

Coalition skill games with incomplete information and over-abundance of tasks

Mihail Bogojeski, Tobias Pockrandt, Mitja Phillip Richter, Björn Fischer

Abstract

In diesem Paper stellen wir einen Lösungsansatz für das Matchen von Agenten mit gewissen Fertigkeiten auf Aufgaben welche besagte Fertigkeiten benötigen. Das Ziel ist es bei fehlen von vollständiger Information eine Maximierung der sozialen Wohlfahrt zu erreichen. Hierzu werden verschiedene Methoden und Konzepte aus Bereichen wie der Spieltheorie angewandt. Der vorgestellte Lösungsansatz wird ca. 80% der theoretisch möglichen optimalen Lösung erreichen und wir liefern eine Reihe von Maßnahmen die eine Verbesserung dieses Wertes ermöglichen sollten, bzw. die Rechenzeit bei minimalem Verlust verbessern.

Keywords

coalition games — skills — incomplete information — matching problem

Contents

Introduction	1
1 Versuchsumgebung und Beschreibung	1
2 Algorithmusbeschreibung	2
2.1 Nutzenkalkulation	2
2.2 Bewerbungsphase	2
2.3 Koalitionsermittlung	2
2.4 Beste Koalition	2
2.5 Überschusstilgung	2
2.6 Gewinnverteilung	2
2.7 Festlegung auf Quest/Koalition	2
2.8 Schliessung erfüllter Quests	2
3 Methoden und Konzepte	2
4 Evaluation	2
5 Future Work	2
6 Conclusion	3
Acknowledgments	3

Introduction

Dieses Paper stellt einen Abschlussbericht für den Kurs *Agententechnologien in der Forschung* dar und repräsentiert die im Semester erarbeiteten Resultate. Das Matchen von Agenten und Aufgaben ist ein äußerst komplexes Problem. Aus diesem Grund werden einige Einschränkungen vorgenommen. Es existiert ein Pool aus Fertigkeiten und ihren Intensitäten. Aus diesem Pool werden die Aufgaben und die Fertigkeiten der Agenten entnommen. Zwischen der Menge von Agenten

und der, der Aufgaben soll ein möglichst effizientes Matching gefunden werden. Effizienz definiert hierbei der Grad in dem die soziale Wohlfahrt maximiert wurde. Um einen erhöhten Bezug zu Realität zu erreichen besitzen die Agenten nur ein beschränktes Wissen über ihre Umgebung. Dies führt zusätzlich dazu dass die beste Lösung nicht durch einfaches propagieren aller Permutationen ermittelbar ist. Für die beschriebene Spezifikation des Problems wird in diesem Paper eine Lösung dargestellt. Hierfür wird zunächst auf die genauen Rahmenbedingungen des Versuchs eingegangen (KAPITEL 1). Im Anschluss erfolgt eine genauere Betrachtung des angewandten Algorithmus (KAPITEL 2) sowie der genutzten Methoden und Konzepte (KAPITEL 3). Um die Qualität der präsentierten Lösung zu bewerten wurde eine Evaluation durchgeführt (KAPITEL 4). Den Abschluss bilden ein Ausblick über mögliche Verbesserungen (KAPITEL 5), sowie eine Zusammenfassung (KAPITEL 6).

1. Versuchsumgebung und Beschreibung

Um eine spannendere Umgebung zu erschaffen, bezieht sich der gewählte Algorithmus nicht einfach auf Agenten und Aufgaben. Um das gesetzte Problem interessanter und intuitiver darzustellen handelt es sich hier um Abenteurer (Agenten) die in einer Kaschemme angeheuert werden um gewisse Quests (Aufgaben) gegen die Zahlung einer bestimmten Menge Gold zu erfüllen. Im folgenden wird daher auch von Abenteurern und Quests gesprochen. Die Abenteurer besitzen einen von 3 möglichen Skills (Fertigkeiten). Der Pool aus Skills besteht aus Kämpfen, Verhandeln und Schleichen. Jeder Fertigkeit eines Abenteurers ist ein Power-Wert zugeordnet, die beschreibt wie gut er diese beherrscht. Die Quests benötigen von jedem Skill einen gewissen Wert um erfüllt zu werden. Eine Quest kann nicht abgeschlossen werden, solange nicht alle Anforderungen komplett erfüllt worden sind.

