УДК 681.3

О.Ю. Бочкарьов
Національний університет “Львівська політехніка”,
кафедра “Електронні обчислювальні машини”

ВИРІШЕННЯ ЗАДАЧІ МЕХАНІЧНОГО ВРІВНОВАЖЕННЯ КОЛЕКТИВОМ МОБІЛЬНИХ АГЕНТІВ

Сформульована задача механічного врівноваження в рамках теорії колективної поведінки, зокрема наведена її змістовна інтерпретація та розглянуті різні варіанти задачі. Розглянуті алгоритми врівноважуючих інтелектуальних агентів: централізований, випадковий, педантичний алгоритми, алгоритми на основі цілеспрямованих автоматів (learning automata) та алгоритми на основі навчання з підкріпленням (reinforcement learning). Обговорюються результати чисельного моделювання розглянутих алгоритмів та робляться висновки щодо підтвердження основних принципів теорії колективної поведінки.

Problem of mechanical equilibrium in framework of cooperative behavior theory is stated including the problem practical meaning and problem types' discussions. Algorithms of balancing intelligent agents are considered including centralized, probabilistic, pedantic algorithms, algorithms based on learning automata and algorithms based on reinforcement learning. Results of these algorithms numerical simulation are considered. Correctness of basic principles of cooperative behavior theory is stated on the basis of simulation results.

1. Вступ

У цій роботі вперше розглядається задача механічного врівноваження з точки зору теорії колективної поведінки (ТКП) [1,2,3]. Постановка задачі виникла внаслідок спільного розгляду двох проблем: перенесення вантажів за допомогою роботів-вантажників [4] та підвищення остійності морського вантажного судна в умовах сильної хитавиці [5]. Ключовим моментом в цій задачі є принцип рівноваги, який детально досліджується у межах інших близьких за змістом моделях колективної поведіники (економічна рівновага, екологічна рівновага, рівновага за Нешем в теорії ігор). Однак в зазначених моделях увага дослідників концентрується на специфічних для відповідних областей процесах, що призводить, з одного боку, до зловживання змістовною інтерпретацією досліджуваних моделей (у тому сенсі, що реальні процеси, що проходять у відповідних областях – економіка, екологія, гра) "диктують" досліднику певний напрямок розвитку моделей цих процесів з метою більш докладного їх дослідження) та невиправданого ускладнення математичного апарату цих моделей, яке натомість не призводить до якихось суттєвих результатів з точки зору ТКП (у тому сенсі, що ці результати можна легко отримати набагато простішими шляхами). Запропонована в даній роботі модель колективної поведінки на основі вирішення задачі механічного врівноваження дозволяє узагальнити принцип рівноваги та дослідити його найбільш повно саме з точки зору ТКП. При цьому область змістовної інтерпретації запропонованої моделі (механіка) добре досліджена, внаслідок чого модель використовується для дослідження принципів ТКП, а не принципів механіки.

У п. 2 сформульована задача механічного врівноваження, зокрема наведена її змістовна інтерпретація та розглянуті різні варіанти задачі. У п. 3 розглянуті алгоритми врівноважуючих інтелектуальних агентів: централізований, випадковий, педантичний алгоритми, алгоритми на основі цілеспрямованих автоматів (learning automata) та алгоритми на основі навчання з підкріпленням (reinforcement learning). У п. 4 обговорюються результати числового моделювання розглянутих алгоритмів та робляться висновки щодо підтвердження основних принципів ТКП. У п. 5 зроблено висновки щодо отриманих результатів.

2. Постановка задачі

2.1. Змістовна інтерпретація задачі

Розглянемо площину з однією точкою опори, яка обрана так, що ця площина знаходиться у горизонтальному положенні (врівноваженому стані). При цьому точка опори буде знаходитись у центрі ваги даної площини. Розмістимо випадковим чином на цій площині N агентів. Всі агенти мають однакову вагу, яку приймемо за одиницю. Кожний з агентів здатний переміщуватись по площині за власною ініціативою. Внаслідок розміщення агентів площина вийде зі стану рівноваги (новий центр ваги не збігатиметься з обраною раніше точкою опори). Перед колективом агентів ставиться задача врівноважити площину, тобто знайти таке взаєморозміщення, яке б повернуло її у горизонтальне положення. Ускладненням задачі є те, що врівноважити площину треба за якомога меншу кількість часових кроків.

