Monte-Carlo Tree Search for Artificial General Intelligence in Games

Valorization

> **Please note**: the *Regulations for obtaining the doctoral degree* of Maastricht University, dated 1 September 2018, require the addition of a *valorization addendum* to each dissertation. "While the valorisation addendum is included in the thesis, it is not regarded as part of the thesis and must not be taken into account in the assessment of the thesis by the Assessment Committee and the Defence Committee".[1]

Knowledge valorization is "the process of creating value from knowledge, by making knowledge suitable and/or available for social (and/or economic) use and by making it suitable for translation into competing products, services, processes and new activities". This addendum discusses various ways in which the research presented in the thesis might contribute to create social and economic value. We mainly discuss the value that can be created by advancing research on Monte-Carlo Tree Search (MCTS), giving also some insights on how value can be created by using MCTS to advance research towards Artificial General Intelligence (AGI). First, valorization opportunities in game-related domains are discussed. Next, the discussion is extended to other domains. Finally, we argue that valorisation opportunities can arise from research on MCTS that moves in the direction of AGI.

Games

This thesis uses (video) games as a test bed for MCTS, therefore the gaming industry is the first application domain that comes to mind where the presented research has the potential to create value. Video games are getting more and more realistic and sophisticated, demanding smarter AI characters that are able to deal with increasingly complex environments. Moreover, if AI characters show unrealistic behaviors, the interest of the players in the game will decrease and the game will lose its entertainment value. MCTS can be used to model believable AI characters in commercial applications, and there have been already some examples of its success on this task. In the *Spades* mobile phone game by AI Factory MCTS has been used to create engaging AI opponents and allies (Whitehouse *et al.*, 2013), and has later been improved to emulate human play (Baier *et al.*, 2018). Moreover, MCTS has been used in the development of the AI system for the on-line turn-based strategy

[1] `https://www.maastrichtuniversity.nl/support/phds`, retrieved 4 September 2019.

game *Prismata*, by Lunarch Studios (Churchill and Buro, 2015), and in the development of the AI system of *Total War: Rome II*, a strategy game developed by Creative Assembly and published by Sega (Thompson, 2018).

Potential applications of MCTS to games are not only limited to the development of AI characters for the game industry. MCTS can also be useful during the process of game design and game content generation. A possibility is to use it to evaluate generated games. For instance, Browne (2012) computes game evaluation metrics using the performance of UCT-based search on such games. Other approaches apply MCTS to directly generate games or game content. For example, it has been used to generate initial tile placements for Pentominoes puzzles (Browne, 2013) and to generate Sokoban puzzles (Kartal, Sohre, and Guy, 2016). In addition, it has been applied to automatically generate narratives (Kartal, Koenig, and Guy, 2014), which could be used in virtual environments or in video games to dynamically generate new plot lines.

Another application domain for which MCTS can create value are serious games, i.e. games that combine the playfulness of video games with a serious purpose, such as training, education or rehabilitation. An example of a serious game that makes use of MCTS is given by Sanselone *et al.* (2014), which use this search technique to control the behavior of virtual characters in a simulated surgery room environment for medical staff training. Another example is the work of Hocine, Gouaïch, and Cerri (2014), which use MCTS in the development of a game aimed at rehabilitating patients that suffered from a stroke.

Other Domains

The potential to create value from research on MCTS and its enhancements is not limited to games alone. MCTS is a suitable technique for domains that involve planning, optimization and decision making, and there are many examples of successful application of MCTS and its variants in a different number of fields.

In robotics, for instance, MCTS-based algorithms have been used for various purposes, ranging from controlling the movement of robots to managing their perception. Using MCTS variants, Goldhoorn *et al.* (2014) design robots that are able to find and follow people in a real-world scenario, Zech, Xiong, and Piater (2015) propose an optimization method that improves the precision of robotic grasps and Kartal (2015) addresses the multi-robot patrolling problem. Moreover, Patten, Martens, and Fitch (2018) propose an object classification mechanism for robots that operate in an outdoor environment and Best *et al.* (2019) design a decentralized mechanism that controls multiple robots with active perception of the world.

