

# **Anthropomorphe Agenten**

Seminar  
“Multimodale Mensch-Maschine Kommunikation”  
Lutz Krause  
25.01.01

## **Einführung**

Anthropomorphe Agenten sind Schnittstellen zum Anwender, die die menschliche Kommunikation als Vorbild haben. Bei diesem Thema fließen viele Bereiche der MMMMK mit ein.

Als Grundlage dienen Erforschungen der menschlichen Kommunikation von Psychologen und Verhaltensforschern.

Die Idee eines Anthropomorphen Agenten existiert schon recht lange, wird aber erst seit kurzem ernst genommen. Es ist somit ein recht junges Gebiet der Mensch-Maschine-Kommunikation, über das sehr viel diskutiert wird.

## **Motivation**

Menschen sind soziale Wesen, für die die Kommunikation eine sehr wichtige Rolle spielt.

Die menschliche Kommunikation ist sehr komplex, eine einfache Schnittstelle ist also nicht an die Bedürfnisse des Menschen angepasst und kann ihre Möglichkeiten nicht ausschöpfen. Desweiteren personifizieren Menschen Computer und behandeln sie oft nach sozialen Regeln.

Somit ist es naheliegend ein Interface zu entwerfen, dass sich der menschlichen Konversation anpasst.

## **Anforderungen**

Anthropomorphe Agenten sollten nicht nur lebendig wirken, sondern sollten sich in einer Konversation genau wie Menschen mit folgenden Eigenschaften verhalten:

- nonverbale Eingabe erkennen und darauf adäquat reagieren
- Möglichkeit verbalen und nonverbalen Output zu generieren
- Sie müssen mit Sprach-Funktionen umgehen können
- Sie müssen eigene Vorschläge beisteuern können

Kann man dies nicht adäquat leisten, sollte man auf einen A. Agenten verzichten.

## Schwierigkeiten

Die Menschliche Konversation ist sehr komplex, es müssen Gesten, Betonung, Sprache, Mimik und der Augen-Ausdruck berücksichtigt werden.

Es mussten/müssen erst noch schwierige Voraussetzungen in der Sprachverarbeitung, der Bildverarbeitung und in der Analyse der menschl. Konversation geschaffen werden.

## Gandalf



Gandalf wurde am MIT entwickelt (ca. 1996) und ist ein Prototyp für einen humanoiden Agenten.

Er besteht aus einem Gesicht, einer Hand und einer Stimme und ist ein Experte für unser Sonnensystem. Die Anwender nimmt er mit Hilfe eines Eye-Trackers, eines Mikrophons und einer Data-Suit wahr.

Gandalf führt Echtzeit Dialoge mit Menschen und ist in der Art und Weise wie er spricht, betont und gestikuliert dem Menschen sehr ähnlich.

Seine Antworten kommen jedoch nur aus einer festen Bibliothek von Antworten und der Spracherkennung arbeitet mit einem Wörterbuch mit ca. 100 Wörtern.

## REA



Auch REA wurde am MIT entwickelt, ihr Arbeitsbereich ist der Verkauf von Immobilien.

Sie hat einen menschl. Körper und versucht Augensprache, Körperhaltung, Hand-Gesten und den Gesichtsausdruck wie Menschen zu benutzen.

REA wird auf einer Projektionswand dargestellt, auf der zwei bewegliche Kameras befestigt sind, mit der sie die Hände und das Gesicht des Benutzers passiv wahrnehmen kann.

Im Gegensatz zu Gandalf generiert sie ihre Sprach- und Gesten-Ausgabe basie-

rend auf einem Wörterbuch und grammatikalischen Regeln sowie Wissen über die Funktion von Gesten.

## Beispielkonversation mit REA

*Mike approaches the projection screen. As Mike moves within range of the two cameras mounted above the screen, Rea turns to face him and says:*

REA> Hello. How can I help you?

MIKE> I'm looking to buy a place near MIT. *Rea nods, indicating that she is following.*

REA> I have a house to show you. *A picture of a house appears on-screen behind Rea who blinks and looks at the house and then at Mike.*

REA> It's in Somerville.

MIKE> Tell me about it. *Rea looks up and away while she plans what to say*

REA> It s big. *Rea makes an expansive gesture with her hands. Mike brings his hands up as if he is about to speak and so Rea does not continue; instead waiting for him to speak.*

MIKE> Tell me more about it ...

## YMIR, die Architektur hinter Gandalf und REA

YMIR ist eine generelle Architektur, mit der man die Fähigkeit menschl. zu Kommunizieren mit multimodalem In- und Output implementieren kann.

Sie ist modular, verteilt und ist sehr flexibel, gestattet dem Anwender somit zu experimentieren.

Mit YMIR können so viele Ein- und Ausgabemodalitäten implementiert werden, wie es Ein- und Ausgabe-Geräte gibt. Seine Kern-Module arbeiten auf Funktionen, also nur unimodal, die Ein- und Ausgabe ist multimodal. Die verschiedenen Eingabemodalitäten werden in Funktionen umgewandelt, b.z.w. Funktionen in Verhalten. D.h. die verschiedenen Modalitäten haben intern die gleiche Repräsentation und werden gleich behandelt.

## Nutzen von A. Agenten:

- Der Anwender muss nicht lernen den Computer zu bedienen
- Hemmschwelle einen Computer zu bedienen kann abgebaut werden
- Menschen reagieren emotional auf Gesichter und übertragen menschl. Eigenschaften auf Agenten
- Motivation kann gesteigert werden
- A. Agenten erscheinen Anwendern oft glaubwürdiger

- Überprüfung von Konzepten der KI
- A. Agenten bieten Psychologen und Sprachforschern eine gute Plattform für die Erforschung der menschlichen Konversation

## Nachteile

A. Agenten können zu einer Verlangsamung der Arbeit führen. Desweiteren können menschl. Züge eines Agenten zu einer Übererwartung des Anwenders gegenüber dem Agenten bewirken.

## Einsatzgebiete

A. Agenten können besonders gut dann eingesetzt werden, wenn Computer ohne Tastatur eingesetzt werden oder wenn Benutzer ohne Computerkenntnisse Computern bedienen müssen. Z.B. können sie in Spielzeug, Fahrkartenautomaten, in öffentl. Gebäuden wie Museen, Schulen, Universitäten e.c.t. eingesetzt werden.

## Zusammenfassung

A. Agenten können ein sehr wichtiges und leistungsstarkes Interface werden, aber bis dahin muss die menschl. Kommunikation noch weiter erforscht werden, und die Techniken der Sprach- und Gestenerkennung und Auswertung müssen noch verbessert werden.

Leider wurden in diesem Bereich bisher nur einzelne, wenige Projekte gestartet, die nicht aufeinander aufbauten.

In diesem Bereich kann noch viel erforscht und verbessert werden, z.B. kan die Leistung von A. Agenten durch vermehrten Einsatz von Neuronalen Netzen stark gesteigert werden.

Ob es jedoch jemals Agenten mit menschlichen Eigenschaften wie das Empfinden von Emotionen geben wird ist fraglich.

Können jedoch die grundlegenden Fertigkeiten der menschl. Konversation nicht adäquat implementiert werden, sollte man auf den Einsatz von a. Agenten verzichten.

## Literatur

Thorisson, K.: Communicative Humanoids - A Computational Model of Psychosocial Dialogue Skills. Ph.D. Thesis, Massachusetts Institute of Technology, 1996. y

Cassell, J.: More than Just Another Pretty Face: Embodied Conversational Interface Agents. Communications of the ACM 43(4), 2000, 70-78.