

Simulierte Bewerbungsgespräche mit dem virtuellen Agenten

Realitätsnahe Trainings zum Umgang mit eigenen Emotionen und zur Performance-Steigerung



PATRICK GEBHARD
Dr., Deutsches Forschungs-
zentrum für Künstliche
Intelligenz (DFKI),
Saarbrücken

Sich in Bewerbungsgesprächen erfolgreich zu präsentieren, ist vor allem für unerfahrene Bewerber/-innen eine schwierige Situation. Der Umgang mit eigenen Emotionen ist dabei ein zentraler Aspekt. Um dies zu trainieren, erforschen das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) und die Universität Augsburg seit 2011 den Einsatz von sozialen Agenten in simulierten Bewerbungsgesprächen. Eine Herausforderung ist dabei, das Verhalten dieser Agenten möglichst realitätsnah zu gestalten und kommunizierte Emotionen durch Computersysteme zu deuten. Der Beitrag gibt Einblicke in diese Forschungsarbeiten, beleuchtet Potenziale virtueller Bewerbungstrainings und benennt abschließend Perspektiven für die weitere Forschung und Entwicklung.

Herausforderung Bewerbungsgespräch

Bewerbungsgespräche dienen dazu, den positiven Eindruck aus den Bewerbungsunterlagen in einem persönlichen Gespräch zu bekräftigen. Dabei spielen neben fachlichen Kriterien auch personale Eigenschaften der Bewerber/-innen eine Rolle. Die Scham, im Bewerbungsgespräch etwas nicht zu wissen oder womöglich eine falsche Antwort zu geben, kann gerade bei unerfahrenen Bewerberinnen und Bewerbern Unsicherheit auslösen und im schlimmsten Fall zu einem Blackout führen. Ein derartiges Erlebnis kann Angst vor zukünftigen Bewerbungsgesprächen auslösen. Wie kann ein Training aussehen, das es Bewerberinnen und Bewerbern ermöglicht, die Fähigkeiten zu entwickeln, eine Bewerbungssituation souverän zu meistern?

Für DANIELS/ROBINSON (2019) ist der Umgang mit der Emotion Scham bei Bewerbungen von zentraler Bedeutung. Als interne Emotion tritt sie vor allem dann auf, wenn eigene Kompetenzen oder Eigenschaften angezweifelt werden (vgl. HILGERS 1996). Scham kann meistens nicht direkt im Verhalten erkannt werden, da sie in der Regel unmittelbar (unbewusst) reguliert wird. Bei der Emotionsregulation (vgl. TAMIR 2011) wird eine schwer ertragbare Emotion (zum Beispiel Scham) durch eine bewusst erlebte »leichtere« Emotion (zum Beispiel Ärger oder sogar Freude) reguliert. Dadurch wird eine unangenehme Situation emotional »aushaltbar«. Doch wie können solche komplexen

Prozesse im Rahmen eines virtuellen Bewerbungstrainings von einem Computersystem erkannt werden und der/dem Trainee bewusst gemacht werden?

Affective Computing – Übertragung von sozio-emotionalen Fähigkeiten in Computersysteme

1963 identifizierte der Begründer der Kognitiven Psychologie ULRIC NEISSER die wohl wichtigste Herausforderung des Forschungsgebiets der künstlichen Intelligenz: die Nachbildung beziehungsweise Modellierung der menschlichen Motivation (vgl. NEISSER 1963). Sie macht uns Menschen einzigartig und treibt uns immer wieder an. Persönliche Wünsche und das individuelle Erleben von Situationen sind eng damit verbunden. 1997 benennt die Computerwissenschaftlerin und Elektroingenieurin ROSALIND PICARD dieses Teilgebiet der KI als Affective Computing (vgl. PICARD 1997). Dort wird erforscht, wie mithilfe von Modellen über Emotionen und soziales Verhalten Computersysteme so verbessert werden können, dass mit ihnen natürlich interagiert werden kann. Sie beschreibt unter anderem, wie derartige Systeme zum Training von Bewerbungsgesprächen eingesetzt werden können (ebd., S. 85 ff.). Damals beschränkte sich die Hilfe durch ein solches System auf beobachtbare und messbare Signale und ob diese in einer Situation angemessen sind. Nervöses Verhalten galt demnach zum Beispiel als unangemessen.

