

УДК 004.023

Марченко О.О., аспірант, Марченко О.І., к.т.н., доцент,
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

ЛОГАРИФМІЧНИЙ КРИТЕРІЙ КОНТРОЛЮ ФОРМИ ДЕРЕВА ДЛЯ ПОКРАЩЕННЯ ПОШУКУ ПО ДЕРЕВУ МЕТОДОМ МОНТЕ-КАРЛО

Марченко О.О., Марченко О.І. Логарифмічний критерій контролю форми дерева для покращення пошуку по дереву методом Монте-Карло. Апробація запропонованих авторами раніше критеріїв типу «глибина-ширина» при контролі форми дерева пошуку та їх використання для покращення пошуку по дереву методом Монте-Карло (MCTS) з'ясували деякі слабкі місця використання цих критеріїв. У статті, на основі результатів останніх досліджень, для контролю форми дерева пошуку пропонується новий логарифмічний критерій, який дозволяє більш точно оцінити процес побудови та форму дерева пошуку MCTS. Передбачається, що запропонований критерій дозволить точніше вибрати потенційно кращі варіанти продовження такого пошуку.

Ключові слова: задачі штучного інтелекту, дерева ігор, пошук в дереві, метод Монте-Карло, MCTS, методи покращення MCTS.

Марченко А.А., Марченко А.И. Логарифмический критерий контроля формы дерева для улучшения поиска по дереву методом Монте-Карло. Апробація предложених авторами ранее критериев типа «глубина-ширина» при контроле формы дерева поиска и их использование для улучшения поиска по дереву методом Монте-Карло (MCTS) выявили некоторые слабые места использования этих критериев. В статье, на основе результатов последних исследований, для контроля формы дерева поиска предлагается новый логарифмический критерий, который позволяет более точно оценить процесс построения и форму дерева поиска MCTS. Предполагается, что предложенный критерий позволит точнее выбирать потенциально лучшие варианты продолжения такого поиска.

Ключевые слова: задачи искусственного интеллекта, деревья игр, поиск в дереве, метод Монте-Карло, MCTS, методы улучшения MCTS.

Marchenko O.O., Marchenko O.I. Logarithmic criterion for tree shape control for improvement of the Monte-Carlo tree search method. Approbation of the proposed earlier by authors "depth-width" kind criteria for controlling of search tree shape and applying of the criteria for Monte-Carlo tree search (MCTS) method improvement revealed some weak points of these criteria exploitation. The paper, basing on the results of the latest investigations, proposes a new logarithmic criterion for tree shape control which allows to estimate MCTS tree building process and tree shape more precisely. It is supposed that this criterion would allow more accurate selecting potentially better variants for continuation of such kind of search.

Keywords: artificial intelligence tasks, game trees, tree search, Monte-Carlo method, MCTS, MCTS improvement methods.

Рис. 5. Лім.8.

Постановка наукової проблеми.

Однією з найбільш поширених форм подання інформації в задачах штучного інтелекту є форма у вигляді дерев послідовних рішень, а одним з методів штучного інтелекту, що був запропонований порівняно нещодавно і призначений для виконання швидкого та високоефективного пошуку правильних рішень у дереві інформації, є метод пошуку в дереві з використанням методу Монте-Карло (Monte Carlo Tree Search – MCTS) [1]. Особливо добре він зарекомендував себе в задачах з дуже великим ступенем розгалуження пошуку, таких, як гра Го (ступінь розгалуження дорівнює 250). Завдяки цьому методу відбувся прорив у створенні комп'ютерного гравця в Го, рівного по силі гри найкращим професійним гравцям. Для прискорення та підвищення ефективності процесу пошуку методом MCTS, його вдосконалення можливо виконувати на різних етапах його загальної схеми, і за час існування цього методу було створено вже багато його варіантів та модифікацій [1, 2, 3], але потужність концепції методу MCTS залишає ще дуже великий простір для нових досліджень з метою покращення пошуку в задачах штучного інтелекту.

Аналіз досліджень.

