одномерных условных плотностей вероятности с эффективным интегрированием в ТТ-формате. При этом способ позволяет не выполнять нормирование вероятностей, поэтому он не требует вычисления коэффициента нормализации.

В соответствии с раскрытым изобретением этап уточнения 500 ожиданий на максимум включает в себя шаг, на котором определяются (502) степени выполнения задачи для предложенных наборов значений параметров.

$$y_1 = f(\mathbf{x}_1), y_2 = f(\mathbf{x}_2), \dots, y_K = f(\mathbf{x}_K)$$

Этап определения 502 степеней выполнения задачи для предложенных наборов значений параметров в предпочтительном варианте осуществления изобретения выполняется путём записи предложенных наборов значений параметров на носитель 22 значений параметров метода системы 20 управления, последующей реализации метода управления с каждым из предлагаемых наборов параметров и последующим определением степени выполнения задачи для каждого из наборов значений параметров. В различных вариантах осуществления на данном этапе может использоваться как робот 10 и среда 11 как таковые, так и их модели.

В соответствии с раскрытым изобретением этап уточнения 500 значений ожиданий на максимум включает в себя шаг, на котором на основе определённых степеней выполнения задачи из предложенных наборов значений параметров отбирается 503 заданное количество отбираемых значений параметров, степени выполнения задачи при которых являются наибольшими. Индексы таких значений параметров можно представить следующим образом:

$$\mathcal{S} = \{s_1, s_2, \dots, s_k\}$$

В соответствии с раскрытым изобретением этап уточнения 500 значений ожиданий на максимум может включать в себя шаг, на котором параметры разложения тензорного поезда значений ожиданий на максимум уточняются с использованием метода градиентного подъёма (Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv: 1609.04747) таким образом, чтобы значения ожиданий на максимум для отобранных наборов значений параметров повысились. Это достигается путём применения одного или нескольких шагов градиентного подъёма к следующему функционалу

$$\hat{L}_\theta(\{x_{s_1}, x_{s_2}, \dots, x_{s_k}\}) = \sum_{i=1}^k \log(\mathcal{P}_\theta[x_{s_i}])$$

представляющему собой сумму логарифмов значений тензора значений ожиданий на максимум в отобранных к наборах параметров.

Значение шага градиентного подъёма α (или скорости обучения - learning rate) может определяться в зависимости от особенностей конкретной задачи и может составлять, например, от 0,0001 до 0,1. При этом градиентный подъём в некоторых случаях может выполняться на величину нескольких шагов, что не выходит за рамки раскрытого изобретения.

В некоторых вариантах осуществления изобретения уточнение ожидания на максимум может выполняться путём сложения функции ожидания на максимум с вспомогательным тензором, который, имея аналогичную размерность, равен во всех индексах нулю, кроме индексов кандидата на максимум, ожидание для которого необходимо повысить. Для индексов кандидата на максимум, ожидание для которого необходимо повысить, значение тензора может быть равно некоторому δ , который определяет шаг уточнения ожидания на максимум. В некоторых вариантах осуществления значения тензора в соседних (по отношению к индексам кандидата на максимум, значение для ожидания на максимум для которого

необходимо повысить) индексах может также спадать от δ до нуля. Операция сложения может выполняться путём поэлементного сложения в ТТ-формате. При сложении ранг тензора повышается, поэтому после этого необходимо выполнить операцию округления (Tensor-Train Decomposition, Oseledets I., SIAM Journal on Scientific Computing, (2011), 2295-2317, 33(5), 10.1137/090752286), чтобы снизить ранг до исходного значения.

Способ в соответствии с раскрытым изобретением может выполняться с помощью системы 30 формирования компьютерно-реализуемого метода управления функционированием в заданной среде 11 робота 10, включающего в себя один или несколько приводов, каждый из которых выполнен с возможностью управляться посредством сообщения ему известного множества управляющих сигналов с помощью системы 20 управления.

При этом система 20 управления в соответствии с раскрытым изобретением включает в себя блок 21 управления, выполненный с возможностью формирования и сообщения приводам управляющих сигналов на основании метода управления, определяемого набором параметров, каждый из которых может принимать значения в известном множестве, а также выполненный с возможностью считывания наборов значений параметров с носителя 22 значений параметров метода. В качестве блока 21 управления может выступать контроллер, персональный компьютер или другое вычислительное устройство, выполненное с возможностью реализовывать формируемый метод управления роботом, а также выполненное с возможностью формировать на основании метода управляющие сигналы и сообщать их приводам робота.

