POLITECHNIKA WARSZAWSKA

WYDZIAŁ MATEMATYKI I NAUK INFORMACYJNYCH

Rozprawa doktorska

mgr inż. Karol Walędzik

Połączenie algorytmów symulacyjnych oraz dziedzinowych metod heurystycznych w zagadnieniach dynamicznego podejmowania decyzji

Promotor

prof. dr hab. inż. Jacek Mańdziuk

WARSZAWA 2017

Streszczenie

Rozprawa poświęcona jest tematyce budowy hybrydowych algorytmów łączących metody symulacyjne, na przykładzie algorytmu UCT (ang. *Upper Confidence bounds applied to Trees*), oraz dziedzinowe metody heurystyczne. Jej celem jest zbadanie skuteczności zastosowania tego typu podejść w problemach dynamicznego podejmowania decyzji, na przykładzie dwóch problemów:

- budowy agenta grającego w gry zdefiniowane w formule General Game Playing

oraz

- konstrukcji strategii postępowania w problemie harmonogramowania projektów z uwzględnieniem ryzyk (ang. *Risk-Aware Project Scheduling Problem – RAPSP*).

General Game Playing to platforma definiująca język opisu reguł oraz środowisko prowadzenia rozgrywek w skończone, symultaniczne, deterministyczne gry jedno- i wieloosobowe o pełnej informacji. Zadaniem agenta GGP jest autonomiczne opracowanie w ograniczonym czasie wygrywającej strategii dla gry o nieznanych wcześniej regułach.

Risk-Aware Project Scheduling Problem to zaproponowane w rozprawie uogólnienie stochastycznego problemu harmonogramowania projektów przy ograniczonych zasobach. Bazowa formuła projektu uzupełniona jest w tym podejściu o podstawowe elementy zarządzania ryzykiem: elastyczny model ryzyk oraz odpowiedzi na nie.

Algorytm UCT, dominująca w literaturze odmiana przeszukiwania drzew metodą Monte Carlo (ang. *Monte Carlo Tree Search* – MCTS), swoją popularność zawdzięcza przede wszystkim skuteczności w obszarze gier dwuosobowych, w pierwszej kolejności Go. Rozprawa obejmuje weryfikację hipotezy jego użyteczności w znacznie szerszej klasie problemów: zagadnieniach dynamicznego podejmowania decyzji. Są to problemy, w

których agent zmuszony jest podejmować akcje w środowisku, które zmienia się w czasie – zarówno w efekcie podjętych decyzji, jak i czynników niezależnych od zachowania agenta. Naturalnym obszarem zastosowania MCTS są takie instancje problemów, dla których ekstensywne przeszukanie przestrzeni stanów jest w praktyce niemożliwe, a dostarczenie wiarygodnej miary jakości stanów pośrednich – problematyczne.

Choć UCT jest metodą aheurystyczną, jej elastyczność sprawia, że w zagadnieniach, w których jest to możliwe, naturalnym jest uzupełnienie jej o wykorzystanie dziedzinowych metod heurystycznych. W ramach rozprawy proponowane są różne możliwości łączenia tych podejść oraz badane efekty ich synergii.

Rozprawa formułuje oraz pozytywnie weryfikuje, na dwóch wyżej przywołanych przykładach, hipotezę o możliwości budowy tego typu algorytmów hybrydowych, które swoją skutecznością przewyższą wyniki zastosowania każdej z metod składowych samodzielnie. W ramach opisywanych badań, jako ilustracja powyższej tezy, proponowane są nowe metody łączenia podejść symulacyjnych oraz dziedzinowych metod heurystycznych, konstruowany jest skuteczny agent GGP oraz rodzina strategii rozwiązywania zagadnienia RAPSP.

Karol Walędzik

Abstract

This thesis is concerned with the problem of constructing hybrid algorithms combining simulation-based tree search methods, represented by UCT (*Upper Confidence bounds applied to Trees*), and domain-specific heuristic approaches. Its main aim is assessment of applicability of this approach to dynamic decision-making problems, represented by:

- development of General Game Playing agent

and

- solving Risk-Aware Project Scheduling Problem.

General Game Playing is a framework defining testbed for building and testing multi-game playing agents. It offers a game description language capable of defining any finite, discrete, deterministic single- or multi-player game. GGP agents are expected to be able to autonomously analyze the rules and devise successful game playing strategies for games they have never encountered before.

Risk-Aware Project Scheduling is a very flexible generalized version of the Stochastic Resource-Constrained Project Scheduling Problem, proposed by the author of the thesis. It extends the project model with elements of risk-management processes: risks and risk responses.

UCT algorithm, the currently dominant Monte Carlo Tree Search method, is popular mainly thanks to its spectacular successes in the domain of two player game-playing, especially the game of Go. This thesis proves its applicability to a wider class of dynamic decision-making problems. These include problems in which the agent is forced to take a sequence of actions in a dynamic

environment, which may change from one point in time to another both in response to the agent's choices as well as on its own (in a non-deterministic fashion or in response to stimuli outside of the agent's control).

While one of the strengths of the UCT method in many research areas is the fact that it is truly aheuristic, the algorithm is at the same time flexible enough to easily be augmented by inclusion of problem-specific heuristic approaches. The thesis explores several possibilities of constructing such hybrid methods, verifies their efficiency and analyzes the synergy effects present in the resulting combinations.

It proposes and proves true the hypothesis that it is possible to build such hybrid methods that surpass the quality of each of their components on their own. This is done using the two above-mentioned examples: GGP and RAPSP. While doing so, new methods of designing hybrid UCT-based methods are designed and tested, a successful GGP agent is constructed and a family of RAPSP solving strategies is devised, implemented and thoroughly analyzed.

Karol Walędzik