Найбільш фундаментальною узагальнюючою роботою з методу MCTS є огляд 2012 року, підготовлений десятима провідними вченими під керівництвом Камерона Броуне, що досліджують цей метод [1], але протягом наступних років з'явилися не тільки нові модифікації та покращення існуючих способів, але й принципово нові варіанти реалізації MCTS. Ще один погляд на структуру та взаємозв'язки між ідеями існуючих способів покращення MCTS був запропонований авторами у

вигляді класифікації [2, 3]. В результаті виконаного аналізу способів покращення MCTS був визначений напрям в дослідженнях MCTS, якому приділялось ще недостатньо уваги попередніми дослідниками, а саме дослідження залежності ефективності пошуку MCTS від форми дерева, а також знаходження можливості покращення пошуку методом MCTS спираючись на певні критерії оцінювання поточної форми дерева пошуку.

У [4, 5] авторами були запропоновані критерії типу «глибина-ширина» DWCs (Depth-Width Criteria) для контролю форми дерева пошуку при використанні методу Монте-Карло, а також спосіб покращення ефективності роботи MCTS. Цей спосіб вдосконалює формулу UCSB1 стандартного варіанту UCT (Upper Confidence bounds applied to Trees) реалізації MCTS [1] використанням ще декількох доданків, які встановлюють залежність остаточного значення цієї формули від форми дерева пошуку на основі зазначених критеріїв. Такий спосіб покращення MCTS був названий пошуком по дереву методом Монте-Карло з контролем форми дерева (Monte-Carlo Tree Search with Tree Shape Control – MCTS-TSC) [5] і 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. Наприклад, досить очевидно, що для ігор з великим ступенем розгалуження буде будуватись ширше дерево, ніж для ігор з малим ступенем розгалуження. Крім того, Гілмар Фінссон та Інгві Бьорнссон досліджували ситуацію, чи віддає MCTS перевагу деревам з певною формою, більш глибоким чи більш широким, чи дотримує якогось певного балансу між глибиною і шириною. Вони з'ясували, що такі «вподобання» алгоритму MCTS залежать власне від самої гри, а також те, що, навіть, невеликі нюанси в правилах гри та системі підрахунку балів можуть змінювати структуру дерева пошуку, якій буде віддавати перевагу MCTS [6]. Тому, якщо для певних задач відомо, яка форма дерева MCTS (більш глибока чи більш широка) є ефективнішою, то враховуючи таку наявну початкову інформацію, для цих задач з'являється можливість доцільного виконання динамічного контролю форми дерева за критерієм «глибина-ширина», направляючи подальшу побудову дерева пошуку у потрібному напрямку.

Метою даної роботи є продовження дослідження залежності ефективності пошуку MCTS від форми дерева, а також знаходження можливості покращення пошуку методом MCTS за допомогою оберненого впливу, тобто скерування процесу побудови дерева пошуку до більш широкої чи більш глибокої форми, в залежності від характеру конкретної гри (задачі), на основі певних критеріїв, що дозволяють оцінювати форму поточного дерева пошуку і вибирати потенційно кращі варіанти продовження такого пошуку.

Критерії типу «глибина-ширина» для контролю форми дерева пошуку MCTS

Константа C_{UCB1} основної формули UCSB1 для оцінювання перспективності наступних ходів в методі MCTS [1, 5], фактично є параметром, який контролює форму дерева пошуку MCTS. При менших значеннях C_{UCB1} створюється глибше і вужче дерево пошуку, в той час, як при більших значеннях C_{UCB1} буде побудоване менш глибоке але більш широке дерево MCTS [7]. Рівняння UCSB1 автоматично фокусує пошук на поточних найкращих варіантах подальших ходів одночасно з наданням можливості достатнього дослідження для того, щоб знайти потенційно найкращий хід, якого в дереві ще нема.

Важливо зазначити, що в даній роботі аналіз критеріїв оцінювання форми дерева пошуку навмисно виконується без враховування значень формули UCSB1 для вершин дерева, оскільки наша мета оцінити саме потенційний вплив форми дерева, щоб зробити відповідне зміщення значення UCSB1 в більшу чи меншу сторону в залежності від поточної форми дерева і незалежно від поточних оцінок вершин за формулою UCSB1.

