Streszczenie

Aproksymacja stanu równowagi Stackelberga w grach wielokrokowych o sumie niezerowej z niepełną informacją z użyciem metod Monte Carlo

Rozprawa opisuje dwie metody, będące oryginalnym wynikiem, które pozwalają przybliżać strategię lidera ze stanu Równowagi Stackelberga w grach wielokrokowych o sumie niezerowej z niepełną informacją. Hipotezą badawczą stawianą w rozprawie jest twierdzenie, że możliwe jest zbudowanie efektywnej metody poszukiwania Równowagi Stackelberga w tych grach w oparciu o próbkowanie Monte Carlo.

Gry wielokrokowe, to rodzaj gier, w których gracze mają wiele punktów decyzyjnych, w każdym z tych punktów muszą wykonać akcję. W teori gier często do reprezentacji gier wielokrokowych używa się postaci ekstensywnej – drzewa, gdzie węzłami są stany gry, a krawędziami ruchy możliwe do wykonania w tych stanach. Gry z niepełną informacją charakteryzują się tym, że gracze nie mają pełnej informacji o całym stanie gry, w szczególności o działaniach przeciwnika, a mogą obserwować jedynie projekcję stanu udostępniającą wybrane informacje. Dodatkowo, w tej rozprawie rozważane są tylko gry z doskonałą pamięcią, to znaczy takie, gdzie udostępniana graczowi informacja zawsze uwzględnia rozróżnienie stanów na podstawie wszystkiego co gracz dotąd zaobserwował, w tym wykonanych przez niego akcji. Suma niezerowa oznacza, że w grze dla dwóch graczy wypłaty otrzymane na koniec gry przez graczy nie muszą sumować się do zera.

Równowaga Stackelberga jest pojęciem z obszaru teorii gier. W Grze Stackelberga bierze udział dwóch asymetrycznych graczy, lider i naśladowca. Lider wybiera strategię mieszaną jako pierwszy, następnie upublicznia ją. Naśladowca wybiera swoją strategię znając już strategię lidera. Model Stackelberga zakłada pełną racjonalność naśladowcy, czyli zachowanie gdzie naśladowca zawsze wybierze strategię, która daje mu najlepszą możliwą wypłatę. Stanem Równowagi Stackelberga nazywamy układ strategii lidera i naśladowcy, gdzie strategia naśladowcy jest optymalną odpowiedzią na strategię lidera, a strategia lidera daje liderowi najwyższą możliwą wypłatę spośród wszystkich układów strategii spełniających warunek optymalnej odpowiedzi naśladowcy. Rozprawa wskazuje pozycje w literaturze, które mówią o praktycznym zastosowaniu Równowagi Stackelberga w sytuacjach związanych z interakcją pomiędzy siłami bezpieczeństwa (np. policją, strażą graniczną), a łamiącymi prawo (przemytnikami, terrorystami).

Po zdefiniowaniu potrzebnych pojęć z teorii gier zaprezentowane są istniejące w literaturze podejścia do obliczeniowego wyznaczania Równowagi Stackelberga. Duża część z istniejących w literaturze podejść jest dedykowana bardzo szczególnym podklasom wielokrokowych Gier Stackelberga o sumie niezerowej z niepełną informacją. Są to metody, które wykorzystują specyficzne cechy w strukturze gry, aby znacznie przyspieszyć obliczenia. Tych metod nie da się uogólnić na całą klasę gier rozważaną w rozprawie. Zaprezentowane są również metody z literatury dedykowane całej wspomnianej klasie gier. Wszystkie prezentowane metody, zarówno te specyficzne dla danej klasy, jak i ogólne wykorzystują programowanie liniowe jako istotny

na trzech zbiorach gier testowych i porównana z metodami z literatury. Dla wszystkich zbiorów testowych, dla dużych instancji gier testowych O2-UCT jest szybsza od metod z literatury. Wartości wypłat uzyskiwane przez strategie są bardzo bliskie optymalnym dla wszystkich gier, dla których udało się policzyć rozwiązania dokładne.

Wyniki eksperymentalne badające zaproponowane w tej rozprawie metody potwierdzają hipotezę badawczą postawioną w rozprawie.

Słowa kluczowe: Równowaga Stackelberga, UCT, MCTS

Abstract

Monte Carlo methods for approximation of Stackelberg Equilibium in sequential multi-act general sum games with imperfect information

This thesis describes two original methods that approximate leader's strategy of a Stackelberg Equilibrium in multi-act non-zero sum games with imperfect information. The research hypothesis checked in this thesis is that it is possible to construct a method based on Monte Carlo sampling that seeks Stackelberg Equilibrium in such games.

Multi-act games are a type of games where players encounter multiple decision points in the game and they have to act multiple times, once in every point. The popular representation of such games in game theory is extensive game — a tree where each node is a game state and arcs between nodes are actions possible to play in the given state. Imperfect information is a property that means that players in the game do not see the whole state of the game, in particular they may have limited information about opponents' actions. They can see only a projection of the state which reveals some information. In this is only games that satisfy additional property called perfect recall are considered. Perfect recall means that player can always tell the difference between two states where previous observations and actions of that player were different from each other. Non-zero sum game means that in two player game players' payoffs do not necessary sum up to 0.

Stackelberg Equilibrium is a game-theoretic concept. Stackelberg Game is played by two asymmetric players, the leader and the follower. The leader commits to a mixed strategy first and makes it public. The follower chooses a strategy already knowing leader's commitment. The assumption is that the follower is perfectly rational and chooses the strategy maximizing their payoff. The strategy profile is called Stackelberg Equilibrium where the follower's strategy is the best response to the leader's strategy and the profile maximizes leader's utility across all profiles satisfying the best response restriction. The literature review in the thesis cites papers that describe practical applications of Stackelberg Equilibrium in scenarios where interactions between law enforcement (police, border guards) and criminals (smugglers, terrorists).

After defining all required game-theoretic terms the thesis presents existing approaches do finding Stackelberg Equilibrium found in the literature. Majority of existing methods is dedicated to specific subclasses fo Stackelberg non-zero sum games with imperfect information. Those methods exploit specific properties of a game structure in given class to speed up computation. It is not possible to generalize such approaches to general class of games. Methods that are applicable to broad class of all games with the mentioned properties are also presented. All the presented methods, both game-specific and general ones employ linear programming as a base of the solutions. Some of the methods are composed of a linear program that defines the equilibrium and simply solve it, other methods solve many linear programs and use some additional procedures to obtain the final solution. During the analysis of the existing methods some recurring elements were noted, as they might be useful when building new methods for Stackelberg Equilibrium. Those include: browsing all follower's strategies and for each of the