Розглянемо більш простий випадок одновимірної “площини”, тобто звичайних важільних терезів у вигляді коромисла з однією точкою опори, що як завгодно точно збігається з його центром ваги. Припустимо, що агенти мають можливість пересуватись по всьому коромислу, яке надає для них одновимірний обмежений з двох боків лінійний простір розміром М дискретів. Кожний з агентів може реалізувати в цьому просторі одну з трьох дій: зробити крок вліво (d_1), зробити крок вправо (d_2), залишитись на місці (d_3).

Змістовною інтерпретацією (фізичною моделлю) цієї задачі є теорія важільних терезів, яку вперше розробив Л.Ейлер. За теорією Ейлера відхилення коромисла терезів під дією навантаження P описується такою формулою [6, 7]

$$tg(\varphi) = \frac{Pl}{x_0 P_0}, \tag{1}$$

де φ – кут відхилення коримисла терезів; l – довжина плеча коромисла; x_0 – відстань від точки опори до центра ваги коромисла; P_0 – вага коромисла. При цьому ми розглядаємо лише випадок ідеального коромисла, яке абсолютно не деформується і маса якого не залежить від довжини.

Як можна бачити з (1), на кут відхилення φ впливають два змінні у часі параметри: вага навантаження P та довжина плеча коромисла l. Роль навантаження в даному випадку грають агенти, що пересуваються по коромислу. Відстань агента від точки опори грає роль плеча коромисла. Оскількі агентів багато, то потрібно розглядати сумарний момент сил, з яким агенти діють на терези. З точки зору зовнішнього спостерігача одна частина агентів буде розміщена зліва від точки опори, а друга – справа. Назвемо лівим моментом суму

$$m_l = \sum_{i=1}^{N_l} p_i d_i \,, \tag{2}$$

де p_i – вага i–го агента, що знаходиться зліва від точки опори, d_i – відстань між цим агентом і точкою опори, N_l – кількість агентів, що знаходяться зліва від точки опори. Відповідно правим моментом назвемо суму

$$m_r = \sum_{i=1}^{N_r} p_i d_i \ , \tag{3}$$

де p_i – вага і–го агента, що знаходиться справа від точки опори, d_i – відстань між цим агентом і точкою опори, N_r – кількість агентів, що знаходяться справа від точки опори. При цьому

$$N_l + N_r = N. \tag{4}$$

Абсолютне значення різниці лівого та правого моментів буде визначати кут відхилення, а знак різниці буде визначати, в який бік відхилився важіль: вправо чи вліво. Відповідно формула (1) набуде такого вигляду

$$tg(\varphi) = \frac{|m_l - m_r|}{x_0 P_0} . \tag{5}$$

Прийнявши, що вага кожного агента дорівнює одниниці, то з урахуванням того, що кут відхилення дорівнюватиме нулю при нульовій різниці правого та лівого моментів, отримаємо рівняння задачі

$$\sum_{i=1}^{N} (x - c_i) = 0 \ , \tag{6}$$

де x – координата точки опори (в межах простору коромисла), а c_i – координата і–го агента. При цьому завжди має виконуватись нерівність

$$c_{i-1} < c_i < c_{i+1}. \tag{7}$$

2.2. Обговорення різних варіантів задачі

Можна розглядати різні за змістом задачі залежно від того, чи володіє агент інформацією про

– координату точки опори;
– свою власну координату;
– розмір простору (кількість дискретів простору коромисла M);
– розмір колективу (кількість агентів N).

Надалі будемо розглядати задачі врівноваження, в яких вся ця інформація є недоступною для агента.

Крім цього, різні варіанти задачі виникають завдяки таким моментам:

– яка інформація про стан важеля доступна кожному окремому агенту (це може бути кут відхилення φ (або більш простий варіант – різниця моментів) і тоді всі агенти отримують на вхід однаковий відгук середовища ("кутові" агенти), або це може бути висота агента над рівнем рівноваги важеля і тоді всі агенти отримують різні відгуки середовища ("висотні" агенти));

– чи можуть агенти "перестрибувати" один через одного;

– чи можуть агенти знаходитись у одному дискреті простору коромисла (тобто вставати один одному на "плечі");

– який колектив утворюють агенти: несамовиявлений (кожний окремий агент не володіє ніякою інформацією про інших агентів) чи самовиявлений (кожний окремий агент володіє певною інформацією про інших агентів);

– який рівень складності агента у випадку самовиявленого колектива (перший варіант: лише детектування в заданому радіусі видимості: єдиний результат – це відстань до виявленого сусіда; другий варіант: обмін інформаційними повідомленнями між агентами);

– якою саме інформацією обмінюються агенти у разі самовиявленого колективу з обміном інформацією.

