

# Realisierung eines Tic-Tac-Toe-spielenden NAO-Roboters mittels automatischen Erlernens der Spielstrategie

Andy Klay

Masterarbeit • Studiengang Informatik Master • Fachbereich Informatik und Medien • 13.08.2015

## Aufgabenstellung

Ziel der Arbeit ist die Entwicklung einer Applikation, die einen NAO-Roboter (Abb. 1) gegen einen Menschen Tic-Tac-Toe spielen lässt. Ein Schwerpunkt der Arbeit ist die geeignete Realisierung eines Lernvorgangs mit dem eine Spielstrategie automatisch erlernt wird. Der Lernvorgang soll durch die Messung der Spielstärke evaluiert werden. Die Applikation soll modular entworfen werden, so dass ein einfacher Austausch oder Erweiterung von Komponenten ermöglicht wird. Die Komponenten Spielsteuerung und Lernmodul sollen möglichst unabhängig vom konkreten Spiel sein. Eine Teilaufgabe besteht in der Erkennung der Spielsituation mit Hilfe der Bildverarbeitung. Zur Ansteuerung der Aktorik ist eine sinnvolle Schnittstelle unter Berücksichtigung der NAO-Plattform zu realisieren.

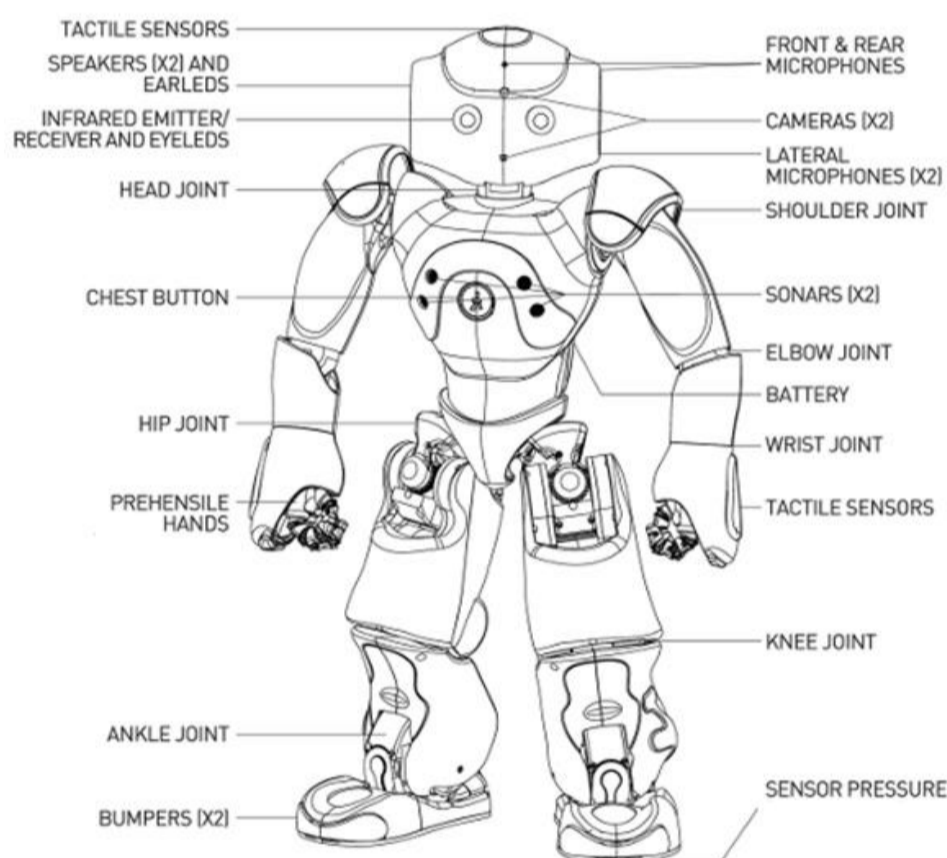


Abb. 1: Übersicht eines NAO-Roboters [Ald151]

## Konzept

Die Arbeit ist in drei große Problemfelder aufgeteilt. Dazu zählt die optische Erkennung der Spielsituation, die Entwicklung der Lernstrategie für das Spiel und die Schnittstelle zur Aktorik. Des Weiteren werden alle drei Teile im Spielablauf miteinander verknüpft. Die Entwicklung der Lernstrategie steht dabei im Vordergrund. Hierbei werden verschiedene Konzepte aus dem Reinforcement Learning betrachtet und diskutiert. Der implementierte Lernprozess wird durch Tests gegen einen Zufallsspieler geprüft. Notwendige Mittel zur Implementierung der optischen Erkennung werden geklärt und angewendet. Schließlich wird das Gesamtsystem mit NAO-Roboter gegen einen menschlichen Gegner getestet und Ergebnisse dokumentiert. Zuletzt werden die Ergebnisse und zukünftige Erweiterungen zusammengefasst.