Запропоновані у [4, 5] критерії оцінювання форми дерева DWCs обчислюються для кожної вершини на основі двох властивостей форми дерева – глибини D та ширини W піддерева цієї вершини. Обчислення значень цих властивостей виконується за наступними правилами [4]:

- у всіх термінальних вершин дерева (листіків) і глибина, і ширина їх піддерев дорівнюють одиниці;
- глибина піддерева нетермінальної вершини обчислюється як максимальне значення зі значень глибин піддерев її нащадків плюс одиниця;
- ширина піддерева нетермінальної вершини обчислюється як сума значень ширин піддерев її

нащадків.

Для оцінювання та контролю форми всіх піддерев поточного дерева пошуку в [4, 5] авторами було запропоновано два критерії

$$DWC1 = C_{DW} * D / W,$$

$$DWC2 = C_{DW} * (C_D * D - W) / W,$$

де C_{DW} – коефіцієнт визначення форми дерева (пропорції глибини до ширини),

C_D – додатковий коефіцієнт глибини дерева.

Призначення коефіцієнтів C_{DW} та C_D у цих критеріях є подібним до призначення коефіцієнту C_{UCB1} у основній формулі UCB1 [1, 5] метода MCTS. Значення критеріїв $DWC1$ та $DWC2$ використовуються як доданки до значення формули UCB1 і можуть використовуватись як окремо, так і разом.

Подальші дослідження показали, що просте відношення глибини піддерев до їх ширини не завжди достатньо точно відображає логіку корекції формули вибору наступної вершини в залежності від форми піддерев поточної вершини, де приймається рішення. Розглянемо приклад характерного випадку. На рисунку 1 показана ситуація, коли у поточної вершини DN (Decision Node), де потрібно прийняти рішення про вибір наступного ходу, є дві вершини-дитини CN_1 та CN_2 (CN – Child Node), серед яких потрібно вибрати. Перша цифра відповідає глибині піддерева, а друга цифра – ширині піддерева. Щоб відрізнити ситуацію, коли у вершині показаний критерій «глибина-ширина», а не критерій «кількість вигравів/кількість моделювань», на відміну від останнього, для розділу двох значень використовується символ «\» (back-slash), а не «/».

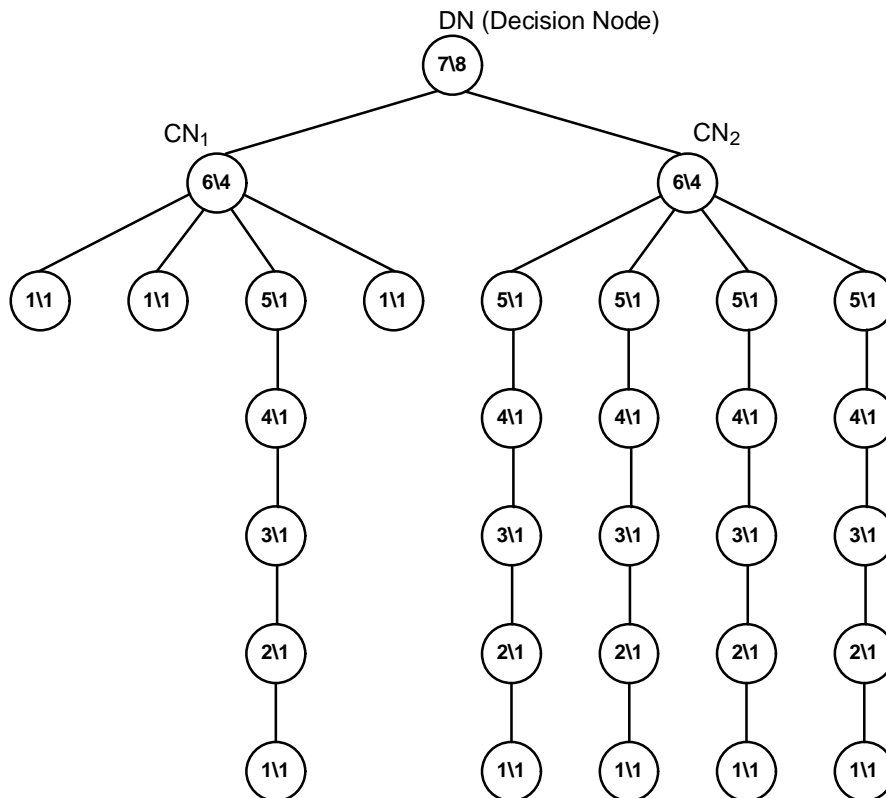


Рис.1. Приклад недостатньої точності критеріїв $DWC1$ та $DWC2$.