Um das Training nahe an das echte Erleben zu bringen, bedarf es des Einsatzes sozialer Agenten. Soziale Agenten sind (teil-)autonome Computerprogramme, die bestimmte Software- und Hardwareeigenschaften aufweisen (vgl. VINAYAGAMOORTHY u.a. 2006) und menschenähnliche Gesprächspartner simulieren (vgl. Foto). Sie können eine virtuelle oder physische Präsenz haben. Visuell reicht das Spektrum von einer »cartoonhaften« bis hin zu einer detailreich nachgeahmten Darstellung eines Menschen.

Zur Sprachausgabe nutzen sie aktuelle Text-to-Speech-Systeme. Zur Erfassung ihrer Gesprächspartner und deren Umgebung in Echtzeit (mind. 25-mal in der Sekunde) benötigen sie Hardware-Sensoren wie Kameras oder Mikrofone. Die Verarbeitung der Sensordaten geschieht auf unterschiedlichen Ebenen und ist an ein sozio-emotionales Benutzer- und ein Agentenverhaltensmodell gekoppelt. Damit können soziale Agenten ihr eigenes Verhalten an das der Interaktionspartnerin/des -partners anpassen. Eine empathische Adaption und eine facettenreiche Analyse der Interaktion erfordert eine Echtzeitsimulation von Benutzeremotionen (im Sinne des individuellen Erlebens einer Situation). Dank dieser technischen Ausstattung kann man mit ihnen wie gewohnt interagieren, also mit Gestik, Mimik und natürlicher Sprache.

Die Affective-Computing-Gruppe des DFKI in Saarbrücken und der Lehrstuhl für Multimodale Mensch-Technik-Interaktion der Uni Augsburg forschen seit 2007 gemeinsam am Einsatz von sozialen Agenten zu Trainings- und Assistenzzwecken. Bewerbungstrainings sind ein Fokus, der mit dem EU-Projekt TARDIS begann, mit dem BMBF-Projekt EmpaT weitergeführt wurde und jetzt im DFG-Grundlagenforschungsprojekt DEEP vertieft wird (vgl. Infokasten).

Projekte zum Einsatz von sozialen Agenten zu Trainings- und Assistenzzwecken

EU-Projekt TARDIS (2011–2014)

Erforschung des Einsatzes von sozialen Agenten für das Bewerbungstraining von jungen Erwachsenen, die keine Schule besuchen, keiner Arbeit nachgehen und sich nicht in beruflicher Ausbildung befinden. Der Fokus liegt auf der Erkennung des beobachtbaren Benutzerhaltens (Gesicht und Körper) und der Nutzung für ein adaptives Training von Bewerbungsgesprächen.

BMBF-Projekt EmpaT (2015–2018)

Erforschung von Computermodellen über interne Benutzeremotionen im Kontext von Bewerbungstrainings. Die Modelle benötigen eine Kopplung mit einer Echtzeitanalyse für soziale Signale. Diese erweiterten sozialen Agenten sollen ein empathisches Trainieren von Bewerbungsgesprächen mit tiefergehenden Analysen ermöglichen.

DFG-Projekt DEEP (2018–2021)

Grundlagenforschung der Repräsentation interner Emotionsregulation mit Simulation von internem Erleben. Kopplung mit einer erweiterten Analyse sozialer Signale unter Berücksichtigung aller Gesprächspartner. Erstellung eines erklärbaren, hybriden kognitiv- und machine-learning-basierten Computeremotionsmodells. Evaluation der Anwendbarkeit für empathische Assistenz- und Trainingsagenten.



Interaktion mit einem sozialen Agenten (Foto: DFKI 2019)

Was macht das Training mit sozialen EmpaT-Agenten so besonders?

