

# **Personalisierung der Mensch-Roboter-Interaktion durch sozial-sensitives Lernen**

M.Sc. **Hannes Ritschel**, M.Sc. **Tobias Baur**, Prof. Dr. **Elisabeth André**, HCM-Lab, Universität Augsburg

## **Kurzfassung**

Persönlichkeit ist ein wesentliches Element der Mensch-Roboter-Interaktion, sowohl von Seiten des Nutzers als auch des Roboters. Aus Studien geht hervor, dass die Anpassung der Persönlichkeit eines Roboters an das menschliche Profil die Interaktion ansprechender gestalten kann. Aus neuesten Erkenntnissen zeichnet sich jedoch auch ab, dass die Persönlichkeit des Roboters nicht zwingend der des Menschen entsprechen muss, sondern dass der Aufgabenkontext bei der Frage nach einem geeigneten Roboterprofil eine entscheidende Rolle spielt. Gegenstand der aktuellen Untersuchungen ist deshalb ein Anpassungsprozess, der die Präferenzen des Nutzers automatisch erlernen soll. Als algorithmische Basis dient hierbei das Bestärkende Lernen, in Kombination mit menschlichen, sozialen Signalen. Als Beispielszenario dient ein Roboter in der Funktion eines Geschichtenerzählers, der seine Persönlichkeit in Form von linguistischem Stil durch Generierung natürlicher Sprache ausdrücken und unter Rücksichtnahme auf soziale Signale für den jeweiligen Nutzer anpassen kann.

## **1. Hintergrund**

Für soziale Roboter spielt Persönlichkeit eine wesentliche Rolle. Durch Ausdruck von Persönlichkeit wird Mensch-Roboter-Interaktion interessanter und attraktiver, zudem begünstigt dies die Schaffung sozialer Bindung zwischen Mensch und Maschine [1]. Ergebnisse der aktuellen Forschung weisen darauf hin, dass viele Menschen Roboter bevorzugen, die ein ähnliches Persönlichkeitsprofil zeigen wie sie selbst, und dass die Anpassung der Persönlichkeit des Roboters an den Menschen dazu führt, dass die Interaktion als ansprechender wahrgenommen wird [2]. Weitere Ergebnisse deuten aber auch an, dass der Aufgabenkontext einen wesentlichen Einfluss darauf hat, ob ein ähnliches oder entgegengesetztes Persönlichkeitsprofil präferiert wird [3]. Unabhängig von der konkreten Ausprägung der Persönlichkeit ist Anpassung an den Menschen im Allgemeinen wichtig, um menschliche Bedürfnisse und Präferenzen erfüllen und lernen zu können. Dies

wiederum erlaubt es dem Roboter, Personalisierung zu realisieren und die Interaktion auch langfristig ansprechend zu gestalten [4].

Bei der Implementierung und Realisierung neuer Mensch-Roboter-Interaktionsszenarien ist zunächst nicht bekannt, welche Persönlichkeit der Roboter verkörpern und ausdrücken sollte, da es hierfür eine Vielzahl möglicher Einflussfaktoren gibt. Aus diesem Grund beschäftigt sich unsere Forschung als Alternative zum Festlegen auf ein ähnliches oder entgegengesetztes Persönlichkeitsprofil mit einem maschinellen Lernverfahren, um menschliche Präferenzen automatisch schätzen und den Roboter daran anpassen zu können. Als Prototyp wurde für diesen Zweck ein Dialogszenario entworfen, in dem ein „Reeti“-Roboter dem Menschen Fakten über die Hauptcharaktere des Buches „Alice im Wunderland“ präsentiert [5] [6]. Die Persönlichkeit des Roboters äußert sich in der von ihm verwendeten Sprache. Dabei sind die Formulierungen nicht vordefiniert, sondern werden automatisch mit unterschiedlichem Extraversionsgrad generiert. Extraversion ist eine der „Big Five“ Dimensionen [7], mit deren Hilfe Persönlichkeit beschrieben werden kann, welche sich unter anderem auch in den sprachlichen Äußerungen widerspiegelt. Ein Anpassungsprozess auf Basis von Bestärkendem Lernen und sozialen Signalen kontrolliert und optimiert die Persönlichkeit des Roboters, um das Engagement des Menschen längerfristig aufrecht zu erhalten.

