RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

Spieltheorie als Entscheidungswerkzeug für patrouillierende Roboter

Seminarvortrag
Tobias Schmidt

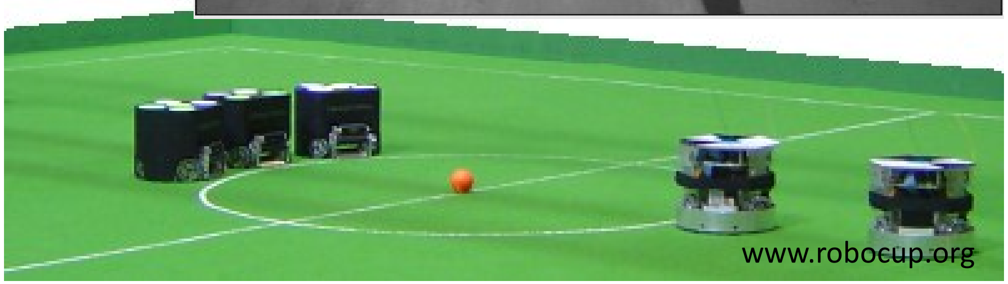
Betreuer: Dipl.-Inf. Alexopoulos
Lehrstuhl für Automation

Motivation


www.netzwelt.de


www.robowatch.de


www.iosb.fraunhofer.de


www.robocup.org

Problemstellung

- Gegeben:
  - Umgebungsbeschreibung
  - Aufgabenbeschreibung
  - Bewegungseigenschaften
  - Informationsstruktur
- Gesucht für patrouillierende Roboter:
  - Umfassendes Modell mit Berücksichtigung der „Gegner“
  - Effiziente Patrouillierungsstrategie
- Lösungsansatz:
  - Spieltheoretische Problembeschreibung
  - Spieltheorie als Entscheidungswerkzeug für optimale Strategie

Stand der Technik

- Patrouillierende Roboter
  1. Suche von Eindringlingen
  2. Bewachung von Objekten
- Spieltheorie als Werkzeug zur (optimalen) Wegplanung
  - Optimierungsprobleme
  - Spiel in Normalform
  - Spiel in extensiver Form

Ziele

- Spieltheoretische Grundlagen vermitteln
- Modell von Amigoni et al. formalisieren
- Erkenntnisse und Ergebnisse zusammenfassen

Spieltheorie

Was ist Spieltheorie?

- Analyse strategischer Entscheidungssituationen
- Ergebnis hängt von Entscheidungen anderer ab
- Jeder kennt diese Abhängigkeit
- Jeder berücksichtigt diese Abhängigkeit

Spiele in Normalform

- Konfliktsituationen mit wenig formalen Konzepten
- Konzept der Strategie und Auszahlung entscheidend
- Komponenten eines Normalformspiels:
  - Anzahl Spieler
  - Spielregeln
  - Strategien jedes Spielers
  - Auszahlungsfunktion für jede Strategiekombination

Spiele in Normalform

Beispiel *Gefangenendilemma*

- Zwei Tatverdächtige
- Getrennte Vernehmung
- Nachweis geringerer Vergehen
- Beide können wählen:
  - „Ich sage nichts!“
  - „Ich gestehe und belaste den anderen ebenfalls!“
- Je nach Wahlkombination unterschiedliche Haftstrafen

Spiele in Normalform

Beispiel *Gefangenendilemma*

- Anzahl der Spieler: zwei
- Spielregeln: Keine Absprache möglich
- Strategien: Beide können „schweigen“ oder „gestehen“
- Informationsstruktur:
  - Beide Spieler kennen beide Strategiemengen
  - Beide Spieler kennen alle Auszahlungen
  - Beide Spieler wissen, dass obiges bekannt ist
  - Kein Spieler kennt gespielte Strategie des anderen
- Auszahlungen:
  - Beide schweigen: 2
  - Einer gesteht und einer schweigt: 3 & 0
  - Beide gestehen: 1

