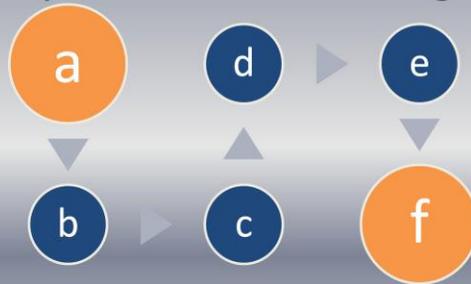


Künstliche Intelligenz – Intelligente Agenten & komplexe Entscheidungen

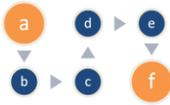


Jochen Bernhardt
Fabian Gippert

Vortrag vom 19.12.2007 im Rahmender Vorlesung:

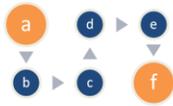
Künstliche Intelligenz / Künstliche Agenten im WS 07/08 -FH Köln Campus
Gummersbach

Jochen Bernhardt 11042430
Fabian Gippert 11043446



Inhalt

- Sequenzielle Entscheidungen
- MEP's => Markov Entscheidungsprozesse
- Taktiken und Horizonte
- Spieltheorie
- Dynamische Entscheidungsnetzwerke
- Entwurfsmechanismen



Sequenzielle Entscheidungen

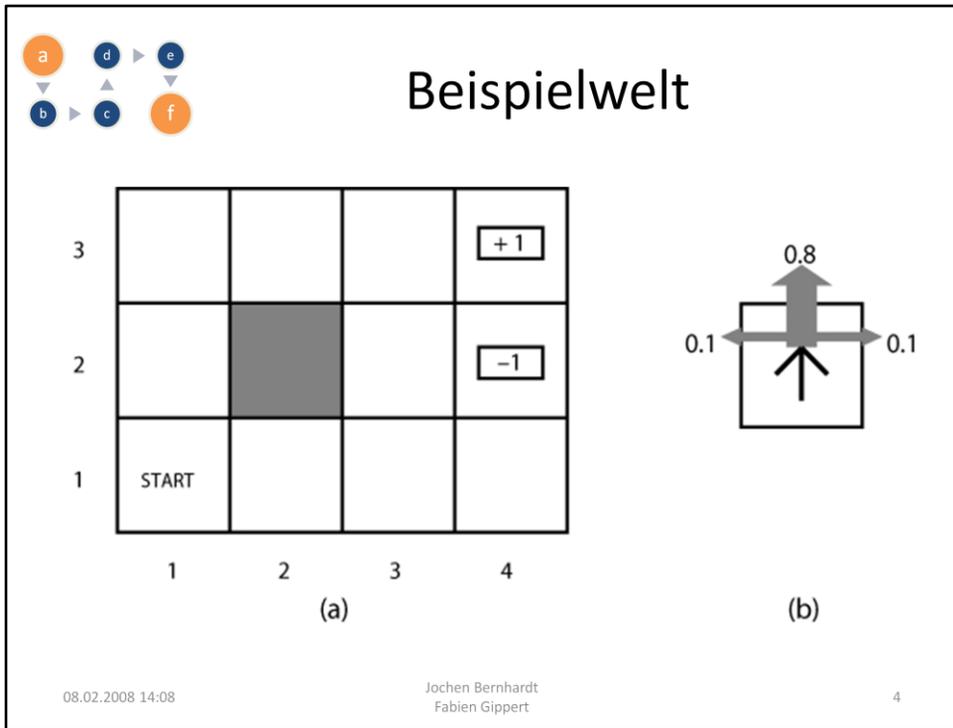
- Wie bewerten wir Folgen von Entscheidungen?
- Wie ermitteln wir den besten „Weg“ zum Ziel?

08.02.2008 14:08

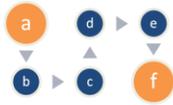
Jochen Bernhardt
Fabien Gippert

3

Was ist das? Vorstellen worum es dabei geht.

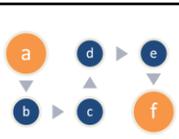


Beispiel Welt Abb 17.1 erklären und vorstellen
Welt, Übergangsmodelle, Bewegungswahrscheinlichkeit



Markov Entscheidungsprozesse

- „*Sequentielle Entscheidungsprobleme in unsicheren Umgebungen*“
- Ausgangszustand: S_0
- Übergangsmodell: $T(s, a, s')$
- Gewinnfunktion: $R(s)$
- Markov Übergang:
Wahrscheinlichkeit von s zu s' zu gelangen
ist nicht von Vorgängeraktionen abhängig.