## **2. Anpassungsprozess**

Um menschliche Präferenzen zu erlernen und Anpassung realisieren zu können, bedarf es einer wesentlichen Information: Rückmeldung bzw. Feedback vom menschlichen Nutzer, das Rückschlüsse auf die tatsächlichen Bedürfnisse und Präferenzen des Nutzers ermöglicht. Eine einfache, aber gleichzeitig auch sehr aufdringliche und unelegante Möglichkeit stellt zu diesem Zweck die explizite Befragung des Nutzers dar. Um dies zu umgehen, basiert unser Ansatz auf sozialen Signalen, die während der Interaktion ohnehin zu Verfügung stehen. Basierend auf einem Microsoft Kinect 2 Sensor und dem von uns entwickelten Social Signals Interpretation (SSI) Framework [8] werden die menschlichen sozialen Signale wahrgenommen, ausgewertet und darauf basierend das aktuelle menschliche Engagement mit Hilfe eines dynamischen Bayes'schen Netz geschätzt [9]. Diese Information gilt als Ansatzpunkt, ob die vom Roboter gezeigte Persönlichkeit den Nutzer engagiert oder nicht. In Kombination damit nutzen wir Bestärkendes Lernen als algorithmische Grundlage, das die Manipulation der Persönlichkeit und das Lernen menschlicher Präferenzen ermöglicht.

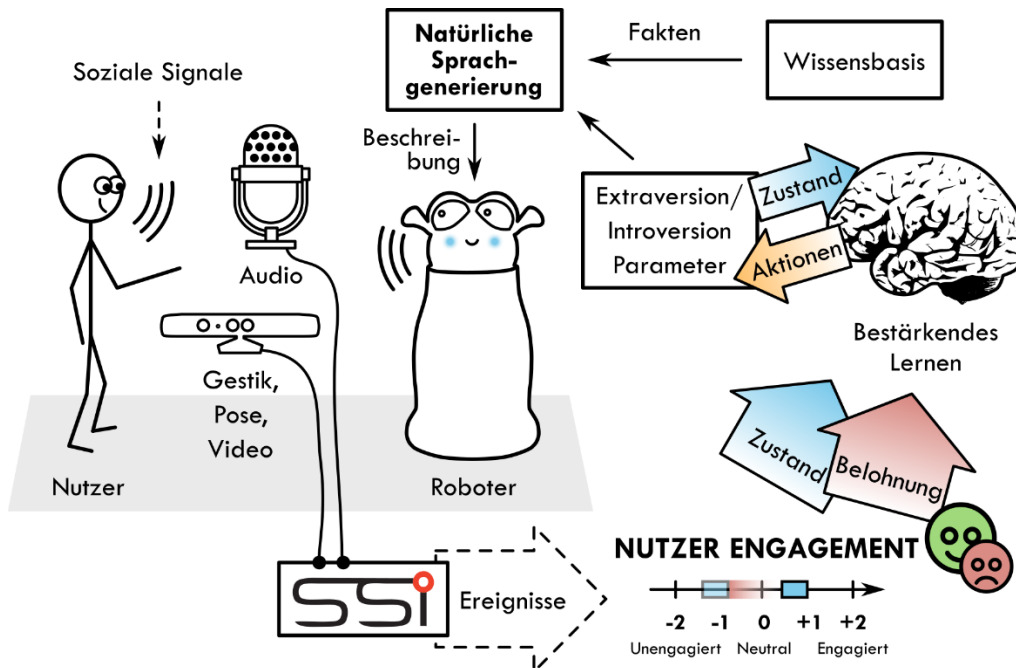


Bild 1: Überblick über den Prototypen

Bild 1 veranschaulicht die Interaktion und den Anpassungsprozess. In jedem Zeit- bzw. Lernschritt präsentiert der Roboter eine Beschreibung eines Charakters aus dem Buch, indem zur aktuellen Roboter-Persönlichkeit passende Sätze generiert werden. Während der Präsentation werden soziale Signale des Nutzers ausgewertet und für die Berechnung des Nutzer-Engagements herangezogen. Bestärkendes Lernen optimiert dann die Persönlichkeit des Roboters für die nächste Beschreibung.



