

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП'ЮТЕРНА ТЕХНІКА

УДК 004.02

О. О. Марченко¹
 О. І. Марченко¹
 Б. О. Щербина¹

ПОБУДОВА ДЕРЕВА ПОШУКУ СПОСОБОМ З ВИКОРИСТАННЯМ МЕТОДУ МОНТЕ-КАРЛО І КОНТРОЛЕМ ФОРМИ ДЕРЕВА

¹Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Детально описано механізм впливу способу контролю форми дерева пошуку на процес виконання пошуку в дереві методом Монте-Карло (MCTS). Виконано порівняння способу реалізації MCTS з контролем форми дерева зі стандартним способом реалізації цього методу. Показано, що за правильного підбору коефіцієнтів використання способу реалізації MCTS з контролем форми дерева приводить до підвищення результативності пошуку і, в результаті, до прийняття кращих рішень.

Ключові слова: дерева ігор, пошук в дереві, метод Монте-Карло, MCTS, способи покращення MCTS.

Вступ

Пошук в дереві з використанням методу Монте-Карло (Monte Carlo Tree Search — MCTS) [1] є методом пошуку оптимізованих рішень у заданій області за допомогою випадкових значень із заданого простору значень і побудови дерева пошуку за отриманими результатами. Наразі вже запропоновано багато варіантів реалізації MCTS та способів покращення цього методу [1, 2, 3]. Древа пошуку різних ігор мають різний степінь розгалуження на вершинах дерева. Наприклад, шахи мають степінь розгалуження 35, а найскладніша з відомих ігор гра Го має набагато більший степінь розгалуження 250 [4]. Очевидно, що побудова повного дерева рішень для ігор з таким ступенем розгалуження є неможливою. З однієї сторони, чим більша кількість гілок дерева, вже побудованих в дереві в поточний момент часу, тим кращий аналіз подальших ходів можна зробити. З іншої сторони, побудова більшої кількості гілок в кожній вершині дерева рішень вимагає занадто багато часу і занадто великого обсягу пам'яті. Тому, знаходження дійсно найкращого ходу в поточній позиції за заданого обмеження на час виконання ходу є задачею нетривіальною.

Вдосконалення методу MCTS з метою прискорення процесу пошуку кращих рішень може бути виконано на різних етапах такого пошуку. Один зі способів покращення ефективності роботи MCTS за допомогою контролю форми дерева пошуку був запропонований авторами у [5].

Цей спосіб покращення відштовхується від деяких результатів, отриманих іншими дослідниками, основна ідея яких полягає в тому, що різні ігри (задачі) мають різні типи для них форми дерева пошуку побудовані методом MCTS. Очевидно, що для ігор з великим ступенем розгалуження буде будуватись ширше дерево, ніж для ігор з малим ступенем розгалуження. Крім того, Гілмар Фіннссон та Інґві Бьорнссон з'ясували важливі для нашого дослідження факти: “relative to the number of nodes in the trees the depth and width can be varied, allowing us to answer the question if MCTS favors one over the other. ... We found that it depends on the game itself whether MCTS prefers deep trees, big branching factor, or a balance between the two. Apparently small nuances in game rules and scoring systems, may alter the preferred game-tree structure” [6].

Тому, на відміну від більш традиційних способів налаштування MCTS з метою його покращення, в тому числі на основі різних варіантів розподілу ймовірностей, метою дослідження авторів є підтвердження наявності залежності ефективності пошуку MCTS від форми дерева, а також знаходження можливості покращення пошуку методом MCTS за допомогою оберненого впливу, тоб-

то корекцією поточної форми дерева під час процесу пошуку. Теоретичне підґрунтя цього способу розглянуто авторами в [5], а у цій статті розв'язується конкретна задача аналізу процесу побудови дерева пошуку за використання варіанта MCTS з контролем форми дерева в порівнянні з побудовою дерева стандартним варіантом реалізації методу MCTS, а також підтверджується існування залежності ефективності пошуку MCTS від форми дерева пошуку на основі окремих отриманих результатів. Детальний порівняльний аналіз отриманих результатів має великий обсяг і може бути предметом окремої роботи.