„Langfristiger“ Nutzen

- Welchen Ansporn hat der Agent das Ziel zu erreichen?
- Negativer Nutzen Pro Feld (-0,04)
- Nutzen der Terminalsymbole (+1 / -1)

=> Der Agent fühlt sich „ein wenig unwohl“ und versucht das sichere Ziel zu erreichen.

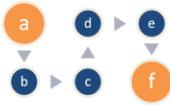
08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

6

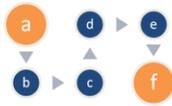
Additive Nutzen Funktion, welcher Weg bringt welchen Nutzen, welchen nutzen bringen Terminalzustände.

Was macht der Agent abhängig seines „wohlbefindens“?

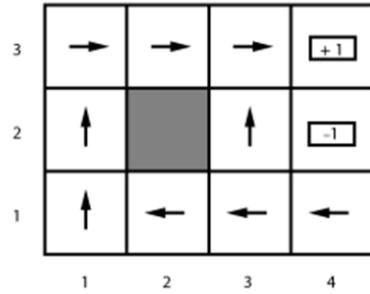


Nutzen und Glück

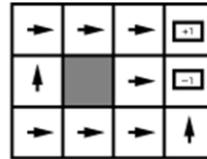
- Wann ist der Agent wie (un-)glücklich?
- Was macht der Agent wenn er
 - Unglücklich ?
 - ein wenig unglücklich ?
 - tottraurig oder ?
 - glücklich ist ?



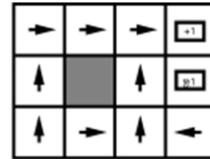
Verschiedene Taktiken



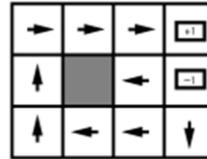
(a)



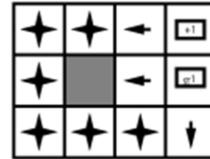
$$R(s) < \approx 1.6284$$



$$\approx 0.4278 < R(s) < \approx 0.0850$$

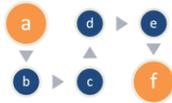


$$\approx 0.0221 < R(s) < 0$$



$$R(s) > 0$$

(b)



Taktik und Horizonte

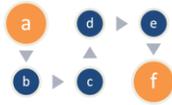
- Eine Taktik ist der Plan das Ziel zu erreichen!
- Taktik wird als π , die von ihr empfohlene Aktion als $\pi(s)$ in der Situation s bezeichnet.
- Endlicher Horizont: Die Aufgabe muss in einer festen Zeit beendet sein.
- Unendlicher Horizont: keine zeitliche Terminierung

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

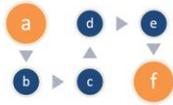
9

Mündlich eingehen auf den die Beeinflussung von Taktiken durch Hotizonte!



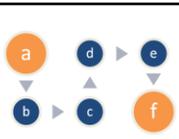
Problem unendlicher Horizonte

- Unendliche Horizonte
- Auswirkungen auf Situationen „Far Far Away“
- Unendliche Folgen => Unendlicher Nutzen
 - Lösung 1 : Verminderte Gewinne
 - => $\sum_{t=0}^{\infty} \gamma^t R(s_t)$ statt $\sum_{t=0}^{\infty} R(s_t)$
 - Führt immer zu einem endlichen Nutzen!
 - Lösung 2 : Garantieren das Taktik Terminalzustände erreicht („Richtige Taktik“)
 - Lösung 3 : Algorithmen zum Vergleich des Wachstums unendlicher Folgen



Partielle beobachtbare MEPs

- Wesentlich unklarere Umgebung
- Zustand des Agenten unklar
 - $\pi(s)$ nicht ausführbar (s unbekannt)
- Nutzen nun abhängig von Wissen und Glauben des Agenten



Partielle MEPs

- Partielles MEP:
 - Ausgangszustand: S_0
 - Übergangsmodell: $T(s, a, s')$
 - Beobachtungsfunktion: $O(s, o)$
 - Gewinnfunktion: $R(s)$

- Neues für den Agenten:
 - Glaubenzustand $b(s)$

08.02.2008 14:08 Jochen Bernhardt
Fabien Gippert 12

Das Beobachtungsmodell ist die Wahrscheinlichkeit die Beobachtung o im Zustand s wahrzunehmen.

