

Prüfungsprotokoll

Diplom HF

Jörg Müller

SS 2007

(Advanced AI Techniques, Introduction to Mobile Robotics)

1 Advanced AI Techniques

1.1 Reinforcement Learning (RL), Burgard

Worin besteht RL?

Bestimmung von optimalen Aktionen, z.B. in Markov Decision Processes (kurz erklärt)

Unterschied zwischen RL, Supervised Learning (SL)

Trainingsinput ist bei RL nur Reward, bei SL die optimale Aktion

Dynamische Programmierung (DP): Wie wird die optimale Aktion bestimmt?

Mit Policy-Iteration oder Value-Iteration, (zu) kurz erklärt, hat nochmal nachgefragt, Bellman Equation hingeschrieben.

Monte Carlo (MC): Was ist der Unterschied zu DP?

Es werden nicht alle möglichen Nachfolgezustände betrachtet, sondern anhand einer gesampleten Sequenz gelernt.

Vorteile von MC?

Kann bei Interaktion mit Welt lernen; kein Modell notwendig; geringerer Einfluss einer Markov-Verletzung.

Temporal Difference (TD)?

Zu MC abgegrenzt: Zusätzlich zu Sampling Bootstrapping. Backup-Schritt aufgeschrieben.

Was ist das exploration exploitation Dilemma?

Man will möglichst oft die beste Aktion ausführen, um den Reward zu maximieren (exploit), muss aber auch explorieren, um den gesamten Zustands-Aktions-Raum abzudecken.

1.2 Game Theory, Nebel

Unterschied der Spieltheorie zu RL?

Mehrere Agenten, ...

Strategische Spiele: Definition

$G = \langle N, (A_i), (U_i) \rangle$, erklärt ...

Was ist ein Nash Equilibrium (NE)?

Gleichgewichtszustand (Aktionsprofil), bei dem es sich für keinen Spieler lohnt, davon abzuweichen, d.h. die Auszahlung bei Abweichung nur geringer werden kann.

Gibt es immer ein NE?

Nicht für pure Strategien. Nach Satz von Nash gibt es aber immer ein NE in gemischten Strategien.

Was sind denn gemischte Strategien

$N = \langle N, (\Delta(A_i)), (U_i) \rangle$, erklärt ...

Wie berechnet man NE in gemischten Strategien?

Mit Support-Lemma (Support und Support-Lemma aufgeschrieben/erklärt): Nutzen der puren Strategien aus Support sind gleich. Lösung des Gleichungssystems für zwei Spieler und zwei Aktionen einfach.

Kompetitive Spiele: Zentrales Konzept?

Maximinimierer definiert, Maximinimierer Satz aufgeschrieben.

Berechnung des NE?

Aktionen $A_1 = \{a_{11}, \dots, a_{1n}\}$, $A_2 = \{a_{21}, \dots, a_{2m}\}$.

Lineare Gleichungen:

- $\sum_i \alpha(a_{1i}) = 1$
- $\alpha(a_{1i}) > 0 \quad \forall i \in \{1, \dots, n\}$
- $\sum_i \alpha(a_{1i}) u_1(a_{1i}, a_{2j}) \geq u \quad \forall j \in \{1, \dots, m\}$

Lösung durch Maximierung von u mit linearer Programmierung.

Definition Bayessche Spiele

...

2 Introduction to Mobile Robotics, Burgard

Wo wir gerade bei Bayes sind: Was sind Bayessche Filter?

Recursive Bayes Filter Gleichung hingeschrieben.

Wozu wird der Bayes Filter verwendet?

Z.B. Roboter Lokalisierung, allgemein Abschätzung eines Zustands mit Observationen und Zustandsübergängen.

Was für Implementierungen des Bayes Filters gibt es?

Kalman Filter, Diskrete Filter, Particle Filter. Jeden in einem Satz beschrieben.

Wie funktioniert ein Particle Filter?

Repräsentation der Wahrscheinlichkeitsverteilung durch Particles. Algorithmus beschrieben:

- Resampling nur sehr kurz erklärt. Hat nachgefragt → Stochastic Universal Resampling genau erklärt
- Motion Modell: Für jeden Particle Sampling normalverteilter Fehler in Odometrie-Messung $\langle \delta_{\text{rot1}}, \delta_{\text{trans}}, \delta_{\text{rot2}} \rangle$ und Integration der Odometrie mit diesem Fehler.
- Sensor Modell: Gewichtung der Particles $= \frac{\text{target distribution}}{\text{proposal distribution}} = p(z_t | x_t)$.

Dabei jeweils auf die entsprechenden Teile der Recursive Bayes Filter Gleichung hingewiesen.

Wie bekommt man das Sensormodell?

Beam-basierte Sensormodell mit Annahmen beschrieben; Modellierung des Sensors und der Fehlerquellen, Parametrisierung der Fehlerquellen, Lernen der Fehlerparameter.

Wie funktioniert ein Kalman Filter?

Modell ist:

$$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t$$

$$z_t = C_t x_t + \delta_t$$

Komponenten der Gleichungen beschrieben. Mehr wollte er nicht hören.

Nachteile des Kalman Filter?

nur für lineare Systeme mit normalverteiltem Rauschen.

Was macht man im nichtlinearen Fall?

EKF: linearisierung durch Taylorentwicklung 1. Ordnung.

$$x_t = g(x_{t-1}, u_t)$$

$$z_t = h(x_t)$$

Jakobi-Matrix definiert, Taylorentwicklung hingeschrieben.

$$\text{Übergang z.B. von } A_t \text{ zu } G_t = \left. \frac{\partial g}{\partial x_{t-1}} \right|_{\mu_{t-1}}.$$

Malt eine eindimensionale Normalverteilung und einen Pfeil für die Odometrie-Messung auf. Wie sieht die Abschätzung des Kalman-Filters aus?

Erwartungswert wird um u verschoben. Varianz wird größer, also Kurve flacher.

Malt jetzt noch eine Observation dazu. Wie geht die ein?

Observation hat geringere Varianz, also wird der neue Erwartungswert zwischen dem bisherigen und dem der Observation, aber näher an dem der Observation liegen. Die Varianz ist kleiner, als die der bisherigen Abschätzung und die der Observation.

Warum wird die Varianz kleiner?

Regel für Multiplikation zweier Normalverteilungen aufgeschrieben: $\Sigma' = \frac{1}{\Sigma_1^{-1} + \Sigma_2^{-1}}$.

Was macht man bei SLAM?

(Wollte erst Rao-Blackwellization erklären, aber er war noch beim Kalman Filter.)

Zustand enthält beim EKF-SLAM Roboterposition und (x, y) -Positionen

der Landmarks.
Kovarianzmatrix mit Kreuz-Korrelationen.

Wieso sieht die Kovarianzmatrix wie ein Schachbrett aus?

x -Positionen der Landmarks korrelieren untereinander, genauso y -Positionen.
Im Grenzfall sind die Landmarks voll korreliert. Da die Anordnung gewöhnlich $l_{1x}, l_{1y}, l_{2x}, l_{2y}, \dots$ ist, entsteht ein Schachbrett.

Bemerkung

Es war insgesamt eine entspannte Prüfungsatmosphäre
Prof. Burgard hat manchmal nachgefragt, weil ich wohl nicht genau das gesagt habe, was er hören wollte oder für wichtig hielt, aber sonst haben beide ruhig und präzise Fragen gestellt.

Viel Glück bei Euren Prüfungen