

Beispiele multimodaler Systeme

im Rahmen des Seminars: multimodale Mensch-Maschinen-Kommunikation

2. und 9.11.2000

Jan Kleinlützum, Henning Mersch

Motivation:

Vorbild für eine verbesserte multimodale Mensch-Maschinen-Kommunikation ist die Mensch-Mensch-Kommunikation. Sie wird als optimaler Weg angesehen, Informationen und Intentionen weiterzugeben und zu empfangen. Die Mensch-Mensch-Kommunikation läßt sich in verschiedene Modalitäten aufteilen, die jeweils ihre eigenen Schwächen besitzen. Der Mensch ist in der Lage einige davon durch vorhandenes Weltwissen auszugleichen. Gleiches möchte man nun in der Mensch-Maschinen-Kommunikation erreichen. Auch hier wird versucht die verschiedenen Schwächen durch Kombination der Modalitäten aufzulösen. Ein so erweitertes Mensch-Maschinen-Interface bietet die Möglichkeit einer wesentlich effizienteren und intuitiveren Kommunikation.

Im Laufe der letzten 30 Jahre hat sich keine einheitliche Strategie durchgesetzt, da die verschiedenen Ansätze jeweils ihre Vorteile haben. Exemplarisch werden hier einige Systeme vorgestellt.

1970	SCHOLAR TutorSystem Südafrika
1984	Put-That-There
1986	Xtra Wahlster, W.: User and Discourse Models for multimodal Communication. In Sullivan, J.W., & Tyler S.W. (Eds.): intelligent User Interface (pp.45-67): Reading (MA): ACM Press 1991 unter Federführung von: Wolfgang Wahlster - Universität des Saarlandes als universelle Schnittstelle zwischen Mensch und einem Programm geplant, aber nur als Steuerformularinfosystem realisiert. <i>Input:</i> Sprache, Tastatur, Pointing (exakt) durch TACTILUS 3 unscharfe Pointing Methoden: (Finger, Handfläche, einkreisen) Muß vorher selektiert werden <i>Output:</i> Sprache, Bildschirm (Dialogfenster) und durch das unterstützte Programm selbst. Spezialwissen muss im Programm implementiert sein.
1991	CUBRICON Neal, J.G., Shapiro, S.C.: Intelligent Multi-Media Interface Technology. In Sullivan J.W. & Tyler S.W. (Eds.): Intelligent User interface (pp.11-43). Reading (MA): ACM Press 1991 Jeannette G. Neal, Stuart C. Shapiro - CUBRICON militärisches Karteninformations- und Navigationssystem. <i>Input:</i> Pointing durch Maus exakt und grob, Spracheingabe, auch Kombination <i>Output:</i> Sprachausgabe, 2 Monitore (Kartenanzeige und Tabelle der Objekte) 4 Wissensbasen: Lexikon & Grammatik, bisherige Kommunikation, Wissen über den Benutzer, Welt-/Domänenwissen <i>Verstehen der Eingabe:</i> kann mit Ambiguitäten umgehen über Hierarchie der Wissensbasen. <i>Erzeugen der Ausgabe:</i> Wahl der Methode durch statische Regeln. Zur Not: natürliche Sprache Sollte erweitert werden um Formular für Missionsplanung.

1993	<p>AIMI Burger, J.D., Marshall, R.J.: The Application of Natural Language Models to Intelligent Multimedia. In Maybury, M.T. (Ed): Intelligent Multimedia Interfaces (pp.174-196) Menlo Park (CA), AAAI Press/The MIT Press 1993. Burger, Marshall - Rome Laboratories Frachttransport Zeit- und Routenplanung; integriert sämtliche Eingaben in eine Zwischensprache. <i>Input</i>: Natürliche Sprache und Pointing (Maus) <i>Output</i>: Visuell: Landkarten, Tabellen, BusinessCharts, Bilder akustisch: natürliche Sprache und nonverbale Signale. Extrem modular gehalten um vielseitige Verwendung zu ermöglichen. komplexe Ausgabengenerierung aus einer Zwischensprache (KL-1, KL-2)</p>
1993	<p>AlFresco Stock, O., et al.: AlFresco: Enjoying the Combination of Natural Language Processing and Hypermedia for Information Exploration. In Maybury, M.T. (Ed): Intelligent Multimedia Interfaces (pp.197-224). Menlo Park (CA): AAAI Press/MIT Press, 1993 O. Stock, AlFresco Project Team - Italien interaktives Informationssystem über italienische Meister des 14. Jahrhundert. Fresken und Monumente <i>Input</i>: natürliche Sprache, Pointing (Maus) <i>Output</i>: Bilder, Hypertext, natürliche Sprache Kombination aus explorativem Suche (Browsen) und zielgerichteter Suche Repräsentation: YAK (Abkömmling von KL-One)</p>
1994	<p>Edward Bos, E., Huls, C., Claassen, W.: Edward: Full Integration of Language and Action in a multimodal Interface. International Journal of Human-Computer Studies 40, 1994, 473-495 Carla Huls, Edwin Bos - University of Nijmegen (NL) Grundsatzfrage nach Sinn & Zweck von M4K Bringt Spracheingabe einen Geschwindigkeitsvorteil? → Nein! Was für pos./neg. Auswirkungen hat Supplemental Linguistik Output (SLO)? → Grundsätzlich positive. Verändert sich die Akzeptanz einer Modalität mit Erfahrung am System? → keine Tendenz zu erkennen. <i>Input</i>: SLO über Keyboard/Spracheingabe <i>Output</i>: SLO über Sprache und Bildschirm, Highlighting auf dem Bildschirm</p>
1995	<p>GEORAL Siroux, J., et al.: Modeling and Processing of Oral and Tactile Activities in the GEORAL System In Bunt, H. et al. (Eds.): Multimodal Human-Computer Communication (pp.101-110). Berlin. Heidelberg: Springer-Verlag 1998. J. Siroux, M. Guyomard et. al. Touristeninformationssystem, takiler Input zur Unterstützung des Sprachsystems <i>Input</i>: natürliche Sprache, Touchscreen (Pointing & komplexere Gesten) <i>Output</i>: Bildschirm (Highlighting, Zoom etc.), natürliche Sprache Simple System, Sprache dominiert die Gestik im Zweifelsfall.</p>

Fazit:

Die vorgestellten Lösungen sind alle auf ihre Spezialdomäne beschränkt. Daraus ist ersichtlich, daß eine allgemeine Lösung für die Mensch-Maschinen Kommunikation noch in weiter Ferne liegt. Die vorgestellten Ansätze in den Systemen stellen durch- aus eine Grundlage für weitere Entwicklungen dar. Über die weitere Entwicklung ist leider nichts bekannt.