Glaubenzustand speichert für jeden Zustand die Wahrscheinlichkeit gerade in ihm zu sein.

Glauben und Beobachten

			+1
			-1
START			
1	2	3	4

(a)

(b)

0.111	0.111	0.111	0.000
0.111		0.111	0.000
0.111	0.111	0.111	0.111

(a)

0.300	0.010	0.008	0.000
0.221		0.059	0.012
0.371	0.012	0.008	0.000

(b)

0.622	0.221	0.071	0.024
0.005		0.003	0.022
0.003	0.024	0.003	0.000

(c)

0.005	0.007	0.019	0.775
0.034		0.007	0.105
0.005	0.006	0.008	0.030

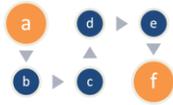
(d)

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

13

- a) Am Anfang – Zustand unbekannt!
- b) Nach fünf mal Links
- c) Nach fünf mal oben
- d) Nach fünf mal rechts



Partielle MEPs- Abschluss

- Aktionen und Taktiken werden auf Basis der Glaubenszustandes ausgewählt
- Der Glaubensraum ist stetig
- Informationen ändern Glaubensraum
- Können Wert von Informationen beinhalten

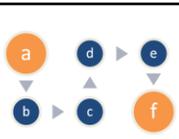
- Lösbar über Umwandlung in MEP!

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

14

In unserem Beispiel ist der Glaubensraum gerade mal ein 11 Dimensional-Stetiger Raum.



Entscheidungstheoretische Agenten

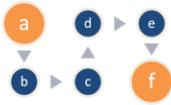
- Beispiel:
 - Früher: Diagnosen über Wahrscheinlichkeiten ermitteln
 - Heute: Diagnose über den Nutzen ermitteln
- Aber wie ermitteln ich die Diagnose über einen Nutzen?
- Kurzform: Die Aktionen auswählen, welche die Wahrscheinlichkeit auf richtige Diagnose maximieren

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

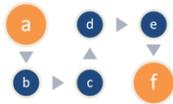
15

Beispiel : Wahrscheinlichste Diagnosen: „Gesund“, „Erkältet“, „Lungenkrebs“
Russel: 739



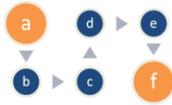
Entwurfsmuster

- Kausales Modell anlegen
- Vereinfachung zum qualitativen Entscheidungsmodell
- Wahrscheinlichkeiten zuordnen
- Nutzen zuordnen
- Modell überprüfen & verfeinern
- Sensibilitätsanalyse

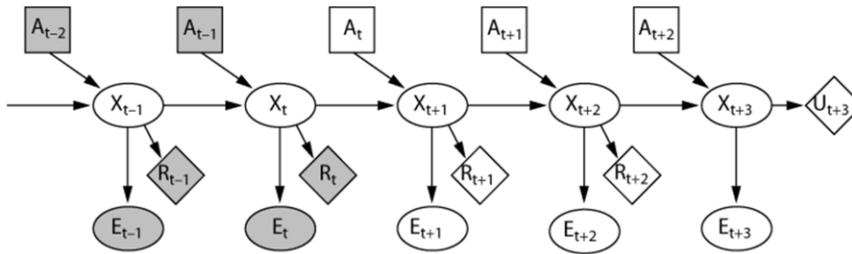


Nutzenbasierter Agenten (Entscheidungstheoretischer)

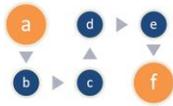
- Bestandteile:
 - Übergangs- und Beobachtungsmodelle
 - (dynamische Bayessche Netzwerke)
 - dynamisches Entscheidungsnetzwerk
 - Filteralgorithmus
 - Einpflegen von Wahrnehmungen und Aktionen
 - Aktualisieren der Glaubenszustände
 - Vorwärtsprojektion zur Entscheidungsfindung



