

Intrinsische Motivation und künstliche Neugier

Michael Lampert

Institut für Informatik,
Humboldt-Universität zu Berlin

SE »Biologisch motivierte Lernverfahren« 2012

HUMBOLDT-UNIVERSITÄT ZU BERLIN



Gliederung

- 1 Intrinsische Motivation
 - Der Begriff
 - Am Bsp.: Das Tischtennispielen erlernen
- 2 Künstliche Neugier
 - Die gewünschten Effekte
 - Der Ansatz
- 3 Intelligent Adaptive Curiosity (IAC)
 - Was es ist
 - Ein paar Definitionen

Gliederung

- 1 **Intrinsische Motivation**
 - Der Begriff
 - Am Bsp.: Das Tischtennispielen erlernen
- 2 **Künstliche Neugier**
 - Die gewünschten Effekte
 - Der Ansatz
- 3 **Intelligent Adaptive Curiosity (IAC)**
 - Was es ist
 - Ein paar Definitionen

Allgemein

Der Begriff intrinsische Motivation bezeichnet das Bestreben, etwas um seiner selbst willen zu tun (weil es einfach Spaß macht, Interessen befriedigt oder eine Herausforderung darstellt).

Bei der extrinsischen Motivation steht dagegen der Wunsch im Vordergrund, bestimmte Leistungen zu erbringen, weil man sich davon einen Vorteil (Belohnung) verspricht oder Nachteile (Bestrafung) vermeiden möchte.[2]

als innerer/eigener Antrieb Lernbares zu Lernen

- Qualitäten einer Beschäftigung
 - »Neuigkeit«
 - Komplexität
 - Überraschung (bietet Unerwartetes)
 - Herausforderung
- Belohnung / Verstärkung durch Lernerfolg erfahrbar durch empfundene Steigerung der
 - Selbstwirksamkeit
 - Kompetenz
 - Autonomie
 - oder auch durch die konkrete, bewusste Feststellung etwas begriffen zu haben

- objektive, vollständige Komplexität ist hoch
 - z.B. Bewegungszustand des Balles - Lage, Richtung, Geschwindigkeit, Drall (8D)
- subjektive (Lern-)Komplexität ist beeinflussbar, also variabel
 - von oben durch den »Lehrer« bewusst begrenzbarm
 - Drall des Balles aus dem Spiel nehmen = 2D weniger
 - immer gleiche Geschwindigkeit = 1D weniger
 - etc.
 - von unten durch den Fortschritt des Schülers begrenzt
 - schon perfektes Beherrschen des Spiels ohne Drall
-> Lernkomplexität des Spiels mit Drall nur noch 2D

Gliederung

- 1 Intrinsische Motivation
 - Der Begriff
 - Am Bsp.: Das Tischtennispielen erlernen
- 2 Künstliche Neugier
 - Die gewünschten Effekte
 - Der Ansatz
- 3 Intelligent Adaptive Curiosity (IAC)
 - Was es ist
 - Ein paar Definitionen

- »progressive, inkrementelle« Entwicklung
 - »Komplexität der Aktivitäten« den vorhandenen Fähigkeiten angemessen und
 - dadurch über die Zeit zunehmend
- selbständige, aktive Entwicklung
 - ohne vorgekaute Lerninhalte
 - ohne vorgegebene Lernziele
 - selbstorganisiert

Grober Ansatz

- Bewertung von
 - »Neuigkeit«
 - Komplexität
 - Überraschung (bietet Unerwartetes)
 - Herausforderung
- Belohnung, die maximal ist, wenn diese Qualitäten auf mittlerem Niveau liegen

Praktischer Ansatz

Maschine, die ihren Einfluss auf ihre Umwelt erlernt.

Drei Gruppen von »Intrinsic Motivation Systems«:

- Error Maximization
 - größter erwarteter Vorhersagefehler
 - kann dazu führen, daß nur noch Unlernbares angegangen wird
- Progress Maximation
 - stärkste erwartete Minimierung des Vorhersagefehlers, und
 - Belohnung im Maße der tatsächlichen Minimierung
 - kann dazu führen, daß ständig zwischen Aktionen mit unterschiedlich guten Vorhersagemöglichkeiten gewechselt wird
- Similarity-Based Progress Maximization
 - nur ähnliche Situationen teilen sich den Verlauf / die Entwicklung der ihnen zugehörigen Vorhersagefehler
 - neue Herausforderung: Ähnlichkeit von Situationen bestimmen

Gliederung

- 1 Intrinsische Motivation
 - Der Begriff
 - Am Bsp.: Das Tischtennispielen erlernen
- 2 Künstliche Neugier
 - Die gewünschten Effekte
 - Der Ansatz
- 3 Intelligent Adaptive Curiosity (IAC)
 - Was es ist
 - Ein paar Definitionen

- *It is a motivation, or **drive**, ... being about ... the maintenance of an abstract dynamic cognitive variable: the learning progress, which must be kept maximal. This definition makes it an intrinsic motivation.*
- *It is called **curiosity** because maximizing the learning progress pushes (as a side effect) the robot towards novel situations in which things can be learned.*
- *It is **adaptive** because the situations that are attractive change over time, indeed, once something is learned, it will not provide learning progress anymore.*
- *It is called **intelligent** because it keeps, as a side effect, the robot away both from situations which are too predictable and from situations which are too unpredictable.[1]*

Definitionen

$S(t)$ Vektor aller reellen Sensorwerte $s(t)$

$M(t)$ Vektor aller reellen Aktorwerte $m(t)$

$SM(t)$ der sensomotorische Kontext

$(SM(t), S(t+1))$ die Erfahrungen, die der Roboter sammelt



R_n Region im sensomotorischen Raum

C_1 Kriterium, wann eine Region geteilt wird - hier $|R_n| > 250$

C_2 Kriterium, wie eine Region geteilt wird - hier gewichtete

Varianzen

Quelle

-  Pierre-Yves Oudeyer, Frédéric Kaplan, and Véréna Hafner. Intrinsic motivation systems for autonomous mental development, 2007.
-  Wikipedia. Motivation — wikipedia, die freie enzyklopädie, 2012. [Online; Stand 26. Juni 2012].