

## MMKI-Praktikum SS06: Aufgabe 6D

Die Aufgabe 6D schließt den Aufgabenteil „Reinforcement Learning“ ab.  
Für diese Aufgabe können Sie nun all Ihre Erfahrungen mit den verschiedenen Aspekten des RL einfließen lassen.  
Ganz bewußt lassen wir Ihnen großen Freiheiten bei der konkreten Ausgestaltung der Aufgabe – seien Sie kreativ.  
Ziel ist es, dem Simloid das schnelle Kicken eines Balls beizubringen.

### Erläuterung:

- In der Version 1.2 des Simulators liegt ein (simulierter) Ball direkt vor dem Roboter. Der bereitgestellte DemoClient6D zeigt, wie der Roboter einen sehr einfachen handcodierten Kick ausführen kann.
- Dabei fällt auf, dass weder die Beinbewegung optimal ist, noch der Roboter das Gleichgewicht halten kann.

### Aufgabe:

- Verstehen Sie den bereitgestellten DemoClient als Ausgangsbasis für ihre eigenen Ideen zum Thema Kick.
- Überlegen Sie, bei welchen Aspekten dieser komplexeren Aufgabe sich Verfahren des RL sinnvoll einsetzen lassen.
- Erzeugen Sie einen Controller, der den Simloid reproduzierbar und ohne umzufallen den Ball kicken lässt. Lösen Sie dabei mindestens ein Teilproblem durch ein maschinelles Lernverfahren (vorzugsweise RL).
- 

### Abgabeformat:

Neben der Programmierung / Konfiguration des Lernverfahrens und dessen Anwendung zur Lösung der Aufgabe sind folgende Punkte ebenfalls Teil der Aufgabe: (bitte kurz und wesentlich behandeln)

- Beschreiben Sie die konkrete von Ihnen gestellte Aufgabe, bzw. das Zielverhalten.  
An welchen Stellen verwenden Sie dabei Lernverfahren und warum scheinen Ihnen diese dafür geeignet zu sein?  
Welchen Teil der Gesamtaufgabe lösen Sie ggf. explizit von Hand und warum?
- **Beschreiben Sie das von Ihnen gewählte Lernexperiment und den Verlauf der Lernphase!** Begründen Sie Ihre Wahl.  
Im Fall von Reinforcement Learning-Verfahren sind dabei auch die Parameter wie Lernrate, Trace-Value, Lernpolitik, ... relevant.  
Sofern Sie sich für ein besonders neues oder trickreiches Verfahren entschieden haben, erläutern Sie es kurz.
- Beschreiben und begründen Sie den von Ihnen modellierten Zustandsraum.  
Welche Zustandsvariablen (Dimensionen) gibt es, wie werden diese berechnet, wie ist der Zustandsraum aufgebaut?  
Verwenden Sie einen diskreten Zustandsraum, ist natürlich auch die Form der Diskretisierung wichtig.
- Beschreiben und begründen Sie den von Ihnen modellierten Aktionsraum.  
Welche Aktionstypen oder konkrete Aktionen gibt es, wie werden diese ggf. in die nativen Simloid-Kommandos umgerechnet?  
Verwenden Sie einen diskreten Aktionsraum, ist wiederum die Form der Diskretisierung relevant.
- Sollten Sie, was durchaus erwünscht ist, verschiedene Verfahren, Modellierungen und Experimente probiert haben, um zu einer Lösung zu gelangen, beschreiben Sie kurz ihre Herangehensweise und die Auswirkungen ihrer Änderungen.

Um den Lernerfolg zu dokumentieren, sind mind. 2 graphische Plots abzugeben:

- Für die Lernphase ist ein Diagramm zu erstellen, welches ein geeignetes Maß für den Lernerfolg gegen die Zeit oder die Anzahl der Takte oder Episoden aufträgt.  
Die Zeitskala soll bei Null starten und bis zum Erreichen eines hinreichend stabilen Verhaltens gehen.
- Der gelernte Controller soll drei mal ausgehend vom Initialzustand aufgerufen werden, um die korrekte Lösung zu dokumentieren.  
Für diese Episoden soll jeweils ein **geeignetes** Gütemaß (der Gesamtreward ist i.d.R. keins) dokumentiert werden.

Abzugeben sind ferner die Quellcodes der Programme, sowie zwei Startskripte:

1. Lernphase: Start des Simloid und des Controllers, der die Struktur lernt.
2. Testphase: Start des Simloid und des Controllers mit der gelernten Struktur.

**Hinweise:**

Ausgehend vom DemoClient wäre der Einsatz von Lernverfahren z.B. für folgende Teilprobleme denkbar:

- Sie programmieren eine gute und schnelle Beinbewegung (von Hand).  
Normalerweise wird der Roboter dabei umfallen.  
Lassen Sie also die gleichzeitig notwendige Stabilisierung durch andere Gelenke (z.B. mit den Armen) durch das Lernverfahren lernen. Die Beschleunigungssensoren können dabei hilfreich sein.
- Sie lassen den Roboter die komplexe Beinbewegung für einen guten Schuss lernen (aber bitte nicht nur ein Gelenk) und stabilisieren Sie den Roboter von Hand.  
Ein geeignetes Gütekriterium wäre z.B. die erzielte Ballentfernung nach n Takten
- Sie lassen den Roboter die gesamte Bewegungsfolge für einen guten stabilen Schuss lernen.
- Sie haben haben andere tolle Ideen, was der Roboter mit dem Ball machen könnte.

Viel Erfolg