Система 20 управления также включает в себя носитель 22 значений параметров метода, выполненный с возможностью приёма, записи предлагаемых наборов значений параметров метода, а также с возможностями записи, хранения и предоставления для считывания блоком 21 управления набора значений параметров, при котором степень выполнения задачи максимальна. В качестве носителя 22 значений параметров метода может быть использовано любое энергонезависимое запоминающее устройство, которое соединено с блоком 21 управления с возможностью передачи ему по меньшей мере параметров метода или с возможностью считывания данных, включающих в себя по меньшей мере параметры метода, с носителя 22 значений параметров метода блоком управления 21. Также носитель 22 значений параметров метода выполнен с возможностью приёма и записи на него наборов значений параметров метода. Приём предлагаемых наборов значений параметров может осуществляться от системы 30 формирования компьютерно-реализуемого метода управления, в частности от блока 32 предложения наборов значений параметров.

Система 30 в соответствии с раскрытым изобретением выполнена с возможностью установки 100 задачи функционирования робота 10. Задачей, устанавливаемой для робота 10, может быть, например, приведение некоторой части среды 11 в требуемое положение или в перемещении робота 10 в среде 11 в заданное положение. При этом степень выполнения задачи может оцениваться различными способами, например путём оценки близости к целевому положению робота или части среды или путём оценки количества итераций выполнения этапов способа формирования компьютерно-реализуемого метода, которые позволили выполнить поставленную задачу. Степень выполнения задачи может быть представлена в виде некоторого значения награды, зависящего от результата функционирования робота под управлением метода с предложенным набором значений параметров, которое численно характеризует степень выполнения роботом 10 установленной задачи. Степень выполнения задачи может определяться на основании показаний датчиков или анализа изображений с камер, которые позволяют определить состояние робота: например, положение робота в пространстве (в среде) или части среды (например, груза), либо данных программного счётчика итераций выполнения способа формирования метода. При этом предполагается, что предпочтительным является такой метод, при котором степень выполнения установленной задачи (а вместе с ней и значение награды в случае её применения) максимальна.

Система 30 в соответствии с раскрытым изобретением выполнена с возможностью задания 300 количества K предлагаемых наборов значений параметров и количества k отбираемых наборов значений параметров, причём количество предлагаемых наборов значений параметров меньше количества возможных наборов значений параметров и больше количества отбираемых наборов значений параметров. Выбор количества K предлагаемых наборов значений параметров и количества k отбираемых наборов значений параметров может осуществляться системой из предварительно заданного диапазона или они могут быть предварительно указаны оператором, администратором или пользователем системы 30. Значения K могут составлять, например, от 50 до 250, предпочтительные значения k могут составлять от 5 до 25.

Также система 30 в соответствии с заявленным изобретением выполнена с возможностью задания 400 критерия останова. Критерий останова необходим для обеспечения конечности выполнения итераций уточнения 500 значений ожиданий на максимум. Критерий останова может включать в себя максимальное количество выполняемых итераций уточнения 500 значений ожиданий на максимум, сходимость метода (например, когда предлагаемые наборы значений параметров повторяются в течение нескольких итераций этапа уточнения 500 значений ожиданий на максимум), а также другие условия и их комбинации.

Система 30 в соответствии с заявленным изобретением выполнена с возможностью задания 200

значений ожиданий на максимум случайным образом в форме разложения тензорного пезда в виде трёхмерных массивов в количестве, соответствующем количеству параметров.

Система 30 формирования компьютерно-реализуемого метода включает в себя носитель 31 значений ожиданий на максимум степени выполнения задачи, выполненный с возможностью хранения значений ожиданий на максимум в форме разложения тензорного пезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров, причём параметрами разложения тензорного пезда являются элементы трёхмерных массивов. В качестве носителя 31 значений ожиданий на максимум может быть использовано любое постоянное или оперативное запоминающее устройство. При этом носитель 31 значений ожиданий на максимум выполнен с возможностью передачи или предоставления для считывания значений ожиданий на максимум блоку 32 предложения наборов значений параметров.