Побудова дерева MCTS з контролем форми дерева

Як один з підходів рішення зазначеної вище проблеми підвищення ефективності пошуку в дереві кращих варіантів продовження гри запропоновано спосіб, що базується на контролі форми дерева пошуку [5]. Цей спосіб вдосконалює UCB1 формулу стандартного варіанту MCTS [1] використанням ще одного доданку, який встановлює залежність остаточного значення цієї формули від форми поточного дерева пошуку на основі критеріїв типу «глибина–ширина» DWCs (Depth-Width Criteria) [5, 7]. Такий спосіб покращення MCTS названий пошуком по дереву з використанням методу Монте-Карло з контролем форми дерева (Monte-Carlo Tree Search with Tree Shape Control – MCTS-TSC) [5]. Цей спосіб використовує два критерії DWC1 та DWC2 для оцінювання і контролю форми всіх піддерев поточного дерева пошуку. Критерії DWCs обчислюються для кожної вершини на основі використання двох властивостей форми дерева — глибини (NodeDepth) піддерева та ширини піддерева (NodeWidth) цієї вершини.

$$DWC1 = C_{DWC1} * NodeDepth / NodeWidth;$$

$$DWC2 = C_{DWC2} * (P * NodeDepth - NodeWidth) / NodeWidth;$$

$$DWC3 = C_{DWC3} * (P * NodeDepth + NodeWidth) / NodeWidth,$$

де C_{DWC1} , C_{DWC2} , C_{DWC3} та P — це коефіцієнти, призначення яких є подібним до призначення коефіцієнта C_{UCB1} у формулі UCB1 [1, 5].

Критерії DWC1 та DWC2 використовуються як доданки до значення формули UCB1 і можуть використовуватись як окремо, так і разом. Відповідно до цього отримуємо три варіанти контролю форми дерева (tree shape control — TSC) у разі пошуку методом MCTS [5].

$$TSC1 = UCB1 + DWC1; \quad TSC2 = UCB1 + DWC2; \quad TSC3 = UCB1 + DWC3.$$

Розглянемо принцип впливу форми дерева на хід процесу побудови дерева пошуку на прикладі варіанта TSC1. Очевидно, чим більше значення відношення $NodeDepth/NodeWidth$ і коефіцієнта C_{DWC1} , тим більшим буде вплив форми дерева на поведінку процесу пошуку. Якщо коефіцієнт C_{DWC1} має додатне значення, то додатковий вплив форми дерева буде в напрямку кращого дослідження варіантів можливих ходів у «вузьких» піддеревих і приводить, відповідно, до побудови ширших дерев пошуку порівняно зі стандартним процесом пошуку MCTS. І навпаки, від'ємне значення C_{DWC1} буде примушувати алгоритм MCTS до побудови глибших і вузьких піддерев, оскільки значення DWC1 віднімається від поточного значення формули UCB1. Такий вплив є подібним до впливу збільшення/зменшення коефіцієнта C_{UCB1} стандартної формули UCB1, розглянутого в [1], але відмінність варіанта TSC1 полягає в тому, що цей вплив на процес побудови дерева залежить саме від форми поточного піддерева, в той час як у стандартній формулі UCB1 не залежить.

Як ще один варіант врахування форми піддерев під час побудови дерева пошуку, якщо це буде більш прийнятним для певної гри, можна запропонувати використання оберненого відношення ширини піддерева до його глибини ($NodeWidth/NodeDepth$).

Приклад, наведений на рис. 1, показує як використання способу TSC1, який базується на використанні критерію DWC1, виконує вибір вужчого піддерева від поточного кореня для подальшого розширення на відміну від використання стандартної формули UCB1. В той час, як стандартна формула UCB1 вибере гілку #1 для продовження побудови дерева пошуку, використання способу TSC1 робить зміщення цінності всіх гілок і вибере гілку піддерева #3 з вужчою формою.