Spiele in Normalform

Beispiel *Gefangenendilemma*

- Anzahl der Spieler: zwei
- Spielregeln: Keine Absprache möglich
- Strategien: Beide können „schweigen“ oder „gestehen“
- Informationsstruktur:
  - Beide Spieler kennen beide Strategiemengen
  - Beide Spieler kennen alle Auszahlungen
  - Beide Spieler wissen, dass obiges bekannt ist
  - Kein Spieler kennt gespielte Strategie des anderen
- Auszahlungen:

		Spieler 2	
		„schweigen“	„gestehen“
Spieler 1	„schweigen“	2, 2	0, 3
	„gestehen“	3, 0	1, 1

Nullsummenspiele

- Summe aller Auszahlungen einer Konfiguration gleich Null
- Einsatz bei Verfolgerspielen:
  - Ziel Verfolgter: Abstand maximieren
  - Ziel Verfolger: Abstand minimieren
- Beispielmatrix:

		Spieler 2	
		σ_{21}	σ_{22}
Spieler 1	σ_{11}	2, -2	0, 0
	σ_{12}	-3, 3	1, -1

Gemischte Strategien

- Wahrscheinlichkeitsverteilungen über Strategiemengen
- Nutzen berücksichtigt diese Wahrscheinlichkeiten
  - $u_i = \sum H(\sigma_i, \sigma_k) \cdot qk$
- Beispiel:
  - $q_{\sigma 11} = 0{,}3$ und $q_{\sigma 12} = 1 - q_{\sigma 11} = 0{,}7$
  - $q_{\sigma 21} = 0{,}6$ und $q_{\sigma 22} = 1 - q_{\sigma 11} = 0{,}4$

		Spieler 2		
		$\sigma_{21}(0{,}6)$	$\sigma_{22}(0{,}4)$	
Spieler 1	$\sigma_{11}(0{,}3)$	2, -2	0, 0	→ 1,2
	$\sigma_{12}(0{,}7)$	-3, 3	1, -1	→ - 1,4
		↓ 1,5	↓ - 0,7	

Lösungsansatz *Nash-Gleichgewicht*

- Entscheidungsfrage:
  - „Welche Strategie bringt mir beste Auszahlung?“

  → Beste Antwort
- Strategiekonfigurationen mit bester Antwort für jeden
- Kein Anreiz zur Abweichung aller Spieler

  → Gleichgewicht

Lösungsansatz *Nash-Gleichgewicht*

Beispiel *Gefangenendilemma*

- Spieler 1:
  - „gestehen“ immer besser als „schweigen“!
- Spieler 2:
  - „gestehen“ immer besser als „schweigen“!
- Nash-Gleichgewicht:
  - „gestehen“ & „gestehen“

		Spieler 2	
		„schweigen“	„gestehen“
Spieler 1	„schweigen“	2, 2	0, 3
	„gestehen“	3, 0	**1, 1**

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

Lösungsansatz *Maximin-Lösung*

- Entscheidungsfragen:
  - „Welche Auszahlung erhalte ich worst-case?“
  - „Bei welcher Strategie ist diese worst-case-Auszahlung am größten?“
- Beispiel:

		Spieler 2	
		σ_{21}	σ_{22}
Spieler 1	σ_{11}	2, -2	**0, 0**
	σ_{12}	-3, 3	1, -1

→ Konfiguration $(\sigma_{11}, \sigma_{22})$ ist die Maximin-Lösung

Spiele in extensiver Form

Spiele in extensiver Form

- Beschreibung erweitert Normalformspiel
- Fügt explizit Zugfolge und Historieninformation hinzu
- Darstellung als Spielbaum
  - Zyklusfrei
  - Zusammenhängend
  - Eindeutiger Wurzelknoten
  - Knoten: Entscheidungssituationen (Informationsbezirke)
  - Kanten: Aktionsmöglichkeiten (Strategien)
  - Blätter: Auszahlungen

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

Informationsstruktur

- Definiert Informationsmenge der Spieler

Perfekte Information
- Aktionshistorie aller Mitspieler ist jedem bekannt
- Aktueller Aktionsknoten für Spieler eindeutig