Система 30 формирования компьютерно-реализуемого метода включает в себя блок 32 предложения наборов значений параметров, выполненный с возможностью до выполнения заданного критерия останова итеративно предлагать 501 заданное количество предлагаемых наборов значений параметров случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум. Кроме того, блок 32 предложения наборов значений параметров выполнен с возможностью передачи предлагаемых наборов значений параметров на носитель 22 значений параметров метода для записи их на упомянутый носитель 22.

Система 30 формирования компьютерно-реализуемого метода включает в себя блок 33 определения степени выполнения задачи, выполненный с возможностью определения степени выполнения задачи для предложенных наборов значений параметров, а также с возможностью передачи определённых значений степени выполнения задачи для каждого из предложенных наборов значений параметров блоку 34 сравнения значений степени выполнения задачи. Для определения степени выполнения задачи блок 33 определения степени выполнения задачи может включать в себя чувствительные узлы, с помощью которых может определяться факт выполнения роботом установленной задачи. В качестве таких узлов могут использоваться, например, любые устройства захвата изображения, датчики или средства навигации в среде. Блок 33 определения степени выполнения задачи может быть выполнен с возможностью обработки сигналов с чувствительных узлов и определения на их основании степени выполнения задачи роботом в численном виде, например в виде значения награды. В случаях, когда степень выполнения задачи выполняется на основании результатов управления моделью робота в среде, блок 33 определения степени выполнения задачи определяет степень выполнения задачи на основании состояния модели в результате функционирования робота под управлением метода управления роботом с предлагаемыми наборами значений параметров.

Система 30 формирования компьютерно-реализуемого метода включает в себя блок 34 сравнения, выполненный с возможностью отбора из предлагаемых наборов значений параметров заданного количества отбираемых наборов значений параметров, от которых степень выполнения задачи наибольшая, а также с возможностью передачи данных об отобранных наборах значений параметров блоку 35 уточнения значений ожиданий. В качестве блока 34 сравнения может быть использовано любое вычислительное устройство, выполненное с возможностью выполнения операции сравнения определённых блоком 33 определения степени выполнения задачи степеней выполнения задачи, которых удалось достичь с использованием предложенных наборов значений параметров. Блок 34 сравнения также выполнен с возможностью передачи блоку 35 уточнения значений ожиданий на максимум данных об отобранных k наборах значений параметров из предложенных K наборов значений параметров, для которых степень выполнения задачи оказалась наибольшей. Блок 34 сравнения может включать в себя по меньшей мере одну электронную вычислительную машину, которая содержит машиноисполняемые инструкции, которые позволяют из определённых значений степени выполнения задачи из предложенных K значений параметров отбирать k наборов значений параметров, для которых степень выполнения задачи является наибольшей. Кроме того, блок 34 сравнения выполнен с возможностью передачи данных о наборах значений параметров, соответствующих отобранным k значениям степени выполнения задачи блоку 35 уточнения значений ожиданий. Для этого блок 34 сравнения может быть соединён с возможностью передачи указанных данных блоку 35 уточнения значений ожиданий.

Система 30 формирования компьютерно-реализуемого метода включает в себя блок 35 уточнения значений ожиданий на максимум, выполненный с возможностью уточнения значений ожиданий на максимум, хранящихся на носителе 31 значений ожиданий на максимум. Для реализации такой возможности блок 35 уточнения значений ожиданий может быть выполнен с возможностью уточнения значений параметров разложения тензорного пезда функции ожиданий на максимум с использованием метода градиентного подъёма таким образом, чтобы значения ожиданий на максимум для отобранных значений параметров повысились. Для обеспечения возможности уточнения значений ожиданий блок 35 уточнения значений ожиданий выполнен с возможностью считывания информации о значениях ожиданий на максимум из блока носителя 31 значений ожиданий на максимум, а также с возможностью записи новых значений ожиданий на носитель 31 значений ожиданий на максимум. В качестве блока 35 уточнения значений ожиданий на максимум может быть использована вычислительная машина, выполненная с

возможностью выполнения операций, необходимых для уточнения значений функции ожиданий методом градиентного подъёма. Блок 35 уточнения значений ожиданий на максимум может включать в себя по меньшей мере одну электронно-вычислительную машину, выполненную с возможностью приёма данных о наборах значений параметров, соответствующих отобранным k наборам значений параметров от блока 34 сравнения; с возможностью чтения и записи информации на носитель 31 значений ожиданий на максимум; а также содержащую машиноисполняемые инструкции, которые при их выполнении на электронно-вычислительной машине блока 35 уточнения значений ожиданий на максимум заставляют её уточнять параметры разложения тензорного поезда функции ожиданий на максимум, содержащейся на носителе 31 значений ожиданий на максимум, с использованием метода градиентного подъёма таким образом, чтобы значения ожиданий на максимум для отобранных значений параметров повысились.