Aufgrund des Überangebots von Aufgaben existieren mehr Quests als Abenteurer, bzw. benötigen die Quests mehr Power um gelöst zu werden als die Agenten in ihrer Summe verfügen. Sowohl die Abenteurer als auch die Quests werden in einem gewissen Rahmen zufällig generiert. Für erstere wird ca. die Hälfte der Gesamtpower aller Abenteurer auf die Abenteurer verteilt. Die Quests hingegen sind frei von derartigen Abhängigkeiten. Sie unterscheiden sich allerdings in ihre Größe (klein, mittel, groß), wobei viele kleine und wenig große Quests existieren. Der Erlös einer Quest im Vergleich zu ihrer Größe ist allerdings nicht linear, sodass das ein großes Quest proportional mehr Gold ergibt als die Summe kleinerer mit gleicher Power. Dies entspricht auch der Realitätsnähe des Versuchs, da Großprojekte i.d.R. einen größeren Gewinn ausstoßen als kleinere.

Wie der Titel der Arbeit verrät wird hier mit unvollständigen Informationen gearbeitet. In diesem Fall bedeutet dies speziell, dass zwar jeder Abenteurer alle Quests kennt, inklusive deren Bedarf und Erlös, allerdings besitzt der Agent keine konkreten Informationen darüber, wie viele andere Agenten im Spiel sind und insbesondere nicht wie viel Power sie besitzen.

In der Planungsphase des Versuchsaufbaus gab es die Diskussion ob die Abenteurer mit Solo- oder Multiskilled sein sollten. Für den Besitz von 2-3 Skills sprachen eine interessantere Verhandlungsphase zwischen den Abenteurern. Dagegen standen der stark erhöhte Aufwand, sowie der Realismus, da eine Firma zu Meist auf eine Aufgabe spezialisiert ist. Weiterhin bestand die unbewiesene Annahme, dass sich ein Multiskillagent auch als mehrere Soloskillagenten darstellen lässt. So führte auch nicht zuletzt die Tatsache, dass hier lediglich ein relativ ausgereifter Lösungsansatz vorgestellt werden soll, zu der Entscheidung Soloskillagenten zu verwenden.

Weiterhin besitzen die Agenten für jedes Abenteuer gewisse Grundkosten die anfallen und mit dem Gewinn abgedeckt werden müssen. In unserem Beispiel ist das die Beschaffung von Tränken und die Instandhaltung von Waffen und Ausrüstung. In der Realität könnte sich dies durch Fahrtkosten oder Wartung von Maschinen darstellen lassen. Ein weiteres Ziel welches damit verfolgt wird, ist eine Erhöhung der Individualität der Einzelnen Agenten. Die Abenteurer unterliegen alle dem gleichen Verhalten, sodass sich ihre Handlungen nur durch ihre eigene Power und die Grundkosten unterscheidet. Ergo gibt es hier keine risikoscheuen oder risikoaverse Agenten. Für alle Agenten existiert ein Zeitlimit (Deadline) welches bekannt ist, nachdem keine Bewerbungen auf Quests mehr möglich sind, sodass die Abenteurer bei fortgeschrittener Zeit auch in weniger profitablere Alternativen wechseln um noch Gewinne zu machen.

Die Stellschrauben an denen hier initial gedreht werden kann, sind die Anzahl an Abenteurern und die der Quests. Die Einhaltung der Rahmenbedingungen erfolgen anschließend automatisch. Eine genaue Beschreibung des Ablaufes des Algorithmus erfolgt im nächsten Kapitel.