Як видно, піддерева вершин CN_1 та CN_2 різні, а значення глибини та ширини, а відповідно і відношення глибина/ширина, що є визначальним для остаточної оцінки форми дерева, у цих вершин є однаковими 6/4, що є некоректним з точки зору логіки враховування форми дерева в алгоритмі MCTS. У кожній з цих вершин однакова кількість вершин-дітей (по чотири), але якщо у вершини CN_2 всі

піддерева досліджені однаково і мають невисоку оцінку отримання виграшу, то вершина CN_1 має тільки одну таку саму вершину, а три інших є ще майже не дослідженими, а відтак є більша ймовірність, що саме серед цих трьох дітей знаходиться кращий варіант для продовження гри. Тому, згідно логіки MCTS, вершина CN_1 є більш цінним варіантом продовження гри, але прості критерії DWC1 та DWC2 цього не враховують. Цей випадок показує, що критерій форми дерева повинен враховувати не тільки глибину та ширину піддерев, а й кількість вершин у них.

Подальші дослідження показали, що в цих та інших подібних випадках логіку корекції формули вибору наступної вершини в залежності від форми піддерев поточної вершини, де приймається рішення, більш коректно відображають критерії типу «глибина-ширина», значення яких мають логарифмічні залежності від відношення глибини до ширини піддерев та кількості вершин у них.

В результаті розгляду та аналізу множини подібних випадків, автори прийшли до схеми враховування форми дерева пошуку MCTS, показаній на рисунку 2.

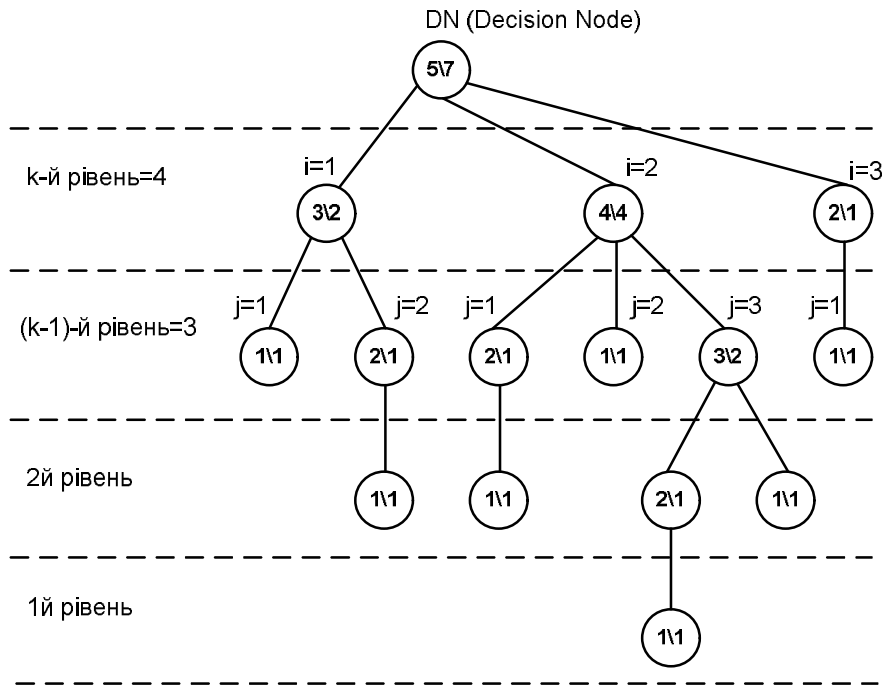


Рис.2. Схема враховування форми дерева пошуку MCTS.

Згідно показаної схеми пропонується виконувати обчислення критерію форми піддерев дітей вершини прийняття рішення DN, тобто піддерев вершин CN_i (що знаходяться на k-му рівні) на основі їх власних значень глибини та ширини D_i^k та W_i^k , а також значень глибини та ширини D_j^{k-1} та W_j^{k-1} всіх дітей (що знаходяться на (k-1)-му рівні) кожної з вершин CN_i за логарифмічною формулою, показаною на рисунку 3.

$$DWC_LOGB_i^k = C_{TSC} * \left(\log_B \left(C_{DW} * \frac{D_i^k}{W_i^k} \right) + \sum_{j=1}^{NB_i^k} \frac{\log_B \left(C_{DW} * \frac{D_j^{k-1}}{W_j^{k-1}} \right)}{B * NB_i^k} - C_{NN} * \log_B(NN_i^k) \right) \quad (1)$$

Рис.3. Формула враховування форми дерева пошуку MCTS.