Bild 2: Interaktionsszenario

## 2.1 Soziale Signale und Bestärkendes Lernen

SSI interpretiert während der gesamten Interaktion die Rohsignale der verwendeten Sensoren und berechnet darauf das aktuelle Nutzer-Engagement  $E_t$  zum Zeitpunkt  $t$ . Dieser Wert wird für die weitere Verarbeitung diskretisiert und kann Ganzzahlen im Intervall  $[-2; +2]$  annehmen, stellvertretend für die Extreme „sehr unengagiert“ sowie „sehr engagiert“. Der Lernprozess wiederum verwendet diese Information, um den Zusammenhang zwischen der vom Roboter ausgedrückten Persönlichkeit und dem Engagement des Nutzers zu erlernen. Dafür manipuliert das Bestärkende Lernen den aktuellen Extraversionswert  $X$  des Roboters, ebenfalls als Ganzzahlen im Intervall  $[-2; +2]$ , von „sehr introvertiert“ bis „sehr extrovertiert“. Drei Aktionen erlauben die Manipulation durch Inkrementieren oder Dekrementieren von  $X$ , oder keiner Änderung.

Der Zustandsraum kombiniert die aktuelle Extraversion des Roboters  $X$  sowie das zuletzt berechnete Nutzer-Engagement  $E_{t-1}$ . Dies ermöglicht dem Roboter das Finden einer optimalen Strategie, um das Nutzer-Engagement zu maximieren, abhängig von  $X$  und  $E_t$ . Als Belohnung dient hierbei die Differenz des Engagements  $\Delta E_t = E_t - E_{t-1}$  zwischen dem letzten Zeitpunkt  $t - 1$  und  $t$ . Diese verrät, ob das Engagement gefallen, gestiegen oder gleichgeblieben ist. Ziel des Roboters ist, das Engagement aufrecht zu erhalten bzw. im Optimalfall zu steigern.



Bild 3: Engagierter und unengagierter Nutzer

## 2.2 Natürliche Sprachgenerierung

Für die Erzeugung jeder Beschreibung kommt ein natürlichsprachliches Generierungssystem zum Einsatz. Einfluss auf die gebildeten Sätze haben hierbei verschiedene linguistische Variationsparameter [10], die speziell für den sprachlichen Ausdruck introvertierter oder extrovertierter Personen typisch sind. Inspiriert vom System Personage [10] werden

Äußerungen mit unterschiedlichem Extraversionsgrad gebildet, indem Parameter gezielt aktiviert und deaktiviert bzw. deren Werte abgeändert werden. Beispiele resultierender Äußerungen in englischer Sprache finden sich in Bild 4.