З огляду на тип середовища на варіант задачі впливають такі моменти:

– якою є кількість дискретів простору коромисла *M*: парною чи непарною;

– де знаходиться точка опори, тобто яка в неї координата (номер дискрету простору): рівно посередені коромисла або зі зміщенням від його центру ваги;

– чи припадає точка опори на дискрет простору (і тоді агент, який знаходиться в цьому дискреті, ніяк не впливає на відхилення важеля, оскільки плече, з яким він діє на коромисло, дорівнює нулю), чи на стик між двома дискретами простору (і тоді агент, який знаходиться в одному з цих дискретів, діє на коромисло з плечем, що дорівнює одиниці).

Способи ускладнення задачі:

– збільшення розмірності простору (1D→2D→3D);

– точка опори переміщується за деяким законом (координата точки опори *x* є функцією від часу: *x(t)*);

– вага агентів є неоднаковою і змінною в часі та просторі (або за ініціативою агентів, або під дією зовнішніх сил);

– радіус видимості змінюється у часі та просторі (наприклад, зі збільшенням відстані до точки опори він зменшується);

– простір може бути викривленим (нелінійним).

3. Алгоритми врівноважуючих агентів

Основна вимога, що висувається до алгоритмів врівноважуючих агентів – це здатність врівноважувати важіль при будь-якому стартовому розміщенні агентів у просторі коромисла (сходимість алгоритму врівноваження). Друга вимога має ознаки критерію якості і визначає кількість часових кроків, яка знадобилась агентам, що виконують той чи інший алгоритм, щоб врівноважити важіль (мінімізація часу врівноваження). Третя і остання вимога полягає у визначенні кількості просторових кроків, яка знадобилась агентам , що виконують той чи інший алгоритм для того, щоб врівноважити важіль (мінімізація кількості просторових кроків).

Розглянемо такі алгоритми врівноваження, які мають ознаки локальності та уніфікованості [2].

1. Централізований алгоритм. Цей алгоритм розглядається як відправна точка у дослідженні гранично децентралізованих алгоритмів ТКП. Порівняння цих алгоритмів з централізованим буде додатковою ознакою їх ефективності. При виконанні централізованого алгоритму всі рішення відносно переміщень агентів приймає єдиний центр управління, який віддає накази підпорядкованим агентам (*slave agents*).

2. Випадковий алгоритм. Цей алгоритм розглядається з метою визначення цілеспрямованності зусиль агентів [2], виконуючих інші алгоритми врівноваження. Згідно з цим алгоритмом випадковий агент (*casual agent*) обирає дії $\{d_1, d_2, d_3\}$ з однаковою ймовірністю *0.3(3)*.

3. Педантичний алгоритм. Це одна з можливих реалізацій гранично децентралізованого алгоритму повного перебору, який задовольняє лише першу вимогу. Цей алгоритм розглядається, щоб оцінити успішність інших алгоритмів щодо задоволення ними другої та третьої вимог. Згідно з цим алгоритмом педантичні агенти (*pedantic agents*) спочатку збираються в *N* крайніх дискретах простору коромисла (справа або зліва) і після визначення моменту повного збору починають по одному переміщуватись на інший бік коромисла до моменту врівноваження.

4. Алгоритми на основі цілеспрямованих автоматів (learning automata) [2, 3]. Розглядаються самовиявлені та несамовиявлені колективи цілеспрямованих автоматів, для яких зменшення кута відхилення еквівалентно сигналу про виграш, а збільшення кута відхилення еквівалентно сигналу про програш. Алгоритми відповідних агентів (*LA-agents*) відповідають конструкціям цілеспрямованих автоматів [2, 3], серед яких докладно розглядались: автомати з лінійною тактикою, ймовірнісні автомати (автомати В.І.Крилова) та стохастичні автомати зі змінною структурою [3].

5. Алгоритми на основі навчання з підкріпленням (reinforcement learning) [8]. Розглядаються самовиявлені та несамовиявлені колективи агентів, для яких так само, як і для цілеспрямованих автоматів, зменшення кута відхилення еквівалентно сигналу про виграш, а збільшення кута відхилення еквівалентно сигналу про програш. Алгоритми відповідних агентів (*RL-agents*) відповідають базовим алгоритмам навчання з підкріпленням [8], серед яких докладно розглядались: "жадібний" алгоритм зваженої оцінки дій (greedy), ε-"жадібний" алгоритм зваженої оцінки дій (ε-greedy), алгоритм навчання з нерівноцінним вибором дій (softmax action selection), алгоритм навчання методом порівняння з підкріпленням (reinforcement comparison method) та алгоритм навчання методом переслідування (pursuit method).