У формулі (1) величини мають наступне значення:

k – рівень вершини рахуючи від листка;

i – номер вершини-дитини поточного кореня дерева гри;

j – номер вершини-дитини i -ї вершини;

DWC_LOG – значення логарифмічного (по основі B) критерію типу «глибина-ширина» i -ї вершини-дитини поточного кореня, яка знаходиться на k -му рівні;

C_{TSC} – загальний коефіцієнт масштабування впливу форми дерева;

C_{DW} – коефіцієнт визначення форми дерева (пропорції глибини до ширини);

C_{NN} – коефіцієнт масштабування впливу кількості вершин у піддеревах;

B – степінь розгалуження (Branching factor) гри;

N – степінь розгалуження i -ї вершини-дитини (Node Branching factor) поточного кореня, яка знаходиться на k -му рівні;

N – кількість вершин у піддереві i -ї вершини-дитини поточного кореня, яка знаходиться на k -му рівні;

l та w – глибина та ширина піддерева i -ї вершини-дитини поточного кореня, яка знаходиться на k -му рівні;

D_j та W_j^i – глибина та ширина піддерева j -ї дитини i -ї вершини-дитини поточного кореня (j -ті вершини знаходяться на рівень нижче від i -х вершин, тобто на $(k-1)$ -му рівні).

Значення критерію $DWC_LOGB_i^k$ для вершин-листіків не обчислюється і приймається рівним нулю, оскільки ці вершини піддерев ще не мають.

Для підтвердження коректності даної формули для відображення випадку, показано на рис.1, наведемо те ж саме дерево (рисунок 4), в вершинах якого знаходяться обчислені значення логарифмічного критерію $DWC_LOGB_i^k$. Обчислення значень вершин для рисунка 4 були виконані при наступних значеннях коефіцієнтів: $C_{TSC} = 1$; $C_{DW} = 1$; $C_{NN} = 0.1$; $B = 7$. Слід зауважити, що дослідження проводились при різних значеннях коефіцієнтів C_{TSC} , C_{DW} та C_{NN} , а ці значення взяті для прикладу. Значення B взяті рівним 7, оскільки дослідження проводились на прикладі гри Connect Four [8], яка має степінь розгалуження 7.

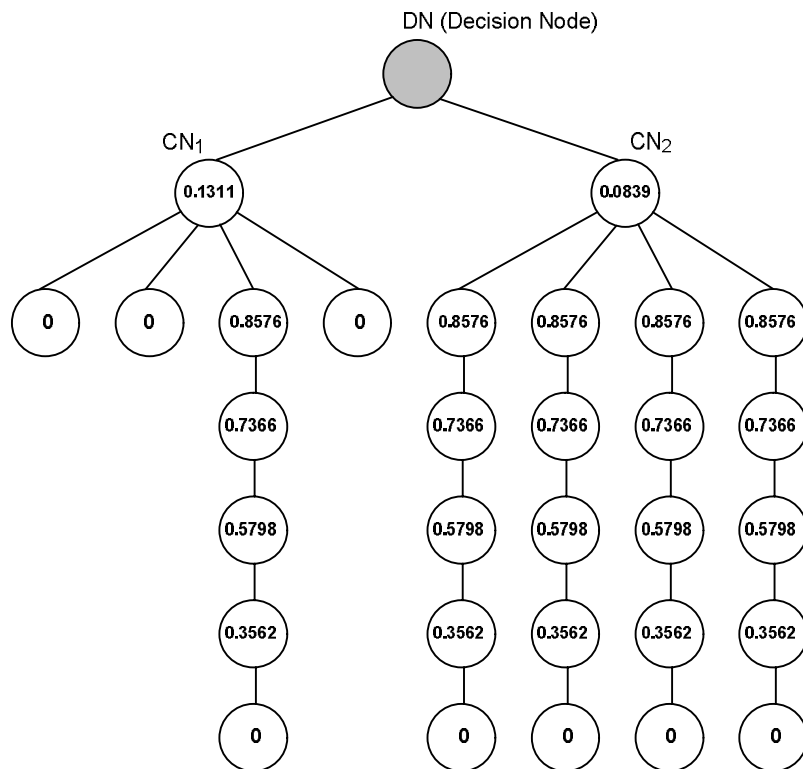


Рис.4. Дерево MCTS з обчисленими значеннями логарифмічного критерію $DWC_LOGB_i^k$.