В настоящем описании под блоками может подразумеваться как отдельные вычислительные машины, так и их группы. В отдельных вариантах осуществления функции нескольких блоков могут выполняться одной вычислительной машиной или группой вычислительных машин.

Для сравнения эффективности раскрытого способа с известными аналогами, все они были развёрнуты на вычислительной машине MacBook Pro 16' (Процессор - 6 ядерный Intel Core i7 2,6 ГГц; Графика: AMD Radeon PRO 5300M 4 Гб, Intel UHD Graphics 630, 1536 Мб; память 16 Гб 2667 MHz DDR4) в виде исполняемых скриптов на языке Python, и каждый из них протестирован на решении задач 3 модели роботов разных типов: 1) перевёрнутый маятник (inverted pendulum, доступный из библиотеки MuJoCo https://www.gymnasium.dev/environments/mujoco/inverted_pendulum/, дата обращения 24.10.2023); 2) пловец (swimmer, доступный из библиотеки MuJoCo <https://www.gymnasium.dev/environments/mujoco/swimmer/>); 3) лунный шаттл (https://www.gymnasium.dev/environments/box2d/lunar_lander/, Lunar lander, дата обращения 24.10.2023). Ко всем трём задачам были применены методы, являющиеся ближайшими аналогами раскрытого изобретения (способ TTOpt - тензорный метод, являющийся ближайшим аналогом раскрытого способа, и способ Portfolio, известный из библиотеки nevergrad, не являющийся тензорным, но применяющийся для аналогичных задач), а также метод и система в соответствии с заявленным изобретением (PROTES). В данном испытании в качестве критерия останова было задано максимальное количество итераций уточнения значений ожиданий на максимум 100 000. Испытания проводились при $K=100$ и $k=10$. Результаты испытания приведены на фиг. 4. Для каждого из экспериментов испытываемый метод был запущен три раза. В таблице приведены средние значения степени выполнения задачи (reward) и времени формирования метода (time) в секундах из трёх экспериментов. Испытание показало, что раскрытый способ формирует более точные методы, так как их степень выполнения задачи максимальна во всех случаях, при этом скорость формирования метода сравнима со скоростью ближайшего аналога (TTOpt), а относительно второго аналога (Portfolio) во всех случаях выше.

Настоящее изобретение было подробно описано со ссылкой на отдельные варианты его осуществления, однако очевидно, что оно может быть осуществлено в различных вариантах, не выходя за рамки заявленного объёма правовой охраны, определяемого формулой изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Система формирования компьютерно-реализуемого метода управления функционированием в заданной среде робота, включающего в себя один или несколько приводов, каждый из которых выполнен с возможностью управляться посредством сообщения ему известного множества управляющих сигналов с помощью системы управления, включающей в себя

блок управления, выполненный с возможностью формирования и сообщения приводам управляющих сигналов на основании метода управления, определяемого набором параметров, каждый из которых может принимать значения в известном множестве, а также выполненный с возможностью считывания наборов значений параметров с носителя значений параметров метода; и

носитель значений параметров метода, выполненный с возможностью приёма, записи предлагаемых наборов значений параметров метода, а также с возможностями записи, хранения и предоставления для считывания блоком управления набора значений параметров, при котором степень выполнения задачи максимальна;

при этом система выполнена

с возможностью установки задачи функционирования робота;

с возможностью задания количества предлагаемых наборов значений параметров, количества отбираемых наборов значений параметров, причём количество предлагаемых наборов значений параметров меньше количества возможных наборов значений параметров метода и больше количества отбираемых наборов значений параметров; а также

с возможностями задания критерия останова, задания значений ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству параметров;

при этом система формирования метода включает в себя

носитель значений ожиданий на максимум степени выполнения задачи, выполненный с возможностью хранения значений ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров, причём параметрами разложения тензорного поезда являются элементы трёхмерных массивов;

блок предложения наборов значений параметров, выполненный с возможностью до выполнения заданного критерия останова итеративного предложения заданного количества наборов значений параметров случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум, а также с возможностью передачи предлагаемых наборов значений параметров на носитель значений параметров метода;