4. Моделювання алгоритмів врівноваження

Центральною проблемою задачі врівноваження є те, що окремий агент діє на власний розсуд і тому в найгіршому випадку знаходиться в ситуації повної невизначеності стосовно правильності своїх дій. Зробивши хід і отримавши відгук (реакцію) середовища у вигляді різниці моментів, агент не може зробити однозначного висновку про те чи був цей його хід виграшним (врівноважуючим) чи програшним (розрівноважуючим), оскільки середовище (терези) реагує на колективну дію усіх агентів, а не на одну лише дію даного агента. Внаслідок цього агент, що не володіє інформацією про дії інших агентів колективу, приречений на сумніви відносно правильності своїх дій. Отже, інтуїтивно можна зробити висновок, що найкращим є те рішення, яке найкраще долає цю невизначеність.

У ході моделювання алгоритмів врівноваження отримані залежності часу врівноваження (кількість тактів моделювання, за які досягнуто стан рівноваги) від типу врівноважуючих агентів та параметрів інформаційної зв’язності колективу агентів. Моделювався одновимірний простір (коромисло терезів) з M дискретами, в яких випадково (рівномірний розподіл) розміщувались N агентів (стартове розміщення), кожний з яких виконував один з вищенаведених алгоритмів врівноваження. У ході моделювання фіксувався факт врівноваження або неврівноваження коромисла терезів. У разі врівноваження фіксувалася кількість тактів моделювання, за які колектив агентів врівноважив коромисло та загальна кількість переміщень агентів колективу.

Отримано підтвердження того факту, що при збільшенні інформаційної зв’язності колективу агентів зменшується час врівноваження (тобто скоріше долається невизначеність щодо дій інших агентів). Отримано експериментальні дані щодо ефективності розглянутих вище алгоритмів врівноважуючих агентів. Зроблено їх порівняльний аналіз. Отримано підвердження принципу співвідношення централізованого та децентралізованого управління [2] та принципу навчання з підкріпленням [8].

5. Висновки

Сформульовано задачу механічного врівноваження, зокрема її змістовну інтерпретацію та розглянуті різні варіанти задачі. Розглянуто алгоритми врівноважуючих інтелектуальних агентів: централізований, випадковий, педантичний алгоритми, алгоритми на основі цілеспрямованих автоматів (learning automata) та алгоритми на основі навчання з підкріпленням (reinforcement learning). Наведено результати числового моделювання розглянутих алгоритмів та зроблено висновки щодо підтвердження основних принципів теорії колективної поведінки.

1. Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, by Gerhard Weiss (Editor), MIT Press, 2000. 2. Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. – М., 1969. 3. Варшавский В.И. Коллективное поведение автоматов. – М.: Наука, 1973. 4. Cem Ünsal, Self-organization in Large Populations of Mobile Robots, Master thesises, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1993. 5. Огирко Н.М. Исследование и разработка электрических приборов контроля загрузки судов. Дис. ... канд. техн. наук. Львов, 1970. –218с. 6. Рудо М.М. Лабораторные весы и точное взвешивание. – М., 1963. 7. Загальна фізика: Лабораторний практикум. / За заг. ред. І.Т. Горбачука. – К., 1992. 8. Richard S. Sutton, Andrew G. Barto, Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 1998.

УДК 681.3, 621.3

О.Ю. Бочкарьов, В.А. Голембо
Національний університет "Львівська політехніка",
кафедра "Електронні обчислювальні машини"

МОДЕЛІ КОЛЕКТИВНОЇ ПОВЕДІНКИ ВИМІРЮВАЛЬНИХ АГЕНТІВ

Розглядаються підходи до вирішення проблеми моделювання розподілених вимірювально-обчислювальних систем з використанням сучасних досягнень теорії колективної поведінки. Наводиться постановка задачі розподілених вимірювань в термінах теорії колективної поведінки. Розглядаються близькі за змістом до розподілених вимірювань моделі колективної поведінки та обговорюються їх недоліки. Запропоновано інтерполяційну та ентропійну моделі колективної поведінки вимірювальних агентів.

Approaches to solving problems of modelling destributed measuring-processing systems using modern achievements of cooperative behaviour theory are considered. Problem of destributed measurements is stated in terms of cooperative behaviour theory. Cooperative behaviour models close to destributed measurements and their drawbacks are considered. Interpolational and entropic models of measurement agent cooperative behaviour are introduced.

1. Вступ

Розглядаються підходи до вирішення проблеми моделювання розподілених вимірювально-обчислювальних систем, зокрема безпровідних вимірювально-обчислювальних мереж [1] з використанням сучасних досягнень теорії колективної поведінки. За