З даного рисунка видно, що, на відміну від значень вершин рисунка 1, вершини CN_1 та CN_2 мають різні значення, причому значення вершини CN_1 є більшим ніж значення вершини CN_2 , що відповідає прийнятій для цього прикладу логіці оцінювання вершин з точки зору форми піддерев, тобто віддавати перевагу менш дослідженим піддеревам. Якщо для іншої гри (задачі) буде потрібна протилежна логіка оцінювання, то останній доданок $C_{NN} * \log_B(NN_i^k)$ формули (1), який залежить від кількості вершин в піддереві, потрібно не віднімати, а додавати. Більш того, якщо для динамічного налаштування процесу оцінювання впливу форми піддерев на ефективність пошуку використовувати методи машинного навчання, то, в додаванням/відніманням останнього доданку можна керувати в залежності від статистики зібраної машинним навчанням.

Іншою важливою відмінністю формули (1) від попередніх критеріїв $DWC1$ та $DWC2$ є використання логарифмічної залежності від відношення глибини піддерев до його ширини. Логарифмічна функція дозволяє досягти двох переваг, в порівнянні з більш простими $DWC1$ та $DWC2$.

Перша перевага полягає в тому, що врахування знаку доданків $\log_B\left(C_{DW} * \frac{D_i^k}{W_i^k}\right)$ та $\log_B\left(C_{DW} * \frac{D_j^{k-1}}{W_j^{k-1}}\right)$

для корекції форми поточного дерева відбувається автоматично в залежності від того, яка з величин, глибина чи ширина, є більшою (враховуючи заданий коефіцієнт форми C_{DW}). Другою перевагою є те, що глибина піддерев впливає на остаточне значення логарифмічного критерію оцінювання форми дерева не лінійно, як у критеріях $DWC1$ та $DWC2$, а більш згладжено. Така властивість більш точно відповідає загальному асиметричному характеру поведінки алгоритму MCTS при побудові дерева пошуку, оскільки у випадках, коли значення головної оцінки вершин згідно формули UCB1 буде стимулювати побудову більш глибокої гілки, то вплив другорядного фактору форми дерева буде зменшуватись. Основа логарифму приймається рівною степені розгалуження гри B , оскільки це дозволяє ввести в формулу (1) залежність від складності конкретних ігор. Крім того, це дозволяє врахувати той факт, що при більших значеннях B кількість доданків в формулі (1) може бути також більшою, в результаті чого сума може стрімко зростати, а відповідне зростання значення основи логарифму буде компенсувати вплив збільшення кількості доданків, оскільки максимально можлива кількість доданків дорівнює B .

Апробація логарифмічного критерію

Спосіб покращення MCTS-TSC з використанням логарифмічного критерію (1) був апробований на грі Connect Four [8], степінь розгалуження якої $B=7$, для трьох випадків: 1) обидва гравці грають за стандартним способомUCT; 2) перший гравець грає за способом з контролем форми дерева TSC, а другий гравець – за стандартним способомUCT; 3) перший гравець грає за способомUCT, а другий гравець – за способомTSC. Для прикладу наведемо результати, отримані при наступних параметрах: кількість ігор для кожного випадку – 10; значення коефіцієнта $C_{UCB1}=0.55$ (найкраще значення отримане для гравця UCT під час експериментів); випадки кількості ітерацій повторення процесу пошуку на один хід були взяті 100000, 150000, 200000, 250000, 500000, 1000000 (кількість ітерацій фактично відповідає часу «роздумів» способу над наступним ходом). Разом виходить по 60 партій для кожного з досліджених варіантів гри. Для реалізації контролю за формою дерева пошуку в способі MCTS-TSC був використаний наведений вище логарифмічний критерій (1) з коефіцієнтами $C_{TSC}=0.0001$, $C_{DW}=10$, $C_{NN}=0.0001$. Результати виконання гри для цих випадків наведені на рисунку 5.