Die im Projekt EmpaT entwickelten Agenten (vgl. Foto) simulieren in natürlicher Weise typische Verhaltensweisen menschlicher Bewerbungstrainer/-innen: Trainees gehen alleine in das Büro, in dem das Interview stattfindet. Darin sitzt ein Agent – lebensgroß dargestellt auf einem 42 Zoll großen Bildschirm – an einem virtuellen Schreibtisch, der in dem Büro durch einen Tisch fortgesetzt wird. Nachdem der oder die Trainee eingetreten ist, beginnt der Agent das Gespräch: »Schön, dass Sie den Weg zu uns gefunden haben, Frau Sonnental*, nehmen Sie doch Platz!« und zeigt auf den gegenüberliegenden Stuhl. Im Anschluss stellt sich der Agent vor und das Interview beginnt – wie bei einem echten Bewerbungsgespräch.

Das Verhaltensmodell der EmpaT-Agenten ist erstmalig direkt an die Interpretation der Kommunikationssignale des menschlichen Dialogpartners gekoppelt; das heißt, der Agent ist in der Lage, zeitlich-dynamisch soziales Kommunikationsverhalten, wie das unbewusste Nachahmen (Mimikry) von Haltung, Gesten und Gesichtsausdrücken, zu zeigen oder eben nicht. Technisch erlaubt dieser Ansatz auch das gegenseitige Unterbrechen von Redebeiträgen, das heißt, man kann dem Agenten beim Sprechen ins Wort fallen und umgekehrt. Auf welche Art und in welcher Geschwindigkeit er darauf reagiert, kann frei modelliert werden.

Soziale Agenten sind flexibel, was die Konfiguration ihres nonverbalen Verhaltens angeht. Es kann je nach Situation beispielsweise zwischen unterstützend und dominant und den damit verbundenen inhaltlichen Herausforderungen angepasst werden (vgl. GEBHARD u.a. 2014). Auf diese Weise können verschiedene sozio-emotionale Interaktionsmuster simuliert werden (vgl. BÄNNINGER-HUBER 1996).

* Der Name wird dem Agenten im Vorfeld mitgeteilt.

Tabelle
Typische Scham-Regulationsstrategien

Strategie	Bedeutung (zusammenfassender Auszug)	Auswahl typischer Verhaltensmerkmale
Rückzug	Jegliche Kommunikationssignale des Gegenübers werden ignoriert.	Ruhige und eingefrorene Bewegungen, abgewandter Blick
Angriff	Die Beziehung zum Dialogpartner wird als nicht gewinnbringend eingestuft.	Aggressives Verhalten, gerichteter Blick und Gesten
Selbstangriff	Selbstkritik nimmt dem Gegenüber die Macht, persönliche Eigenschaften anzugreifen.	Ausdrücke des Ekels, abgewandter Blick
Vermeidung	Verleugnung der Ursache, nach dem Motto »ich überzeuge mich, also sind andere auch überzeugt«.	Ausdrücke der Freude oder Angst, teilweise abgewandter Blick

Quelle: NATHANSON (1992), eigene Darstellung

EmpaT-Agenten sind zudem in der Lage, Signale situationsabhängig zu bewerten und dabei mögliche interne Emotionen (zum Beispiel Scham) und deren Regulation zu berücksichtigen.

Die tiefere Deutung menschlicher Kommunikationssignale ist technisch durch das sozio-emotionale Benutzermodell MARSSI realisiert (vgl. GEBHARD u.a. 2018a). Es modelliert vier typische Scham-Regulationsstrategien (vgl. NATHANSON 1992) und erkennt entsprechende Signale mittels eigens dafür geschaffener Signal-Klassifikatoren. Dabei sind neben dem Kontext Kopf- und Blickrichtung wichtige Indikatoren (vgl. Tab.).

Diese kontextabhängigen Signal-Klassifikatoren sind mittels dynamischer Bayes'scher Netze (DBN) realisiert. Sie beschreiben Verknüpfungen unterschiedlicher Merkmale und zeitliche Änderungen der Verknüpfungen zwischen Merkmalen. DBN können somit die Wahrscheinlichkeit für das Auftreten eines bestimmten Merkmals (hier: Scham-Regulation) in Abhängigkeit zu anderen (hier: nonverbale Signale und Kontextinformationen) vorher-sagen. Die Abhängigkeiten wurden mittels maschineller Lernmethoden ermittelt. Als Datenpool dienten 100 Videoaufnahmen von schamauslösenden Situationen in Bewerbungsgesprächen mit detaillierter Annotation von Kontext und Kommunikationssignalen.