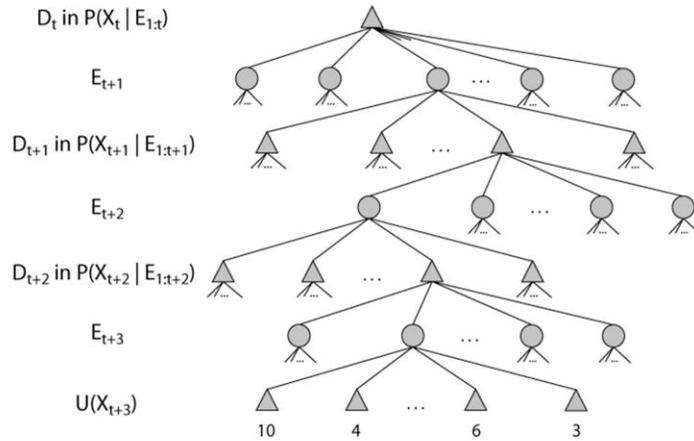
Entscheidungsnetzwerk



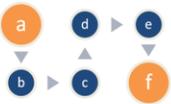
- A_t Aktion zum Zeitpunkt T
- X_t Menge der Zufallsvariablen zum Zeitpunkt T
- E_t Evidenzen zum Zeitpunkt T



Vorrausschaulösung



- U: Nutzen des Pfades,
- D: Aktionen zum Zeitpunkt T, E: Evidenzen zum Zeitpunkt T



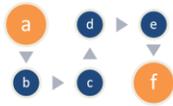
The diagram shows six nodes: 'a' (orange circle), 'b' (blue circle), 'c' (blue circle), 'd' (blue circle), 'e' (blue circle), and 'f' (orange circle). Arrows indicate connections: 'a' to 'b' (down), 'a' to 'd' (right), 'b' to 'c' (right), 'c' to 'd' (up), 'd' to 'e' (right), and 'e' to 'f' (down).

Spieltheorie

- Entscheidungssysteme mit mehreren Agenten
- Recht gut erforscht
- Komponenten:
 - Spieler / Agenten
 - Aktionen
 - Auszahlungsmatrix
 - Gewinne/Verluste für Aktionen

08.02.2008 14:08 Jochen Bernhardt
Fabien Gippert 20

Vorstellung des ganzen, worum geht's, Agenten/Spieler, Aktionen, Auszahlungsmatrix



Einbrecher - Beispiel

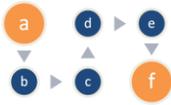
	Anna gesteht	Anna leugnet
Bodo : gestehen	$A = - 5, B = - 5$	$A = - 10, B = - 0$
Bodo: leugnen	$A = 0, B = - 10$	$A = - 1, B = - 1$

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

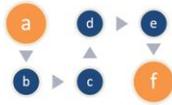
21

Das Beispiel zum erklären



Begriffe Spieltheorie

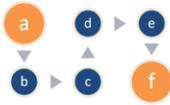
- Strategie = Taktik
- Reine Strategie = deterministische Taktik
- Strategie für ein Spiel mit einem Zug = Aktion
- Gemischte Strategie= zufallsgesteuerte Taktik
- Strategiegleichgewicht
- Nash-Gleichgewicht
- Pareto-Optimal
- Pareto-dominiert



Beispiel 1

	Anna gesteht	Anna leugnet
Bodo : gestehen	A = - 5, B = - 5	A = - 10, B = - 0
Bodo: leugnen	A = 0, B = - 10	A = - 1, B = - 1

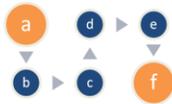
- Dominante Strategie: „gestehen“
- Strategiegleichgewicht: „Beide gestehen“
- Nash-Gleichgewicht: „Beide gestehen“
- Pareto-Dominanz: Gestehen -> Leugnen



Beispiel 2

	Acme : DVD	Acme : CD
Best : DVD	A = 9, B = 9	A = - 4, B = - 1
Best : CD	A = -3, B = - 1	A = 5, B = 5

- Dominante Strategie: keine
- Strategiegleichgewicht: keine
- Nash-Gleichgewicht: DVD/DVD und CD/CD
- Pareto-Dominanz: z.B.: DVD/DVD -> Rest



Nutzen Grundlagen der Spieltheorie

- Dominante Strategie: Vorhersage eines Zuges eines Rationalen Agenten ist.
- Strategiegleichgewicht: Starker Anziehungspunkt – Lösungen werden hinhin gezogen!
- Nash-Gleichgewicht: Jedes Spiel hat eins. Vorhersage wann Strategiewechsel wichtig sind.
- Pareto-Dominanz: Finden besserer Ergebnisse

Mechanismenentwurf

- Das festlegen von Regeln bei denen eine globale Nutzenfunktion maximiert wird
- obwohl mehrere Agenten als einzelne eine eigene rationale Strategie verfolgen.
- Anders: Aus eingeschaenkten Systemen intelligente Systeme formen ohne das diese unbedingt zusammen arbeiten muessen.