## Reinforcement Learning

Reinforcement Learning (Abb. 2) ist das Lernen mittels Interaktion in einer Umwelt. Dabei wird von dem Agenten (dem Lernenden) gelernt, Aktionen zu Zuständen zuzuordnen, so dass die Rückmeldung der Umwelt, der Reward (ein numerischer Wert), maximiert wird. Der Agent lernt anhand der Konsequenzen seiner Aktionen in der Umwelt. Entscheidend ist dabei die sogenannte Policy (Zielvorgabe) und gibt an, nach welchen Regeln der Agent welche Aktion wählt. [Sut98]

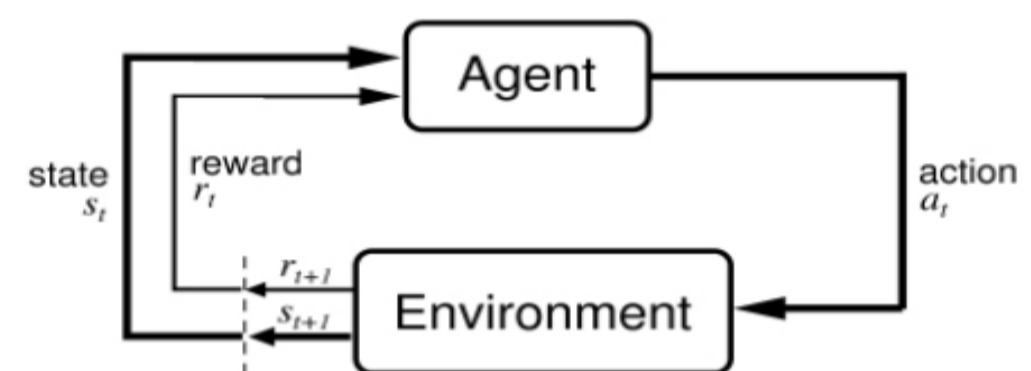


Abb. 2: Prinzip des Reinforcement Learning [Sut98]

## Künstliche Neuronale Netze

Zur Speicherung der Abbildung zwischen Zuständen und deren Güte (Value) für den Spielausgang kommen Künstliche Neuronale Netze zum Einsatz. Ihre Fähigkeit trägt dazu bei, dass auch die Werte für dem Agenten bisher unbekannte Zustände errechnet werden.

## Temporal Difference Learning

Ein Problem beim Erlernen von Spielen ist grundlegend, dass nur für den terminierenden Zug ein Reward bekannt ist. Hierbei kommt das TD-Learning (Abb. 3), eine Reinforcement Learning-Variante, zum Einsatz. Dadurch wird es möglich, den Reward des terminierenden Zustands auf die vorherigen Zustände des Spielverlaufs abzubilden.

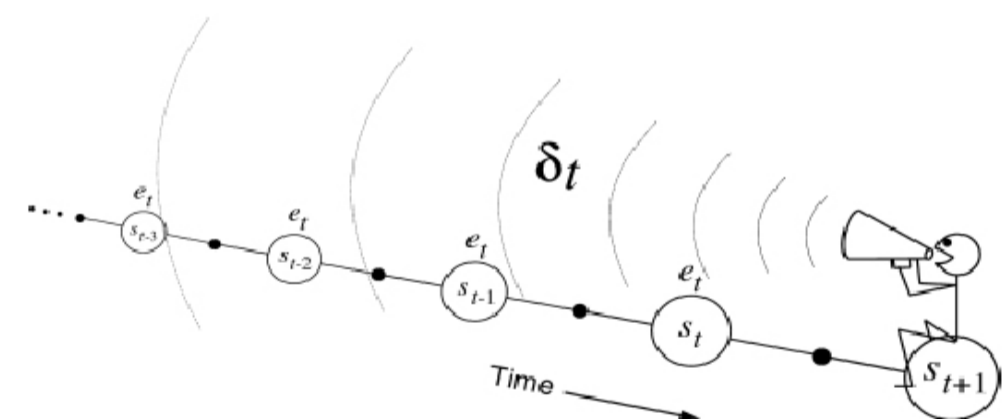


Abb. 3: Prinzip des TD-Learning [Sut98]

## Ergebnisse

Im Großen und Ganzen wurde gezeigt, dass sich das TD-Learning als spezielle Methode des Reinforcement Learning für die Problematik des Erlernens des Tic-Tac-Toe-Spiels eignet. Es ist möglich gegen den Roboter im Tic-Tac-Toe Spiel zu spielen und es wird dabei automatisch aus dem Spielverlauf gelernt. Benötigte Kenntnisse zur Lösung der gestellten Aufgabe wurden geklärt und wichtige Entscheidungen begründet. Geeignete Werte für die verschiedenen Parameter des Lernverfahrens durch stichprobenartige Tests ermittelt.

## Fazit

In der Arbeit wurde gezeigt, dass ein automatisches Erlernen einer Spielstrategie mit Hilfe des TD-Learning auf dem NAO-Roboter möglich ist. Der Lernerfolg konnte deutlich mit Messungen belegt werden. Jedoch lässt die Arbeit viele Möglichkeiten zur Erweiterung und Verbesserungen des Lernvorgangs offen.

## Quellen

[Ald151] Aldebaran: Aldebaran - NAOqi DOCUMENTATION (2015) [<http://doc.aldebaran.com/2-1/index.html>] (Zugriff: 18.03.2015).

[Sut98] Richard S. Sutton, Andrew G. Barto.: Reinforcement Learning: An Introduction (Cambridge, 1998).