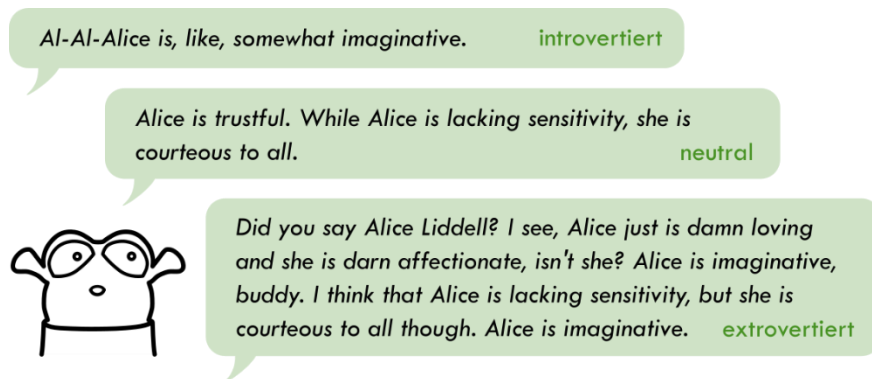


Bild 4: Formulierungen mit unterschiedlichem Extraversionsgrad

Abhängig vom aktuellen Extraversionswert  $X$  beeinflussen verschiedene Parameter in verschiedenen Stadien der Generierung eine Fülle von Merkmalen, wie beispielsweise die Anzahl und Auswahl der Fakten, deren Reihenfolge, Wiederholungen, Neuformulierungen, Satzbau, Wortwahl, Wortfülle, Stottern, etc. Die generierte Äußerung wird dem Nutzer durch das im Roboter eingebaute Text-To-Speech-System präsentiert.

### 3. Simulation

Im Rahmen einer Simulation waren zwei Aspekte von Interesse: die Korrektheit des modellierten Problems sowie die Fehlertoleranz bezüglich sozialer Signale. Um beides zu überprüfen, wurde ein initiales Experiment mit einer Nutzersimulation durchgeführt. Dabei steigt das Nutzer-Engagement, wenn die Roboter-Persönlichkeit den Präferenzen des simulierten Nutzers entspricht, andernfalls sinkt es. Dieses stark vereinfachte, deterministische Verhalten wird allerdings beeinflusst und randomisiert durch verschiedene Arten von Rauschen, die für das Experiment von Interesse waren. Es umfasst drei Aspekte, die wir bei der Arbeit mit sozialen Signalen erwarten: 1. Sensoren bringen immer Rauschen und physikalische Einschränkungen mit sich. Es muss damit gerechnet werden, dass die daraus extrahierte Information fehlerhaft sein kann. 2. Auch die Interpretation der Rohsignale ist nur eine Näherung der Realität und bringt somit weiteres Rauschen in den Lernprozess. 3. Die Nutzerreaktion an sich ist nichtdeterministisch, die gezeigten sozialen Signale müssen nicht nur von der Persönlichkeit des Roboters und dessen Verhalten abhängen.

Um einen ersten Eindruck vom Verhalten des Systems zu erlangen, wurde Q-Lernen mit  $\epsilon$ -greedy und Explorationsrate  $\epsilon = 0.2$  genutzt. Dieser Wert wurde gewählt, da das erwartete Rauschen zu fehlerhaftem Lernen führen kann. Um fehlerhafte Information eliminieren zu können, sollte hinreichend oft exploriert werden. Die Lernrate  $\alpha = 0.5$  wurde gewählt, um zum einen schnell zu lernen, aber gleichzeitig auch nicht alles Wissen zu verlieren, sofern Rauschen den Wert verfälscht. Im Rahmen der Simulation erhält der Roboter im Gegensatz zum echten Prototypen für  $\Delta E_t = 0$  eine Belohnung  $+0.5$  dafür, dass ein Sinken des Engagements verhindert wurde.