Other applications of MCTS can be found in the medical field. For example, MCTS has been used to schedule elective admissions of patients to the hospital (Van Eyck *et al.*, 2013; Zhu *et al.*, 2014) and to improve the process of selecting personalized health-care services in order to reduce the risk of re-hospitalization of patients (Laschet, 2014). In addition, Pinheiro, Kybic, and Fua (2017) design a mechanism based on MCTS to match graph structures to medical images.

Space exploration is another example of task for which the use of MCTS has been considered. Hennes and Izzo (2015) apply MCTS to automatically plan interplanetary trajectories, while Song and Gong (2019) use it to find the optimal exploration sequence of near-earth asteroids. In addition, Arora, Fitch, and Sukkarieh (2017) show that MCTS is suitable to design a robot that is able to perform exploration of a Mars-analog environment autonomously, being able to plan its actions on-line according to the perceptions acquired from the environment.

Various applications of MCTS can be found also in logistic (Edelkamp *et al.*, 2016). For instance, there are multiple studies that address transportation problems. Trunda and Barták (2013) show how MCTS can be used for planning in multiple transportation domains. Moreover, Al-Kanj, Powell, and Bouzaiene-Ayari (2016) present an MCTS-based solution for routing a utility truck that restores outages in the power grid, actively collecting information and updating its beliefs about the state of the network. Abdo, Edelkamp, and Lawo (2016), instead, propose to use MCTS to design an algorithm that is able to deal with different variations of the vehicle routing problem with multiple vehicles, instead of designing a specific solution for each variation. A similar problem is addressed by Jiang, Al-Kanj, and Powell (2017), which design a variant of the MCTS algorithm that optimizes the behavior of a single driver navigating a graph while operating on a ride-sharing platform. Finally, Cazenave, Balbo, and Pinson (2009) show how a variant of MCTS, nested Monte-Carlo search, can be applied to urban transportation in order to minimize the waiting time of bus passengers, and Mirheli and Hajibabai (2019) show how to use MCTS to manage parking utilization in such a way that the travelers' cost is minimized and the parking agency's revenue is maximized.

MCTS has also been successfully applied to energy management. It has been used to solve stochastic energy stock management problems (Couëtoux and Doghmen, 2011), and to balance electricity supply and demand in power grids (Golpayegani, Dusparic, and Clarke, 2015).

Many more examples of the application of MCTS variants to concrete problems can be mentioned. These include designing security models (Marecki, Tesauro, and Segal, 2012; Guo, An, and Kolobov, 2015), military simulation and planning (Marks *et al.*, 2013; Teter *et al.*, 2014), protecting natural resources from illegal extraction (Qian *et al.*, 2014), deciphering encrypted text (Hauer, Hayward, and Kondrak, 2014) and forecasting financial volatility of assets (Cazenave and Hamida, 2015). Further examples are managing autonomous driving vehicles (Lenz, Kessler, and Knoll, 2016), managing wildfires (Bertsimas *et al.*, 2017), improving computer-aided retrosynthesis (Segler *et al.*, 2018) maximizing the performance of job scheduling heuristics (Wimmenauer, 2019), and finding optimal material designs (Dieb *et al.*, 2019). All these examples show the wide applicability of MCTS-based algorithms and many more applications are being explored every day.

Artificial General Intelligence

Research in AGI is based on the idea of having a program that is able to perform multiple different tasks in multiple different environments without any human intervention and with the ability to autonomously adapt to each new, possibly previously unseen task. Although not yet achieving true AGI, programs that are able to cope with tasks of different nature without needing human intervention are suitable to be applied in many real-world scenarios. They are potentially useful for environments in which the type of tasks that the machine has to face might change over time, but they cannot be predicted or even imagined by the programmers. Moreover, such programs might be successful in environments for which it is too time consuming or even physically impossible to re-program the machine for each new type of task. For example, for a robot on a space mission there might be scenarios that the programmers have not thought about, therefore the robot should be able to learn by itself how to cope with them and how to perform the necessary tasks to deal with the unforeseen situation. Moreover, programmers are not able to physically access the robot to re-program it once it has reached space.

The discussion in previous sections highlighted how MCTS is a technique that can be applied successfully to many domains of social and economic interest. Other valorization opportunities created by the research presented in this thesis arise from the possibility of using the proposed domain-independent MCTS enhancements to develop computer programs and systems that are capable of dealing with multiple tasks in various domains at the same time, following the direction of AGI.