Imperfekte Information
- Ein Mitspieler kann andere Aktionen nicht beobachten
- Aktueller Aktionsknoten für Spieler nicht eindeutig

Vollständige Information
- Jeder kennt alle Eigenschaften der Mitspieler

Unvollständige Information
- Mindestens ein Spieler hat private Eigenschaften

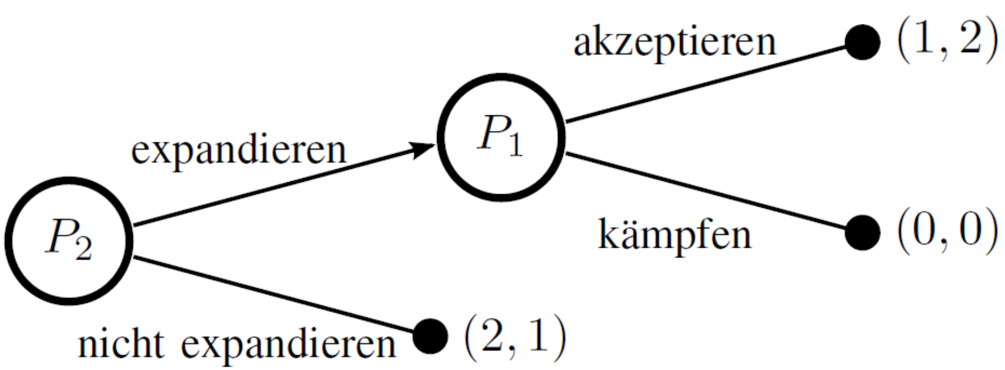
Beispiel *Markteintrittsspiel*

- Ein Unternehmen ist Monopolist
- Zweites Unternehmen möchte in Markt eintreten
- Unternehmen 2 hat Wahl:
  - „In Markt expandieren“
  - „Nicht in Markt expandieren“
- Bei Eintritt hat Unternehmen 1 die Wahl:
  - „Preiskampf betreiben“
  - „Konkurrent akzeptieren“
- Je nach Wahl unterschiedliche Gewinne

Spiele in extensiver Form

Beispiel *Markteintrittsspiel*

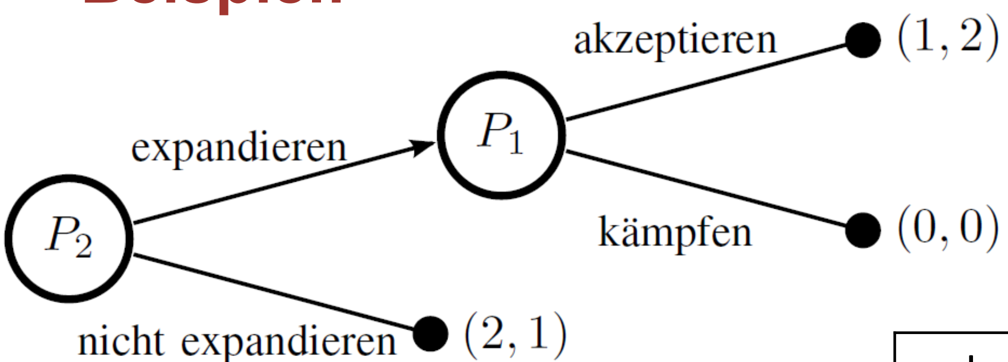
- Anzahl der Spieler:
  - Zwei (P_1 und P_2)
- Zugfolge:
  - Abwechselnd; P_2 zieht zuerst
- Spielregeln:
  - Keine Absprachen möglich
  - Jeder zieht höchstens ein Mal
- Informationsstruktur:
  - Beide Aktionsmengen sind beiden bekannt
  - Alle Auszahlungen sind allen bekannt
  - P_1 kann die Aktion von P_2 beobachten; P_2 weiß das
- Aktionen/Strategien:
  - P_1: „akzeptieren“ oder „kämpfen“
  - P_2: „expandieren“ oder „nicht expandieren“
- Auszahlungen und Entscheidungspfade:



Lösungsansatz *Transformation zu Normalformspielen*

- Induzierte Normalform
- Reine Strategien setzen Strategiemenge
- Auszahlungen = Blätter der Strategiekombination in Graph-Pfaden

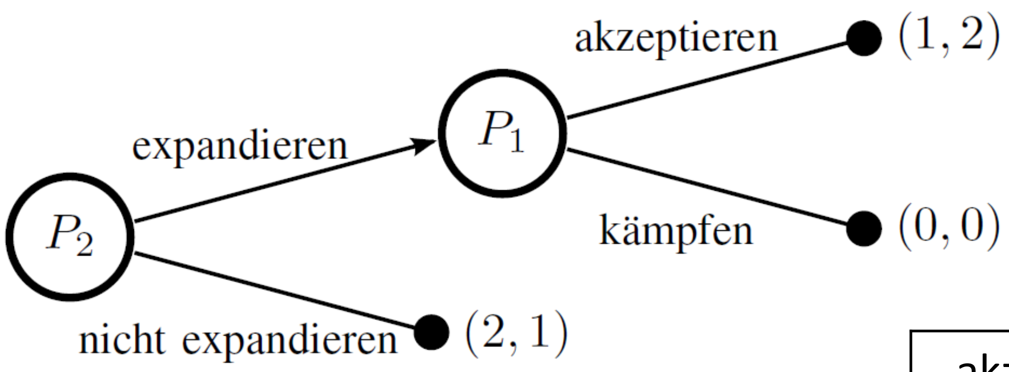
Beispiel:



		P_2	
		„expandieren“	„nicht expandieren“
P_1	„akzeptieren“	1, 2	2, 1
	„kämpfen“	0, 0	2, 1

Lösungsansatz *Transformation zu Normalformspielen*

Beispiel:



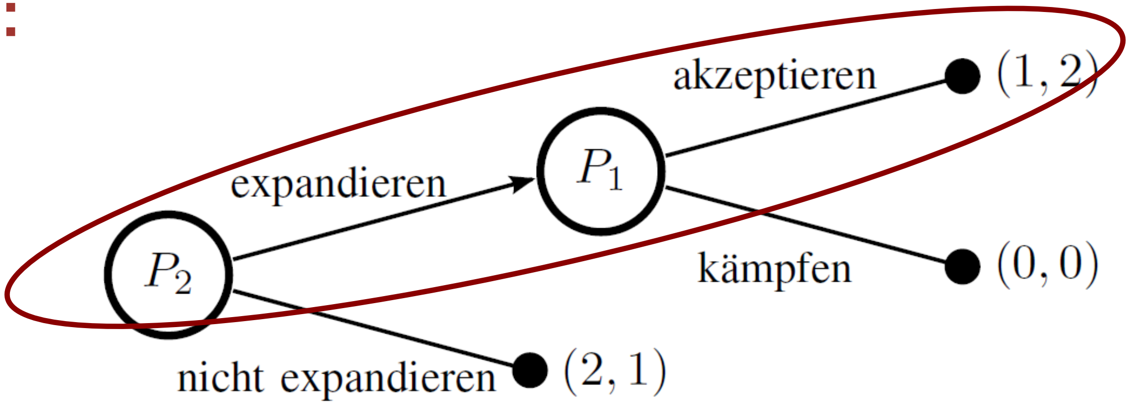
		P_2: „expandieren“	P_2: „nicht expandieren“
P_1	„akzeptieren“	**1, 2**	2, 1
P_1	„kämpfen“	0, 0	**2, 1**

- Unglaubwürdige Drohungen
- Nash-Gleichgewicht beruht auf undefinierter Kombination

Lösungsansatz *Rückwärtsinduktion*

- Rekursive Lösung
- Knoten mit Gleichgewichten in Teilbäumen ersetzen

Beispiel:



→Lösungspfad: P_2: expandieren → P_1: akzeptieren (1, 2)

Spieltheorie bei patrouillierenden Robotern

Spieltheorie bei patrouillierenden Robotern

- Amigoni et al.: „*A Game-Theoretic Approach to Determining Efficient Patrolling Strategies for Mobile Robots*" (2008)
- Annahmen:
  - Zeitdiskretisierung durch Runden
  - n zu patrouillierende Orte
  - Direkte Verbindung zwischen allen Orten
  - Eindringling nur an einem Ort gleichzeitig erkennbar
  - Unterschiedliche Eindringlingsarten
  - Umgebungsänderung während Spiel möglich

Spieldefinition nach Amigoni et al.