Die EmpaT-Agenten nutzen MARSSI, um mögliche Scham-Regulationsstrategien während des Bewerbungsgesprächs zu erkennen und diese Stellen für eine anschließende Besprechung zu markieren. Das folgende Beispiel soll dies veranschaulichen: Der Agent stellt die provokative Frage: »Sagen Sie mal, woher haben Sie denn Ihr Outfit, irgendwie passt das gar nicht zu Ihnen!«. Das Gespräch wird per Video aufgezeichnet und zeitgleich mit einer Abfolge von erkannten Signalen (Kopf-, Blickbewegung, Gestik, Haltung ...) angereichert. Weiterhin können EmpaT-Agen-

ten MARSSI zur Adaption des Schwierigkeitsgrads nutzen. Wenn öfter schwierige Regulationsstrategien (zum Beispiel Angriff oder Selbstangriff) im Traineeverhalten erkannt werden, passt sich der soziale Agent an und stellt anstatt provokativer eher einfache Fragen.

Das aufgezeichnete Bewerbungstraining wird für eine anschließende Besprechung genutzt. Es enthält zusätzliche Informationen über mögliche Emotionsregulationen und Verhaltensbeschreibungen. Beides kann zur Selbsteinschätzung und -reflexion verwendet werden; alleine oder zusammen mit einem (virtuellen) Coach.

Virtuelle Bewerbungsgespräche können frei gestaltet werden. Unterschiedliches Verhalten von virtuellen Mitarbeiterinnen und Mitarbeitern, Umgebungsgeräusche und Wandfarben simulieren verschiedene Arbeitsatmosphären. Auch simulierte Störungen sind möglich (vgl. GEBHARD u.a. 2018b).

Erste Ergebnisse

Viele der über 200 Personen, die seit 2011 an einer simulierten Bewerbung mit sozialen Agenten teilgenommen haben, beschreiben dies als eine beeindruckende Erfahrung. Sie merkten an, dass gerade die natürliche Interaktion mit den sozialen Agenten intensiv und herausfordernd sei, weil man das Gefühl habe, mit einem aufmerksamen Gesprächspartner zu interagieren.

Bezüglich des messbaren Nutzens von simulierten Bewerbungen mit sozialen Agenten wurden zwei Hypothesen formuliert. Das EmpaT-Training reduziert erstens die Angst vor Bewerbungsgesprächen und steigert zweitens die Interviewperformance.

In einer Vergleichsstudie mit insgesamt 70 Teilnehmenden wurde das EmpaT-Training mit klassischen Vorbereitungs-methoden für Bewerbungen verglichen. Die Kontrollgrup-

pe erhielt Lehrmaterial, Bilder und Videos zum Thema »Körpersprache« und »Vorstellungsgesprächsfragen« für das Selbststudium und zum Durchführen von Rollenspielen, die Angst reduzieren, Selbstwirksamkeit erhöhen und nonverbales Verhalten verbessern. Beide Gruppen hatten etwa 20 Minuten Zeit, sich auf das anschließende Bewerbungsgespräch mit einem professionellen Interviewer vorzubereiten. Das Interview wurde per Video aufgezeichnet. Expertinnen und Experten sahen die Aufzeichnungen und beurteilten die Interviewperformance der Teilnehmenden. Die so eingeschätzte Interviewperformance war bei der Gruppe, die das virtuelle Jobinterview-Training erhalten hatte, signifikant höher als bei der Kontrollgruppe. Mit Fragebögen wurde zudem die subjektive Angstreduktion bei den Teilnehmenden ermittelt. Die Ergebnisse zeigen, dass das virtuelle Bewerbungstraining die selbst eingeschätzte Angst signifikant (um ca. 10%) reduzieren konnte.

Die simulierte Bewerbung mit virtuellen Jobinterviewern kann dazu dienen, Bewerbungsgespräche effektiv vorzubereiten. Die erwähnten Ergebnisse beziehen sich auf die Studie von LANGER u. a. (2016), die die Ergebnisse der Studie mit 20 Teilnehmenden von DAMIAN u. a. (2015) bestätigen.