Як видно з отриманих результатів, другий гравець завжди виявився сильнішим, що, очевидно, пов'язано із загальними особливостями роботи пошуку методом MCTS для гри Connect Four. У партіях, де обидва гравці грали між собою за стандартним способомUCT, другий гравець вигравав у 2 рази частіше. Коли до того самого першого гравця додали контроль за формою дерева способомTSC, він став програвати рідше (в 1.6 рази), а коли TSC-контроль додали до другого гравця, то він став вигравати у першого вже майже у 4 рази частіше, а не у 2 рази. Тобто, додавання механізму контролю за формою дерева пошуку, в обох ситуаціях привело до підсилення найкращого гравця гри Connect Four, реалізованого «чистим» стандартним способомMCTS-UCT.

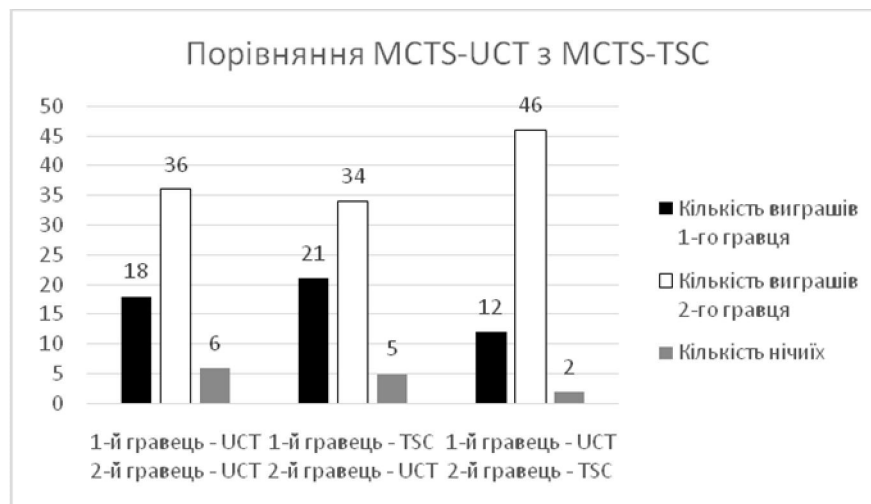


Рис. 5. Порівняння результатів гри гравців реалізованих способами MCTS-UCT та MCTS-TSC

Висновки

У статті запропонований новий логарифмічний критерій, який дозволяє оцінити форму поточного дерева пошуку MCTS та скерувати процес його подальшої побудови більш точно, ніж раніше визначені критерії. Використання способів контролю поточної форми дерева пошуку, не змінюючи загального асиметричного характеру побудови дерева пошуку, може скеровувати побудову дерева гри в сторону більш широких або більш глибоких дерев, відносно стандартного варіанту реалізації MCTS, що дозволяє в певних ситуаціях краще охоплювати аналізом вершини, яким стандартний алгоритм приділяє менше уваги, і, в результаті, при правильному підборі коефіцієнтів, дозволяє покращити ефективність пошуку. Подальшими напрямками досліджень і варіантами застосування способу MCTS-TSC, що базується на критеріях типу «глибина-ширина», можуть бути: сумісне використання цього способу разом з іншими способами покращення пошуку методом MCTS; дослідження ефективності способу контролю форми дерева пошуку на інших іграх з меншими та більшими степенями розгалуження; використання різних методів машинного навчання з метою динамічної корекції коефіцієнтів у критеріях типу «глибина-ширина».

1. Cameron Browne. 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. – vol. 4. – no. 1. – March 2012. – 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. Марченко О.О. Критерій «глибина-ширина» для контролю форми дерева пошуку при використанні методу Монте-Карло. / Марченко О.І., Марченко О.О. // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – 2016. – № 24-25. – С.42-47.
5. Oleksandr I. Marchenko. 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. Joel Veness, Kee Siong Ng, Marcus Hutter, William Uther, David Silver. A Monte-Carlo AIXI Approximation // Journal of Artificial Intelligence Research 40 (2011), pp.95-142.
8. Connect Four // [Електронний ресурс]. – Режим доступу: https://en.wikipedia.org/wiki/Connect_Four