

Human-Agent Teaming

In elektronischen Geräten und Systemen arbeiten zunehmend Algorithmen, die ähnlich dem Menschen Informationen verarbeiten, daraus Entscheidungen ableiten und Aktionen veranlassen. Wenn diese sogenannten softwarebasierten oder künstlichen „Agenten“ uns bei der Erfüllung einer Aufgabe unterstützen, spricht man von Human-Agent Teaming (HAT). Nicht zuletzt von der Weiterentwicklung diesbezüglicher Technologien zur Ermöglichung echter Teamarbeit zwischen Mensch und Maschine wird es abhängen, wie erfolgreich Systeme mit Künstlicher Intelligenz in Zukunft eingesetzt werden können.

Die Interaktion mit Künstlicher Intelligenz bedingt neue Konzepte für die Schnittstelle zum Menschen, damit dieser trotz zunehmender Delegation wichtiger Aufgaben an die Technik eine wirkliche Kontrolle behält. Neue Konzepte betreffen dabei zum einen den physischen Aspekt der ergonomischen Gestaltung der Schnittstelle. Hier spielt z. B. die Sprachsteuerung eine immer wichtigere Rolle. Beim Human-Agent Teaming geht es darüber hinaus in erster Hinsicht um die Frage, wie die Zusammenarbeit zwischen Menschen und künstlichen Agenten organisiert und ausgestaltet sein muss, um eine Leistungsfähigkeit zu ermöglichen, die der eines Teams nur aus Menschen gleicht. HAT würde damit zukünftig die Möglichkeit eröffnen, mit weniger Personal als heute anspruchsvolle und kritische Aufgaben auszuführen, zum Beispiel die Überwachung weitverzweigter Großanlagen wie Kraftwerke oder andere komplexe Infrastruktur, die Organisation von Logistikprozessen, die Auswertung von Satellitenbildern und anderes mehr. Im militärischen Bereich wären die Entscheidungsunterstützung in Leit- und Gefechtsständen und das Führen einer großen Anzahl unbemannter Systeme wichtige Einsatzgebiete. Vor dem Hintergrund des demografisch absehbaren Mangels an Fachkräften, der wachsenden Menge und Bedeutung digitaler Daten sowie der Wichtigkeit kurzer Reaktionszeiten wird HAT daher eine wichtige Funktion zukommen. Bevor es so weit ist, sind aber noch zahlreiche Forschungsfragen zu behandeln.

Dazu gehören technische Aspekte der Kooperation der künstlichen Agenten untereinander wie Architekturen oder Schnittstellen zum Datenaustausch. Zum Beispiel gibt es spezifische Sprachmodelle für Agentensysteme, die den Inhalt eines Satzes ähnlich einem Computercode in Sprachobjekte und Sprechakte gliedern. Damit kann einer Nachricht entnommen werden, ob sie z. B. eine Frage, eine Anweisung oder eine Feststellung enthält. Das ist die Voraussetzung, um dem Inhalt eines Dialogs Intentionen oder Erwartungen des Gegenübers zu entnehmen. Die Prozessketten der Agenten werden in Form sogenannter kognitiver Architekturen organisiert, die sich am Aufbau des menschlichen Gehirns orientieren. Sie besitzen verschiedene Speicherbereiche für Kurz- oder Langzeitwissen, Wissen über Objekte oder Zusammenhänge (semantisches Wissen), Wissen zu Methoden und Abläufen (prozedurales Wissen) und Wissen zu zeitlichen Verläufen (episodisches Wissen). Ein Agent kann also z. B. einen Gegenstand wiedererkennen, kennt dessen Funktion und weiß, wann und warum sein Gebrauch zuletzt scheiterte. Er lernt dazu, indem er die Inhalte verschiedener Langzeitspeicher auf Basis seiner aktuellen Erfahrungen anpasst, etwa durch Trainingssequenzen in Lernverfahren der Künstlichen Intelligenz. Gegenüber ihren Team-Mitgliedern verhalten Agenten sich entsprechend Grundsätzen wie Aufrichtigkeit (d. h. sie äußern sich wahrheitsgemäß und gehen keine unerfüllbaren Verpflichtungen ein), Autonomie (d. h. ein Agent darf von einem anderen nicht zur Dienstleistung gezwungen werden) und Verpflichtung (d. h. ein Agent muss einen Dienst erbringen, zu dem er sich verpflichtet hat).

Da sie mit Menschen zusammenarbeiten sollen, müssen die Agenten darüber hinaus auch die in einem Team aus Menschen relevanten Faktoren berücksichtigen, also physiologische und psychische Aspekte der Teamleistung wie Sympathie, Vertrauen, Stress oder Erschöpfung. Die Herausforderung für die HAT-Forschung liegt darin, objektivierbare Ergebnisse zu erhalten, wie Menschen und virtuelle Agenten bestmöglich zusammenwirken können, wie dabei mit

benutzerindividuellen Einflussfaktoren umzugehen ist und wie Lernen und Training in einem HAT erreicht werden können. Von zentraler Bedeutung sind dabei das Vertrauen des Nutzers in künstliche Agenten, die Nachvollziehbarkeit (auch Transparenz genannt) von Entscheidungen des Agenten und ein teamweit übereinstimmendes Situationsbewusstsein.

Ein Nutzer könnte zum Beispiel dazu tendieren, dem Agenten zu sehr zu vertrauen und deshalb seine Aufsichtsfunktion vernachlässigen. Umgekehrt könnte er bei fehlendem Vertrauen aber auch zu häufig in Situationen eingreifen, in denen der Agent effektiver als der Mensch arbeitet. Offenbar spielen u. a. menschenähnliche Eigenschaften der Agenten eine Rolle dafür, wie sehr ein Mensch ihnen vertraut und wie schnell sie verloren gegangenes Vertrauen wiedergewinnen können. Daher wird z. B. daran gearbeitet, die künstliche Sprachsynthese von Agenten im Hinblick auf Intonation, regionale Dialekte oder Füllwörter und Pausen so zu verbessern, dass der Nutzer keinen Unterschied mehr zu einem Dialog mit einem wirklichen Menschen wahrnimmt. Das ist wichtig, weil mit steigender Autonomie von Systemen auch die Komplexität von Entscheidungsprozessen der Agenten wächst, sodass die Erläuterungsbedürftigkeit gegenüber einem menschlichen Systemoperator steigt. Ein passendes Maß an Transparenz verbessert das gemeinsame Lagebewusstsein und hilft dem Operator, eine möglichst realistische Erwartung an die Leistungsfähigkeit des Agenten zu entwickeln. Dies wiederum gilt als maßgeblicher Faktor für den Aufbau eines austarierten Vertrauensverhältnisses im Mensch-Agenten-Team.

Künstliche Agenten werden also in der Zukunft unverzichtbar sein, damit der Mensch eine wirkungsvolle Kontrolle über zunehmend automatisierte Systeme ausüben kann. Die Zusammenarbeit zwischen Agenten und Menschen muss deshalb so effektiv wie möglich gestaltet werden. Auf dem Weg dorthin werden Forschungseinrichtungen, KI-Firmen, Systemhersteller und Testnutzer noch intensiv zusammenarbeiten müssen.

Dr.-Ing. Guido Huppertz