2. Algorithmusbeschreibung

Zunächst erfolgt eine grobe Zusammenfassung des Algorithmus, wobei die einzelnen Bestandteile in den folgenden Unterkapiteln ausführlicher beleuchtet werden. Zu Beginn berechnen die Agenten für jedes Abenteuer ihren Nutzen. Anschließend dürfen sie sich auf bis zu 4. Abenteurer bewerben. Anhand dieser Bewerbungen werden Koalitionen gebildet und die Beste ausgewählt. Potentieller Überschuss in den Koalitionen wird entfernt und möglichst sinnvoll an die Agenten zurückgezahlt, da nur so ein Erreichen des maximal möglichen Erlöses erst möglich wird. Der zu erwartende Gewinn wird virtuell an die Agenten verteilt, welche sich mit diesen Informationen auf Quests bzw. Koalitionen festlegen. Letztlich werden die geschlossenen Quests aus dem Spiel entfernt, zusammen mit der eingesetzten Power, die Auszahlungen finden statt und sofern noch Power im Spiel ist und die Zeit noch nicht abgelaufen ist beginnt der Algorithmus von neuem.

2.1 Nutzenkalkulation

2.2 Bewerbungsphase

2.3 Koalitionsermittlung

2.4 Beste Koalition

2.5 Überschusstilgung

2.6 Gewinnverteilung

2.7 Festlegung auf Quest/Koalition

2.8 Schliessung erfüllter Quests

3. Methoden und Konzepte

Hier erfolgt eine Zusammenfassung der Methoden und Konzepte, welche unter anderem aus der Spieltheorie kommen um einen Überblick zu erhalten für potentielle Veränderungen aus KAPITEL 5 oder Diskussionen für und gegen die angewandten bzw. nicht angewandten Mittel.

Banzhaf-Power : Um die Banzhaf-Power eines Agenten zu bestimmen wird gezählt in wie vielen möglichen Koalitionen dieser ein kritischer Spieler ist.

----- :

----- :

----- :

4. Evaluation

5. Future Work

Wie bereits aus der Evaluation des vorherigen Kapitels, bekannt erreicht der von uns entwickelte Algorithmus eine recht akzeptable Leistung, wobei durchaus noch Verbesserungen bzw. Veränderungen möglich und notwendig sind. Diese weiteren Arbeiten können und sollen unter anderem die Effizienz (nähe zum Optimum) als auch die Geschwindigkeit verbessern. Es sind unter anderem auf Maßnahmen aufgelistete, welche den Informationsgehalt des Versuchs erhöhen, indem z.B. mehr

(spezifischere) Daten betrachtet und damit auch berücksichtigt und ausgelesen werden können.

Verteilung des Algorithmus: Die hier gewählte Implementation ist kein echter verteilter Algorithmus. Die Hauptarbeit wird von einer zentralen Instanz durchgeführt, welche die Berechnungen und Entscheidungen für die Agenten (in deren Sinne) durchführt. Wenn jeder Agent als Einzelne Entität in einem verteilten Algorithmus realisiert ist, lassen sich sehr einfach individuelle Verhaltensmuster für jeweilige Abenteuerer festlegen. Dann können unter anderem verschiedene Risikoaspekte mit in den Aufbau eingewoben werden. Was bedeutet risikoscheue und aversive Agenten können aufeinander Treffen.

Ein weiterer Aspekt könnte sein, dass ein Agent A unter keinen Umständen mit einem Agenten B eine Koalition bilden möchte, aufgrund schlechter Erfahrungen oder ähnliches. Hierbei könnte untersucht werden in wie weit die Effizienz des gesamten Konstruktes, durch den Ausschluss der Zusammenarbeit zweier Agenten, in Mitleidenschaft gezogen wird. Der Kreativität im Design vom Verhalten einzelner Agenten ist hierbei kein Limit gesetzt.