блок определения степени выполнения задачи, выполненный с возможностью определения степени выполнения задачи для предложенных наборов значений параметров, а также с возможностью передачи определённых значений степени выполнения задачи для каждого из предложенных наборов значений параметров блоку сравнения значений степени выполнения задачи;

блок сравнения, выполненный с возможностью отбора из предлагаемых наборов значений параметров заданного количества отбираемых наборов значений параметров, от которых степень выполнения задачи наибольшая, а также с возможностью передачи данных об отобранных наборах значений параметров блоку уточнения значений ожиданий;

блок уточнения значений ожиданий, выполненный с возможностью уточнения параметров разложения тензорного поезда значений ожиданий на максимум таким образом, чтобы значения ожиданий на максимум для отобранных значений параметров повысились.

2. Система по п. 1, отличающаяся тем, что

блок уточнения значений ожиданий выполнен с возможностью уточнения параметров разложения тензорного поезда значений ожиданий на максимум с использованием метода градиентного подъёма.

3. Компьютерно-реализуемый способ формирования компьютерно-реализуемого метода управления функционированием в заданной среде робота, включающего в себя один или несколько приводов, каждый из которых выполнен с возможностью управляться посредством сообщения ему известного множества управляющих сигналов с помощью системы управления, включающей в себя

блок управления, выполненный с возможностью формирования и сообщения приводам управляющих сигналов на основании метода управления, определяемого набором параметров, каждый из которых может принимать значения в известном множестве, а также выполненный с возможностью считывания наборов значений параметров с носителя значений параметров метода; и

носитель значений параметров метода, выполненный с возможностью приёма, записи предлагаемых наборов значений параметров метода, а также с возможностями записи, хранения и предоставления для считывания блоком управления набора значений параметров, при котором степень выполнения задачи максимальна;

причём способ включает в себя этапы, на которых

устанавливается задача функционирования робота;

случайным образом задаются значения ожиданий на максимум степени выполнения задачи в форме разложения тензорного поезда в виде трёхмерных массивов в количестве, соответствующем количеству дискретных параметров, причём параметрами разложения тензорного поезда являются элементы трёхмерных массивов;

задаётся количество предлагаемых наборов значений параметров и количество отбираемых наборов значений параметров, причём количество предлагаемых наборов значений параметров меньше количества возможных наборов значений параметров и больше количества отбираемых наборов значений параметров;

задаётся критерий останова;

до выполнения заданного критерия останова итеративно выполняется этап уточнения значений ожиданий на максимум степени выполнения задачи;

определяется и передаётся на носитель значений параметров метода набор значений параметров метода, при котором степень выполнения задачи максимальна, и на его основе формируется метод управления роботом;

при этом этап уточнения значений ожиданий на максимум степени выполнения задачи включает в себя следующие действия

случайным образом с вероятностью, пропорциональной соответствующим значениям ожиданий на максимум степени выполнения задачи, предлагается заданное количество предлагаемых наборов значений параметров;

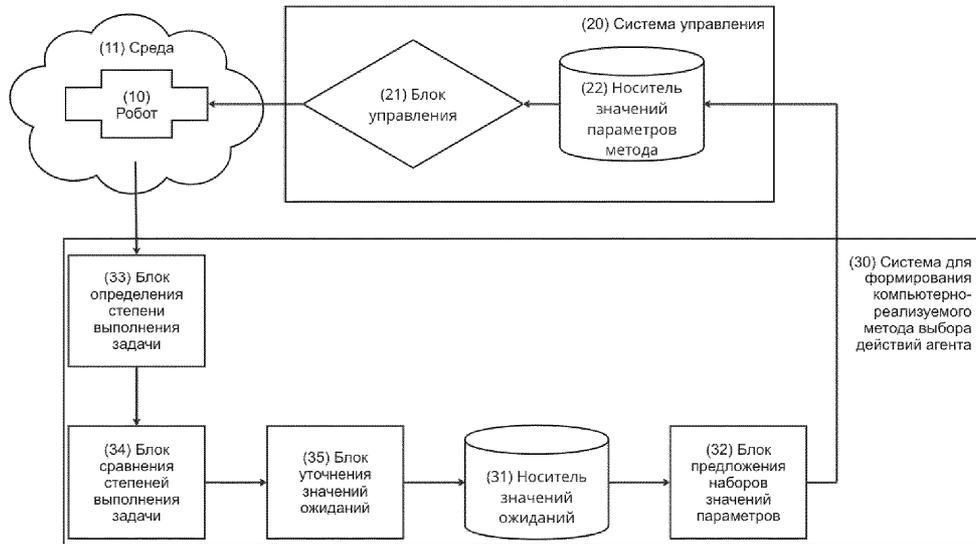
определяются степени выполнения задачи для предложенных наборов значений параметров;