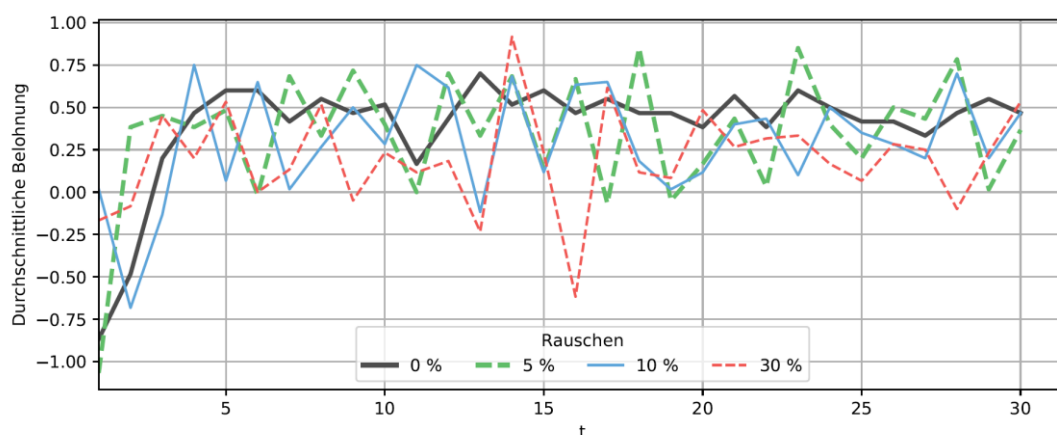


Bild 5: Initiale Simulation

Bild 5 zeigt die durchschnittliche Performanz des Agenten über 30 Nutzerinteraktion hinweg. In jeder dieser simulierten Interaktionen werden 30 Beschreibungen (= 30 Lernschritte) präsentiert, wobei jeder Nutzer initial eine zufällige Präferenz hat und der Roboter ohne Vorwissen mit neutraler Extraversion startet. Ohne Rauschen ist der Lernprozess im Mittel erwartungsgemäß relativ robust und nähert sich 0.5. Dies entspricht dem gewünschten Verhalten, da diese Belohnung genau dann vergeben wird, wenn sich das Engagement des Nutzers nicht mehr ändert, was wiederum primär dann auftritt, wenn der Roboter die der menschlichen Präferenz entsprechende Persönlichkeit gefunden hat. Negative Belohnungen sind auf Exploration zurückzuführen. Mit zunehmendem Rauschen ist erwartungsgemäß auch eine schlechtere Performanz erkennbar.

#### 4. Ausblick

Im Vergleich zur Simulation arbeitet der Prototyp mit echten sozialen Signalen. Wichtig ist deshalb die Evaluation des Systems, um neue Erkenntnisse bezüglich der Praxistauglichkeit,

dem Zusammenhang zwischen Nutzerpräferenz und gelernter Persönlichkeit sowie der Eignung und des Rauschverhaltens der Signalverarbeitung zu gewinnen. Hierfür wird aktuell an der Verbesserung und Optimierung des Bestärkenden Lernens sowie an weiteren Inhalten für die Interaktion selbst gearbeitet.

## 5. Danksagung

Die hier beschriebene Arbeit wurde teilweise unterstützt durch das STMWFK im Rahmen des Forschungsverbundes ForGenderCare sowie durch das Horizon2020 Research and Innovation Programme der Europäischen Union (Grant Agreement No 645378) im Rahmen des Aria-Valuspa Projekts.

- [1] Breazeal, C.: Designing sociable robots. MIT press, 2004.
- [2] Aly, A. und Tapus, A.: Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction. *Autonomous Robots*, 40(2), 2015, S. 193–209
- [3] Joosse, M., Lohse, M., Perez, J. G. und Evers, V.: What you do is who you are: The role of task context in perceived social robot personality. 2013 IEEE International Conference on Robotics and Automation, 2013, S. 2134–2139
- [4] Tapus, A., Mataric M. und Scassellati, B.: Socially assistive robotics [Grand Challenges of Robotics]. *IEEE Robotics & Automation Magazine*, 14(1), 2007, S. 35–42
- [5] Ritschel, H. und André, E.: Real-Time Robot Personality Adaptation Based on Reinforcement Learning and Social Signals. *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 2017, S. 265–266
- [6] Ritschel, H., Baur, T. und André, E.: Adapting a Robot’s Linguistic Style Based on Socially-Aware Reinforcement Learning. *Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 2017, S. 378–384
- [7] McCrae, R. R. und Costa, P. T.: The five-factor theory of personality. *Handbook of personality: Theory and research*, 3, 2008, S. 159–181
- [8] Wagner, J., Lingenfeller, F., Baur, T., Damian, I., Kistler, F. und André, E.: The social signal interpretation (SSI) framework. *Proceedings of the 21st ACM international conference on Multimedia - MM ’13*, 2013, S. 831–834

- [9] Baur T., Schiller, D. und André, E.: Modeling User's Social Attitude in a Conversational System, Human-Computer Interaction Series. Springer International Publishing, 2016, S. 181-199
- [10] Mairesse, F. und Walker, M. A.: Controlling user perceptions of linguistic style: Trainable generation of personality traits. Computational Linguistics, 37(3), 2011, S. 455-488