Треба зазначити, що відношення $NodeDepth/NodeWidth$ не може бути від'ємним і, таким чином, зміщення процесу пошуку в дереві порівняно зі стандартним алгоритмом відбувається завжди, тобто незалежно від того, який параметр поточного дерева є більшим, глибина чи ширина. В залежності від знаку коефіцієнта C_{DWC1} таке зміщення відбувається або в сторону побудови ширших

або в сторону побудови вужчих піддерев. Одним з варіантів застосування способу TSC, що базується на критеріях DWCs, може бути його спільне використання з іншими способами покращення пошуку методом MCTS. Наприклад, ці критерії можна використовувати разом зі способами, що ґрунтуються на критерії «ефективний степінь розгалуження» [8] або підході FPU [9], для яких критерії DWCs можуть точніше показати, що саме треба збільшувати у дереві, глибину, чи ширину. Варіанти контролю форми дерева пошуку TSC2 та TSC3 працюють аналогічно способу TSC1. Детальніше вони описані у [5].

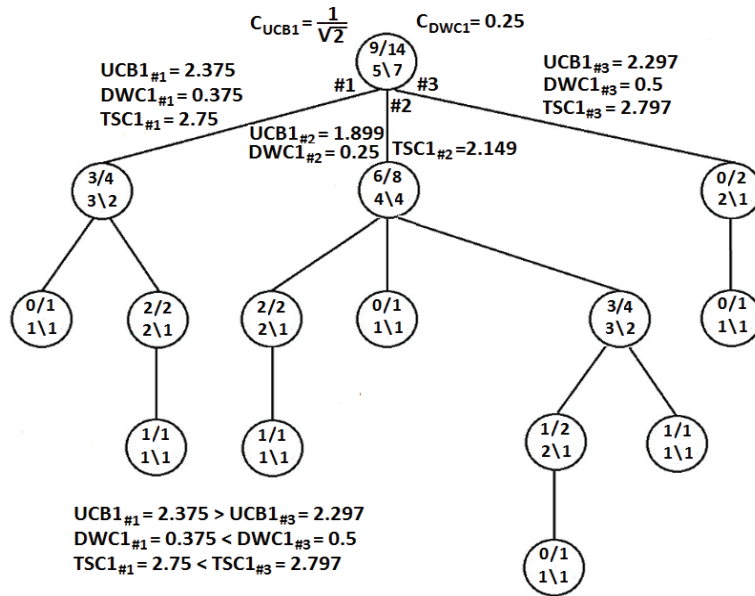


Рис. 1. Механізм вибору наступної гілки дерева способом TSC1

Порівняння побудови дерева MCTS з контролем форми дерева і без

З метою порівняння процесу побудови дерева пошуку стандартним способом UCT (Upper Confidence bounds applied to Trees) реалізації MCTS [1] і способом TSC3 із застосуванням контролю за формою дерева отримані дерева, побудовані після виконання 10 (рис. 2) та 20 (рис. 3) ітерацій процесу пошуку. З цих рисунків чітко видно, що способом TSC3, не змінюючи загального асиметричного принципу побудови, можна скерувати побудову дерева до ширшої і більш регулярної форми, в результаті чого на кожному рівні дерева аналізом буде охоплюватись більша кількість вершин (потенційних ходів). Після 10 ітерацій на першому рівні дерева за способом UCT проаналізовано 2 вершини (хоча одна з них 3 рази), в той час як способом TSC3 на тому ж рівні виконано аналіз для 4 вершин. Після 20 ітерацій ситуація не змінилась: UCT продовжує аналізувати тільки дві вершини першого рівня, а TSC3 — всі чотири з різним ступенем детальності.