- Spieleranzahl:
  - Zwei (patrouillierender Roboter *r*, eindringender Roboter *b*)
- Zugfolge:
  - Gleichzeitige Wahl der Aktion zu Rundenbeginn
- Aktionen/Strategien:
  - *r* wählt Ort zum Patrouillieren
  - *b* wählt, ob er ...
    - *r* beobachtet (und nicht eindringt)
    - zu Ort eindringt (und *r* nicht beobachtet)
- Informationsstruktur:
  - *r* kennt alle Eigenschaften aller Typen
  - *r* kennt Historie von *b* nicht
  - *r* bemerkt Umgebungsänderungen
  - *b* kennt ggf. Historie von *r*
- Spielregeln:
  - Keine Kommunikation
  - *b* braucht Anzahl Runden zum Eindringen
  - *b* darf ein Mal „zu Ort eindringen"
  - *r* kann *b* nur beobachten, wenn
    - beide am selben Ort
    - *r* einen von *b* betretenen Ort patrouilliert
  - Spielende, wenn
    - *b* gefasst wird
    - *b* den Ort erreicht hat
- Auszahlungen:
  - Unterschiedliche Ort-Werte für *r* und *b*
  - Bei Erfassung zusätzliche Auszahlungen für *r* und *b*
  - *r* hat Bewegungskosten

Ansatz von Amigoni et al.

- Extensivform-Spiel in Slices eingeteilt
- Slices entsprechen einer Runde
- Slice ist Normalformspiel mit zwei Spielern
- Alle Slices haben selbe Struktur
  - Nur Auszahlungen unterschiedlich

Ansatz von Amigoni et al.

- Optimierungsproblem
- Unendliche Baumtiefe

→Approximative Lösungssuche mit mathematischen Formulierungen

1. Multi-linear Programming (multi-LP)
2. Mixed Integer Linear Programming (MILP)

- Algorithmen liefern Wahrscheinlichkeitsmenge für Patrouillierungsentscheidungen

Ergebnisse von Amigoni et al.

- Evaluierung des Ansatzes
  1. Laufzeit der Berechnung
  2. Performanz der Patrouillierungsstrategie
     a. Abdeckung der Orte
     b. Unvorhersehbarkeit
     c. Nutzen in einem Slice bei unterschiedlicher Umgebung
  3. Umgang mit Umgebungsänderung

Laufzeit der Berechnung

- multi-LP skaliert besser bei mehr Orten
- MILP skaliert besser bei mehr Eindringlingstypen

multi-LP					
n (# Orte)	1 Typ	2 Typen	3 Typen	4 Typen	5 Typen
3	0,068	0,342	1,278	6,854	62,972
5	0,079	0,796	13,227	544,481	-
10	0,280	16,095	-	-	-
15	0,509	146,665	-	-	-
25	1,085				
50	3,702				
100	16,096				

MILP					
n (# Orte)	1 Typ	2 Typen	3 Typen	4 Typen	5 Typen
3	0,006	0,015	0,031	0,068	0,138
5	0,014	0,033	0,103	0,357	1,128
10	0,029	0,223	1,615	4,320	-
15	0,074	1,351	8,128	-	-
20	0,280	4,706	14,834	-	-
25	0,752	12,955	-	-	-
50	11,220				
100	215,209				

Performanz der Patrouillierungsstrategie

- Vergleich multi-LP mit:
  a) p = 1/n
  b) p ~ v^r – C
  c) p = rand()
- Nutzen höher, andere Kriterien akzeptabel