Optimierung der Koalitionsbildung: Da aktuell jeder Koalition berechnet werden muss, aufgrund des Banzhafpowerindex befindet sich der Aufwand in einem Exponentiellen Bereich, was dazu führt, dass der Algorithmus schon bei einer vergleichsweise geringen Anzahl an Abenteurern extrem viel Zeit benötigt. Es muss demnach eine Möglichkeit gefunden werden ein im besten Fall identisches Ergebnis in geringerer Zeit zu finden. Der intuitive Ansatz ist hierbei nicht alle Koalitionen zu betrachten, sondern eine simple Vorauswahl zu treffen welche die Anzahl einschränkt. Wie diese Auszusehen hat und wie groß der Effizienzverlust ist muss ermittelt und evaluiert werden.

Maschinelles Lernen: In der von uns entwickelten Lösung ist Maschinelles Lernen nur sehr marginal vertreten. Agenten merken sich unter anderem wie oft sie bereits an einer Koalition gescheitert sind. Wenn dieser Anteil erheblich ausgebaut werden würde und sich jede Aktion und ihr Resultat quasi gespeichert wird, lassen sich unter Umständen ganz neue Lösungsverhalten finden, welche das Optimum näher erreichen. Hierbei sind allerdings nicht Lernphase über die Dauer eines Spiels gemeint, sondern lernen anhand von 1000 und oder mehr Spielen das anzustrebende Ergebnis. Sodass sich mittels Lernen unter Umständen neue oder unerwartete Lösungen mit ähnlichem bzw. besserem Ergebnis finden lassen. Dieser Lernende Agent kann gegen "gewöhnliche" Agenten oder sogar Menschen Spielen, diese schnell Anhand ihres Verhaltens identifizieren bzw. kategorisieren und eine geeignete Strategien anzuwenden. Dies sei nur als Anreiz für die Einbettung maschinellen Lernens zu verstehen und erhebt nicht den Anspruch vollständig zu sein.

Untersuchung aller Parameter: Das großflächige Untersuchen sämtlicher Parameter könnte aufschlussreiche Ergebnisse

liefern. In dieser Evaluation wurde nur an der Menge von Agenten und Abenteuern manipuliert. Weitere Stellschrauben an denen zu drehen ist, sind:

- *Belohnung:* Die Verteilung von Auszahlungen könnte variiert werden. Den Anstieg noch mehr Quadratische, oder linear oder mittels einer Beliebigen Fkt darstellen. Auch ein stagnierender Anstieg der Auszahlung für größere Abenteuer ist denkbar.
- *Skillmengen:* Statt 3 Skills, könnte eine Beliebige Anzahl an Skills vorhanden sein (z.B. 5). Eine weitere denkbare Alternative ist, dass nicht jede Baustelle alle Skills braucht, oder die Agenten mehr als nur einen Skill besitzen, wodurch die Verhandlung eine weitere Dimension erreicht.
- *Powergrößen:* Die Power die jeder Agent hat bzw. alle Agenten haben oder die Abenteuer haben, kann weiterhin variiert werden.
- *Angebot:* Zum einen kann das Überangebot der Aufgaben noch verschärft oder zu ein Unterangebot abgewandelt werden. Dann würde sich auch die Frage nach dem Ergebnis abwandeln, da bei dieser Veränderung nicht von Interesse ist, wie viel Gold alle Agenten erwirtschaften, sondern wer bekommt grundsätzlich Gewinn und wie viel.
Ein weiteres Szenario ist eine komplette Gleichverteilung. Es ist demnach möglich alle Quest zu lösen ohne dass ein Skillpunkt eines Abenteurers übrig bleibt. Hier stellt sich die Frage, welcher Agent wie viel vom Gewinn erwirtschaften kann.
- *Agenten:* Die verschiedenen Verhalten die ein Agent aufweisen kann wurden bereits erläutert.

6. Conclusion

Acknowledgments

So long and thanks for all the fish [?].