Reality-Check und Perspektiven

Die im Projekt EmpaT entwickelten virtuellen sozialen Agenten ermöglichen simulierte Bewerbungsgespräche, die echten Situationen sehr nahekommen. Die umfassen-

dere automatische Analyse menschlicher Kommunikationssignale im Kontext der höchst relevanten Emotion Scham erlaubt eine tiefer gehende Deutung des Bewerberverhaltens in simulierten Gesprächen und ein empathisches Trainieren.

Zurzeit simuliert das Benutzeremotionsmodell basierend auf vier Scham-Regulationsstrategien unterschiedliche Facetten, wie es in einem Menschen in einer Situation emotional gerade aussehen könnte. Ob dies jedoch den tatsächlichen emotionalen Zustand repräsentiert, weiß nur der/die Trainee selbst.

Dazu visiert die Forschergruppe »Affective Computing« des Lehrstuhls für Multimodale Mensch-Technik-Interaktion die nächste Generation sozialer Agenten an. Diese sollen in der Lage sein, ihre Annahmen mit dem/der Trainee zu besprechen und zu lernen, wie dieser Mensch die Situation emotional erlebt hat. Die Rückmeldungen sollen dazu genutzt werden, dass der soziale Agent vertieft lernt, indem er die Emotionsmodellierung und das Bayes'sche Netz zur Erkennung selbstständig anpasst, um genauere adaptive Rückmeldungen geben zu können.

Soziale Agenten benötigen aufgrund der Echtzeitanalyse sozialer Signale eine hohe Rechenleistung und verbesserte Sensortechnologie. Beides bringt zum Beispiel das Apple iPhone ab Version X mit. Sobald diese Technologie preislich für einen großen Markt attraktiv ist, kann mit ersten empathischen Assistenz- und Trainingsanwendungen mit sozialen Agenten auch außerhalb von »Laborsituationen« gerechnet werden. ◀

Literatur

BÄNNINGER-HUBER, E.: *Mimik – Übertragung – Interaktion: Die Untersuchung affektiver Prozesse in der Psychotherapie*. Bern 1996

DAMIAN, I. u. a.: *Games are Better Than Books: In-situ Comparison of an Interactive Job Interview Game with Conventional Training*. In: CONATI, C. (Hrsg.): *Artificial Intelligence in Education*. Cham 2015, S. 84–94

DANIELS, M. A.; ROBINSON, S. L.: *The Shame of It All: A Review of Shame in Organizational Life*. In: *Journal of Management* January 2019 – URL: <https://doi.org/10.1177/0149206318817604> (Stand: 10.04.2019)

GEBHARD, P. u. a.: *Exploring Interaction Strategies for Virtual Characters to Induce Stress in Simulated Job Interviews*. In *Proceedings of the 13th International Conference on Autonomous Agents and Multi-Agent Systems*. Paris 2014, S. 661–668

GEBHARD, P. u. a.: *MARSSI: Model of Appraisal, Regulation, and Social Signal Interpretation*. In: *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. Stockholm 2018 a, S. 497–506

GEBHARD, P. u. a.: *Serious Games for Training Social Skills in Job Interviews*. In: *IEEE Transactions on Games* February 2018 (2018b) – URL: <https://ieeexplore.ieee.org/document/8299545> (Stand: 10.04.2019)

HILGERS, M.: *Scham – Gesichter eines Affekts*. Göttingen 1996

LANGER, M. u. a.: *Dear Computer, Teach Me Manners: Testing Virtual Employment Interview Training*. In: *International Journal of Selection and Assessment* 24 (2016) 4, S. 312–323

NATHANSON, D. L.: *Shame and Pride: Affect, Sex, and the Birth of the Self*. New York 1992

NEISSER, U.: *The Imitation of Man by Machine*. In: *Science* 139 (1963) 3551, S. 193–197

PICARD, R. W.: *Affective Computing*. Cambridge 1997

TAMIR, M.: *The Maturing Field of Emotion Regulation*. In: *Emotion Review* 3 (2011) 1, S. 3–7

VINAYAGAMOORTHY, V. u. a.: *Building Expression into Virtual Characters*. In: *Eurographics Conference State of the Art Report*. Wien 2006