на основе определённых степеней выполнения задачи из предложенных наборов значений параметров отбирается заданное количество отбираемых наборов значений параметров, степень достижения цели при которых наибольшая;

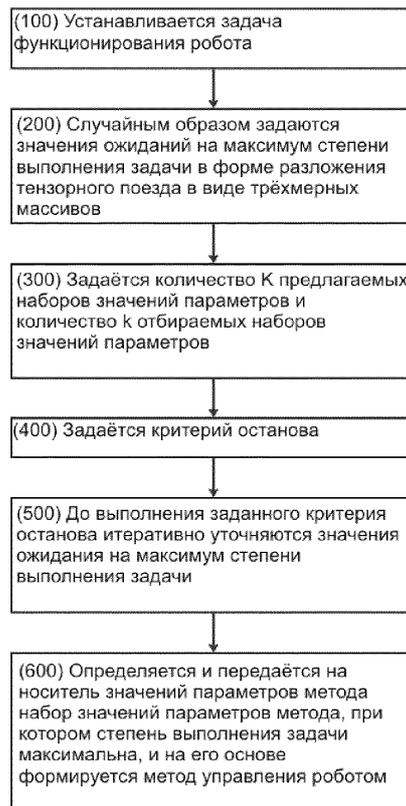
параметры разложения тензорного поезда значений ожиданий на максимум уточняются таким образом, чтобы значения ожиданий на максимум для отобранных наборов значений параметров повыси-

лись.

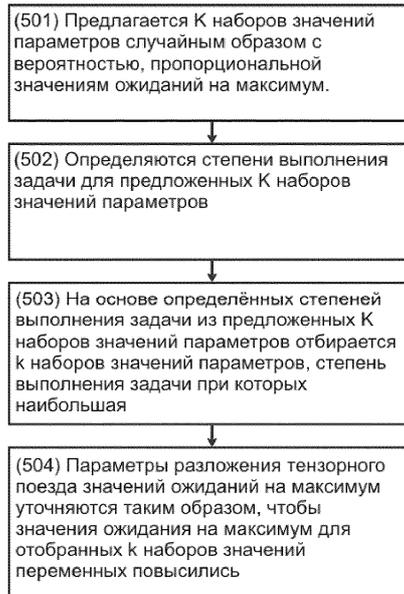
4. Способ по п.3, отличающийся тем, что параметры разложения тензорного произведения значений ожиданий на максимум уточняются с использованием метода градиентного подъёма.



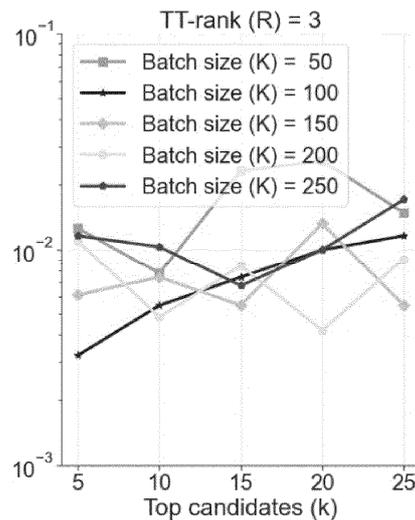
Фиг. 1



Фиг. 2



Фиг. 3



Фиг. 4

	REWARD			TIME		
	PROTES	ТТОПТ	PORTFOLIO	PROTES	ТТОПТ	PORTFOLIO
INVERTED PENDULUM	1.00E+03	7.28E+02	3.49E+02	8.1E+02	6.3E+02	1.3E+04
SWIMMER	3.61E+02	3.43E+02	3.55E+02	9.7E+03	7.7E+03	2.3E+04
LUNAR LANDER	2.80E+02	-1.87E+01	2.64E+02	3.5E+03	8.1E+03	1.5E+04

Фиг. 5