Рис. 2. Форма дерев, отриманих після виконання 10 ітерацій пошуку: а — стандартним способом UCT; б — способом з контролем форми дерева TSC3

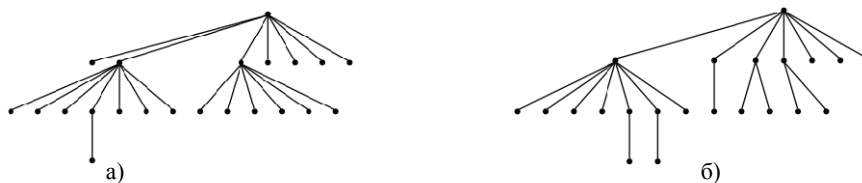


Рис. 3. Форма дерев, отриманих після виконання 20 ітерацій пошуку: а — стандартним способом UCT; б — способом з контролем форми дерева TSC3

Насамкінець наведемо результати виконання гри Connect Four [10] для двох випадків: перший випадок — обидва гравці грають за стандартним способом UCT; другий випадок — перший гравець грає за способом з контролем форми дерева, а другий гравець — за стандартним способом UCT (табл. 1). Виконано по 10 ігор для п'яти значень коефіцієнта C_{UCB1} (0,1, 0,2, 0,5, 1,4, 2,0) за п'яти варіантів кількості ітерацій повторення процесу пошуку (1000, 10000, 50000, 100000, 300000), які фактично відповідають часу «роздумів» способу над наступним ходом. Разом виходить 250 ігор. Коефіцієнти способу TSC3 при всіх іграх були однаковими і дорівнювали $C_{DWC3} = 0,125$; $P = 0,05$.

Таблиця 1

Результати застосування способу TSC3

Спосіб	Кількість вигравів гравця 1	Кількість вигравів гравця 2	Кількість нічиїх	Кількість ігор
UCT vs UCT	112	111	27	250
TSC3 vs UCT	123	109	18	250

Як видно з отриманих результатів, обидва гравці, що грали між собою способом UCT, виграли один у одного приблизно рівну кількість разів, в той час, як гравець, який грав способом TSC3, отримав більш ніж на 10 % вигравів більше (123 проти 109).

Зазначимо, що для інших ігор може дати кращий результат корекція форми дерева у сторону побудови вужчих дерев, а не ширших, що також може бути реалізовано на основі запропонованого способу.

Висновки

Покращення пошуку в дереві з використанням методу MCTS може бути виконано, використовуючи багато різних підходів, одним з яких є врахування форми дерева, що будується, під час роботи пошуку. Для одних задач метод MCTS для ефективного пошуку буде дерева ширшої форми, а для інших — глибшої [6]. Використання способів контролю поточної форми дерева пошуку не змінює загальний асиметричний принцип побудови дерева пошуку MCTS, але дозволяє будувати, якщо це дає кращі результати для конкретної гри, відносно більш регулярні дерева, що дозволяє краще охоплювати аналізом рідко відвідувані вершини (тобто вершини, на яких стандартний алгоритм може «помилятися») і, в результаті, за правильного підбору коефіцієнтів, покращити ефективність пошуку. Надалі в цьому напрямку можуть бути дослідження використання різних методів машинного навчання з метою динамічної корекції коефіцієнтів у критеріях DWCs, а також вплив використання цих методів на ефективність роботи пошуку методом MCTS.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. A Survey of Monte Carlo Tree Search Methods / [Cameron Browne, Edward Powley, Daniel Whitehouse, and others] // IEEE Trans. on Computational Intelligence and AI in Games. — March 2012. — Vol. 4. — No. 1. — P. 1—49.
2. Марченко О. І. Структура та критерії класифікації способів реалізації та покращення пошуку по дереву методом Монте-Карло / О. І. Марченко, О. О. Марченко, М. М. Орлова // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. — 2015. — № 21. — С. 51—57.
3. Марченко О. І. Класифікація способів реалізації та покращення пошуку по дереву методом Монте-Карло / О. І. Марченко, О. О. Марченко, М. М. Орлова // Штучний інтелект. — 2016. — № 2 (72). — С. 59—69.
4. Alan Levinovitz. The Mystery of Go, the Ancient Game That Computers Still Can't Win. [Електронний ресурс] / Alan Levinovitz. — Режим доступу: <https://www.wired.com/2014/05/the-world-of-computer-go/>.
5. Marchenko O. I. Monte-Carlo Tree Search with Tree Shape Control / Oleksandr I. Marchenko, Oleksii O. Marchenko // 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). Conference Proceedings. May 29 — June 2, 2017., Kyiv, Ukraine. — 2017. — P. 812—8173.
6. Hilmar Finnsson. Game-Tree Properties and MCTS Performance / Hilmar Finnsson and Yngvi Björnsson // GIGA 2011: Proceedings of the 2nd International General Game Playing Workshop, 2011. Pp. 23—30.
7. Марченко О. О. Критерій «глибина-ширина» для контролю форми дерева пошуку при використанні методу Монте-Карло / О. І. Марченко, О. О. Марченко // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. — 2016. — № 24—25. — С. 42—47.
8. G. M. J. Williams. Determining Game Quality Through UCT Tree Shape Analysis / G. M. J. Williams // M.S. Thesis, Imperial Coll., London. — 2010.
9. Yizao Wang. Modifications of UCT and sequence-like simulations for Monte-Carlo Go / Yizao Wang, Sylvain Gelly // Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Games. — 2007. — P. 175—182.
10. Connect Four [Електронний ресурс]. — Режим доступу: https://en.wikipedia.org/wiki/Connect_Four.