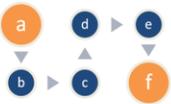
08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

26

Aehnlich wie bei Teams. Mehrere Menschen koennen Ziele erreichen die fuer jeden einzelnen unerreikbaar waeren.

Haeufig gebraucht in Wirtschaft und Politik.



Mechanismenentwurf

- Einsatzbeispiele
 - Versteigern billiger Flugtickets
 - Routing von TCP-Packeten
 - Verteilung medizinischer Ausstattung auf Krankenhäuser
 - Zusammenarbeit von Roboterfußballspielern

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

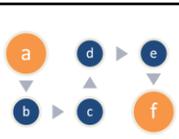
27

Mechanismusentwurf wurde in 90er Jahren populär als mehrere Nationen versuchten Lizenzen zur Funkübertragung zu verkaufen.

-> hundert Millionen Dollars potentieller Gewinn verloren wegen schlechtem Mechanismusentwurf.

Mechanismenentwurf = inverse Spieltheorie

Entwerfen eines Spiels um das Ziel der Zusammenarbeit zu erreichen (Folie siehe vorher)



Mechanismenentwurf

- Idee:
Maximierung des Nutzen der einzelnen Agenten maximiert den Nutzen für alle.

$$- U = \sum_i U_i$$

08.02.2008 14:08

Jochen Bernhardt
Fabien Gippert

28

Ist banal aber funktioniert auch nicht.

Tragedie der Gemeindewiese:

Bauern lassen Vieh kostenlos auf der Gemeindewiese weiden.

Folge: Wiese zerstört. -> negativer Nutzen für jeden einzelnen Bauern

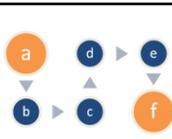
Das Benutzen der Wiese war aber rational, da es den Nutzen des Einzelnen steigerte.

Andere Beispiele:

- Kostenlose Entsorgung von Umweltschmutz in Atmosphäre und Meere

- Kapitalismus (Wenn jeder versucht reich zu werden steigt der Gesamtwohlstand der Gesellschaft)

Agenten können andere Agenten mit Aktionen so beeinträchtigen, dass der globale Nutzen sinkt.



Lösung Mechanismenentwurf

- Agenten müssen für alle externen Ressourcen "mitverantwortlich" gemacht werden.
- Aufgabe ist es Mechanismen zu finden die selbst wenn der Agent die Kausalität seines Handelns nicht selber abschätzen kann, den Agenten eine einfache Entscheidung ermöglichen.

08.02.2008 14:08 Jochen Bernhardt
Fabien Gippert 29

Notizen:

Beispiel der Wiese:

- a) Bauern muessen gleichberechtigt Reparatur der Wiese bezahlen.
- b) Bauern duerfen Wiese nicht mehr Kostenlos Nutzen

zu a) Bauern die die Wiese mehr genutzt haben als andere bezahlen das gleiche. Nutzen dieser Bauern

war goresser und evtl gross genug fuer den einzelnen Bauern das es sich noch lohnt.

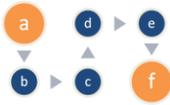
zu b) Preis muesste hoch genug sein das der Nutzen der Bauern zu gering ist die Gemeindewiese zu nutzen.

Durch erstellen von Mechanismen das verhalten von unabhaengigen Agenten steuern. Quasi Spielregeln entwerfen.

Künstliche Agenten einfache und komplexe Entscheidungen



Jochen Bernhardt
Fabian Gippert



Für den Heimweg

- Anwendungsbeispiele mal anders:
 - Schon mal versucht mit Nutzentheorie das Fernbleiben der FH dem Prof zu erklären?
 - Versucht R2D2 zu verstehen?
 - Ne Taktik fürs „Tic Tac Toe“ zu entwickeln?
 - Ne „Auszahlungsmatrix“ einer Beziehung?

Nicht verzagen wenn im nächsten Spiel rausrutscht:
„Es ist egal welche Figur ich ziehe... wir werden vom dominantem Strategiegleichgewicht angezogen“