Lineare Umgebung			
Methode	# nie besuchter Orte	Entropie	Nutzen
multi-LP	15,4	2,081	78,95
(a)	18,7	3,219	52,124
(b)	12,2	3,014	69,528
(c)	0,5	3,027	48,028

Ringförmige Umgebung			
Methode	# nie besuchter Orte	Entropie	Nutzen
multi-LP	15,4	2,117	82,911
(a)	18,1	3,219	56,149
(b)	10,6	3,019	66,450
(c)	1,3	3,033	50,494

Sternförmige Umgebung			
Methode	# nie besuchter Orte	Entropie	Nutzen
multi-LP	14,9	2,169	75,014
(a)	19	3,219	34,831
(b)	16,4	3,007	55,714
(c)	1,2	3,031	40,444

Umgang mit Umgebungsänderung

- Lineare Umgebung mit 3 Orten
- Nach 100 Runden Kosten erhöht

Ort i	v_i^r	# Besuche vor Änderung	# Besuche nach Änderung
1	90	47,9	171,2
2	9	8,5	0
3	30	43,6	228,8

Zusammenfassung

- Spieltheoretische Grundlagen
  - Normal- und Extensivformspiele
  - Gleichgewichte
    - Nash-Gleichgewicht
    - Maximin-Lösung
    - Rückwärtsinduktion
  - Informationsstrukturen
- Ansatz zur Bestimmung einer Patrouillierungsstrategie
  - Spiel in extensiver Form in Slices einteilen
  - Slice als Spiel in Normalform betrachten
  - Wiederholte Optimierung einzelner Slices löst approximativ

Fazit

- Interessanter Lösungsansatz
- Effizienz des Ansatzes berücksichtigt nicht Berechnungsdauer
- Vergleich der Laufzeiten der Methoden wäre interessant
- Test des Modells auf realen Robotern

Vielen Dank

Referenzen von Vortrag und Paper

(1) F. Amigoni, N. Gatti, und A. Ippedico, „A Game-Theoretic Approach to Determining Efficient Patrolling Strategies for Mobile Robots," in Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 02, Serie WI-IAT '08. Washington, DC, USA: IEEE Computer Society, 2008, S. 500–503. [Online]. http://dx.doi.org/10.1109/WIIAT.2008.324

(2) S. K. Berninghaus, K.-M. Ehrhart, W. Güth, S. K. Berninghaus, K.-M. Ehrhart, und W. Güth, Strategische Spiele, Serie Springer-Lehrbuch. Springer Berlin Heidelberg, 2010. [Online]. http://dx.doi.org/10.1007/978-3-642-11651-3

(3) M. J. Osborne, Introduction to Game Theory: International Edition. Oxford University Press, 2009.

(4) D. Fudenberg und J. Tirole, Game Theory. Cambridge, MA: MIT Press, 1991.

(5) A. Wellenreuther, „Spieltheorie," Vorlesungsskript, 2008.

(6) V. Conitzer und T. Sandholm, „Computing the optimal strategy to commit to," in Proceedings of the 7th ACM conference on Electronic commerce, Serie EC '06. New York, NY, USA: ACM, 2006, S. 82–90. [Online]. http://doi.acm.org/10.1145/1134707.1134717

(7) P. Paruchuri, J. P. Pearce, M. Tambe, F. Ordonez, und S. Kraus, „An efficient heuristic approach for security against multiple adversaries," in Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems, Serie AAMAS '07. New York, NY, USA: ACM, 2007, S. 181:1–181:8. [Online]. http://doi.acm.org/10.1145/1329125.1329344

(8) P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordonez, und S. Kraus, „Playing games for security: an efficient exact algorithm for solving Bayesian Stackelberg games," in Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2, Serie AAMAS '08. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2008, S. 895–902. [Online]. http://dl.acm.org/citation.cfm?id=1402298.1402348

(9) R. Bellman, Dynamic Programming. Princeton, New Jersey: Princeton University Press, 1957.