Марченко Олексій Олександрович — аспірант кафедри системного програмування і спеціалізованих комп'ютерних систем, email: alex.mar4elo@gmail.com ;

Марченко Олександр Іванович — канд. техн. наук, доцент, доцент кафедри системного програмування і спеціалізованих комп'ютерних систем, email: oleksandr.i.marchenko@gmail.com ;

Щербина Богдан Олександрович — аспірант кафедри системного програмування і спеціалізованих комп'ютерних систем, email: bogdan.shcherbina11@gmail.com .

Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ

O. O. Marchenko¹

O. I. Marchenko¹

B. O. Shcherbyna¹

Building of Searching Tree by the Technique Using Monte-Carlo Tree Search Method with Tree Shape Control

¹National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»

The paper describes in details mechanism of the tree-shape control technique influence onto searching process with usage of the Monte-Carlo tree search (MCTS) method. Comparison of the tree-shape control technique for MCTS implementation with the standard technique of this method implementation is done. It is shown that in case when coefficients are taken well, usage of the MCTS implementation with tree-shape control follows in increasing of the search performance and better decisions.

Keywords: game trees, tree search, Monte-Carlo method, MCTS, MCTS parallelizing techniques, grid-systems.

Marchenko Oleksii O. — Post-Graduate Student of the Chair of System Programming and Specialized Computer Systems, email: alex.mar4elo@gmail.com ;

Marchenko Oleksandr I. — Cand. Sc. (Eng.), Assistant Professor of the Chair of System Programming and Specialized Computer Systems, email: oleksandr.i.marchenko@gmail.com ;

Shcherbyna Bohdan O. — Post-Graduate Student of the Chair of System Programming and Specialized Computer Systems, email: bogdan.shcherbina11@gmail.com

A. A. Марченко¹

A. И. Марченко¹

Б. А. Щербина¹

Построение дерева поиска способом с использованием метода Монте-Карло и контролем формы дерева

¹Национальный технический университет Украины

«Киевский политехнический институт имени Игоря Сикорского»

Детально описан механизм влияния способа контроля формы дерева поиска на процесс выполнения поиска в дереве методом Монте-Карло (MCTS). Выполнено сравнение способа реализации MCTS с контролем формы дерева со стандартным способом реализации этого метода. Показано, что при правильном подборе коэффициентов использование способа реализации MCTS с контролем формы дерева приводит к повышению результативности поиска и, в результате, к принятию лучших решений.

Ключевые слова: деревья игр, поиск в дереве, метод Монте-Карло, MCTS, способы улучшения MCTS.

Марченко Алексей Александрович — аспирант кафедры системного программирования и специализированных компьютерных систем, email: alex.mar4elo@gmail.com ;

Марченко Александр Иванович — канд. техн. наук, доцент, доцент кафедры системного программирования и специализированных компьютерных систем, email: oleksandr.i.marchenko@gmail.com ;

Щербина Богдан Александрович — аспирант кафедры системного программирования и специализированных компьютерных систем, email: bogdan.shcherbina11@gmail.com