

Anthropomorphe Agenten

Seminar
“Multimodale Mensch-Maschine Kommunikation”

Lutz Krause

25.01.01

Inhalt:

- Einführung und Motivation
- Gandalf/REA/Grandchair: drei Beispiele eines anthropomorphen Agenten
- YMIR: die Architektur hinter Gandalf und REA
- Vor-/Nachteile und Einsatzgebiete
- Zusammenfassung

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.

Motivation

- Menschen sind soziale Wesen, für die die Kommunikation eine sehr wichtige Rolle spielt
 - Menschen lernen schon früh miteinander zu kommunizieren
 - Der menschliche Körper hat sich gut für die Kommunikation entwickelt
- Die menschliche Kommunikation ist sehr komplex, eine einfache Schnittstelle ist also nicht an die Bedürfnisse des Menschen angepasst und kann ihre Möglichkeiten nicht ausschöpfen
- Menschen personifizieren Computer und behandeln sie oft nach sozialen Regeln

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

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 geschaffen werden in der
 - Sprachverarbeitung
 - der Bildverarbeitung
 - der Analyse der menschl. Konversation

Gandalf



- Gandalf ist ein Prototyp für einen humanoiden Agenten.
- Er wurde von Kristinn Thorisson am MIT entwickelt (ca. 1996)

Gandalf



- Gandalf ist ein Experte für unser Sonnensystem
- Anwender müssen Eye-Tracker, Mikrofon und Data-Suit tragen
- Gandalfs Antworten kommen nur aus einer festen Bibliothek von Antworten
- Der Spracherkenner arbeitet mit einem Wörterbuch mit ca. 100 Wörtern

Gandalf



- In der Zeit, in der Gandalf geplant wurde, hatten die Erschaffer von Agenten mit Körpern die Meinung, dass menschlich aussehende Agenten nicht glaubwürdig und lebensnah sind. Deshalb erschufen sie synthetische Charakter, die an Disney-Figuren erinnern.
- Diese erschienen den Anwendern glaubhafter und realer.

REA (Real Estate Agent)



- REA wurde von Justine Cassell am MIT entwickelt
- Ihr Arbeitsbereich ist der Verkauf von Immobilien
- Sie nimmt den Anwender passiv mit Hilfe von Kameras wahr

REA

- REA hat einen menschl. Körper und versucht Augensprache, Körperhaltung, Hand-Gesten und den Gesichtsausdruck wie Menschen zu benutzen
- Sie wird auf einer Projektionswand dargestellt, auf der zwei bewegliche Kameras befestigt sind, die die Hände und das Gesicht des Benutzers verfolgen
- Ihre Sprachausgabe und Gesten werden in Echtzeit erzeugt
- Die Sprach- und Gesten-Ausgabe basierend auf einem Wörterbuch und grammatikalischen Regeln sowie Wissen über die Funktion von Gesten

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 ...

Verbesserungsmöglichkeiten und Einsatzfähigkeit

- Abstimmung von Sprache/Gestik, Klassifizierung von Gesten, Smalltalk, Emotionen, Lernen von Wissen
- Fast alle Testpersonen haben positiv auf REA reagiert und ziehen REA einem Menügesteuerten Interface vor

Grandchair



- These: Grossväter erzählen ihren Enkeln oft Geschichten, und Erfahrungen und Erlebnisse aus ihrem Leben. Dies hilft ihnen ihr Leben zu verarbeiten.
- der Grandchair soll für Grossväter ohne Enkel diese simulieren

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
- Merkmale:
 - modular, verteilt und parallel
 - Haupt-Wissensbasis mit dynam. Wissen über die Konversation
 - Mehrere Fachwissen-Basen, die jeweils Wissen über einen bestimmten Kontext enthalten
 - sehr flexibel und gut zum Experimentieren

YMIR

- Die Kern-Module von YMIR 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.
- Mit YMIR können so viele Ein- und Ausgabemodalitäten implementiert werden, wie es Ein- und Ausgabe-Geräte gibt
- Für die selbe Ausgabefunktion können verschiedene Verhaltensweisen ausgeführt werden

Nutzen von A. Agenten:

Leider gibt es bisher kaum empirische Daten, folgendes lässt sich jedoch sagen:

- Anwender muss nicht lernen den Computer zu bedienen
- wechselnde Initiative
- Menschen reagieren emotional auf Gesichter und übertragen menschl. Eigenschaften auf Agenten
- A. Agenten können grössere Aufmerksamkeit erzielen
- Hemmschwelle einen Computer zu bedienen kann abgebaut werden

Nutzen von A. Agenten:

- Maschine wirkt freundlicher und lebendiger
- 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

- Verlangsamung der Arbeit, ineffizientere Arbeit
- menschl. Züge eines Agenten können zu einer Übererwartung des Anwenders gegenüber Agenten bewirken
- Keiner der bisher erbauten Agenten benötigte wirklich seinen Körper. Nur der schönen Graphik willen lohnt sich der Aufwand nicht. (Schneidermann, Computer Interfaces)

Einsatzgebiete

- virtuelle Welten
- Besonders gut dann, wenn:
 1. Computer ohne Tastatur eingesetzt werden und nicht der PC
 2. Benutzer ohne Computerkenntnisse Computern bedienen müssen
- Beispiele: in Spielzeug, Fahrkartenautomaten, in öffentl. Gebäuden wie Museen, Schulen, Universitäten

Zusammenfassung

- Wir haben hier schon interessante Ansätze gesehen, deren Weiterverfolgung lohnt.
- Leider wurden in diesem Bereich nur einzelne, wenige Projekte gestartet, die nicht aufeinander aufbauten.
- 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.
- Können jedoch die grundlegenden Fertigkeiten der menschl. Konversation nicht adäquat implementiert werden, sollte man auf den Einsatz von a. Agenten verzichten.