

Beispiele multimodaler Systeme

im Rahmen des Seminars:

Multimodale Mensch-Maschinen-Kommunikation

Jan Kleinlützum (jkleinlu@TechFak.Uni-Bielefeld.DE)

Henning Mersch (hmersch@TechFak.Uni-Bielefeld.DE)

Allgemeines

Folgende Informationen werden über alle Systeme (soweit verfügbar)
vorgetragen:

- Formales: Wann, Wo, Von wem entwickelt, grobes Einsatzgebiet
- Grundsätzliche Idee
- Systembeschreibung
- gewonnene Erkenntnisse

Systemhistorie

1970	SCHOLAR TutorSystem Südafrika
1984	Put-That-There
1986	Xtra Interface, Steuerformularhilfe
1991	CUBRICON Militärisch, Karten
1993	AIMI militärisches Interface
1993	AlFresco Bilder, Hypertext+NLP
1994	Edward Sinn & Zweck von M4K
1995	GEORAL Touristen, Karte

Edward: Formales

- Wann:
ca. 1993/1994
- Wo:
University of Nijmegen (NL)
- entwickelt von:
Carla Huls, Edwin Bos
- grobes Einsatzgebiet:
Dateimanager / File-Info-System

Edward: Systembeschreibung

Grundsätzlich handelt es sich um eine Art Filemanager und File-Info-System.

- Input:
ist mit Sprache, Keyboard und Maus (als Pointing) möglich.
- Output:
erfolgt durch Sprache (SLO) und den Bildschirm

Edward: gewonnene Erkenntnisse

- Bringt Spracheingabe einen Geschwindigkeitsvorteil?
Kein Geschwindigkeitsvorteil feststellbar.
- Was für positive/negative Auswirkungen hat SLO?
Grundsätzlich positive Ergebnisse
- Verändert sich die Benutzung der Modalität mit Erfahrung am System?
Keine generelle Tendenz.

⇒ vieles ungeklärt

CUBRICON: Formales

CUBRC Intelligent CONversationalist

Teilprojekt von IMMI (Intelligent MultiMedia Interface)

- Wann:
vor 1991
- Wo:
CUBRC (Calspan-UB Research Center)
- entwickelt von:
Jeannette G. Neal, Stuart C. Shapiro
- grobes Einsatzgebiet:
militärische Karten-Navigation/Information

CUBRICON: Systembeschreibung

- Input:
Pointing (Maus), Spracheingabe. Auch Kombination.
- Output:
Sprachausgabe, 2 Monitore (Tabellen, Grafik, Histogramm)
- 4 Wissensbasen:
 - Lexikon & Grammatik
 - Berücksichtigung der bisherigen Kommunikation
 - vom Benutzer abhängige Einstellungen
 - Weltwissen bzw. Domänenwissen

CUBRICON: Systembeschreibung II

Verstehen der Eingabe

- *Interrogative*: “What is the mobility of <point> SAM ?”
- *Imperative*: “Display the East-West germany region!”
- *Problem*: Ambiguität “This <point> Airbase”

Erzeugen der Ausgabe

- Auswahl des Ausgabe Typs anhand von statischen Regeln
- evtl. mehrfache Darstellung der gleichen Informationen
- natürliche Sprache als letzte Möglichkeit
- Histogramm noch nicht implementiert

XTRA: Formales

eXpert TRAnslator

- Wann:
1986 / 1991
- Wo:
Universität des Saarlandes im Rahmen des SFB314
- unter Federführung von:
Wolfgang Wahlster
- grobes Einsatzgebiet:
geplant als universelle Schnittstelle

XTRA: Grundsätzliche Idee

- Universelle Schnittstelle
- unterstütztes Programm muß kooperieren
- nur als Steuerberaterprogramm realisiert

XTRA: Systembeschreibung

- Input:
Sprache, Tastatur, Pointing in 4 Stufen
- Output:
Sprache, Bildschirm, das Programm selbst

XTRA besteht aus 3 Fenstern.

interne Representation: “organization Graph”

Spezial-Wissensbasis ist im Programm.

XTRA: Pointing durch TACTILUS

XTRA implementiert selbst nur exaktes Pointing

TACTILUS

- unterstützt unscharfes Pointing. (vorherige Auswahl)
- “pars-pro-toto” Interpretation
- benutzt Wissensbasen von XTRA

GEORAL: Formales

- Wann:
1992-1995
- Wo:
Institut de recherche en informatique et systemes aleatoires, Rennes
Ecole Nationale Supérieure de Sciences Appliquées et de Technologie,
Lannion
- entwickelt von:
Siroux, Guyomard, Multon, Remondeau
- Einsatzgebiet:
Touristen-Informationssystem für die Trégor-Region

GEORAL: Übersicht

Idee: Ausnutzung der Stärken der verschiedenen Modi Eingabe:

- (natürliche) Sprache
- Punkte/Regionen auf Touchscreen

Beschränkt: Eingabe nur bei “Prompt”

Ausgabe:

- synthetische Sprache
- Grafik(Highlighting, Zooming)

Integration: Sprache dominiert taktilen Input

GEORAL: Erkenntnisse

- 3 Arten der Beziehung sprachlich/taktil
 - Bindung
Are there any beaches in this <point to Lannion> locality?
 - Bestätigung
Are there any beaches in <point to Lannion> Lannion?
 - Ersetzung
Where are the campsites? <point to Lannion>
- Taktile Geste muss im Kontext betrachtet werden

⇒ vielversprechender Prototyp

AIMI: Formales

AIMI: An Intelligent Multimedial Interface

MACPLAN: Military Airlift Command PLANner

- Wann:
vor 1993
- Wo:
Rome Laboratories/Air Force
- entwickelt von:
Burger / Marshall
- grobes Einsatzgebiet:
Fracht-Transport Zeitpläne/Routen

AIMI: Grundsätzliche Idee

Verbesserung von MACPLAN: Ersetzen von ATN durch KING KONG

- Kontext-Tracking über alle Modi
- Ausgabe in “vernünftigem” Modus
- Ausgabe bietet weiteren Dialog an
- Erweiterbarkeit durch Modularität
- Transparenz für den Nutzer

AIMI: Systembeschreibung

- Input: Natürliche Sprache, Maus-Gesten
- Output: Visuell: Landkarten, Business Charts, Bilder, Tabellen
Audio: Natürliche Sprache, nonverbale Signale
- Abbildung des gesamten Kommunikationsprozesses auf Wissensbasis
- Ausgabemodus kann erzwungen werden

AIMI: gewonnene Erkenntnisse

- Abstraktes Interface ermöglicht Annäherung an natürlich-sprachliche Systeme
- Es wird nicht nur *eine* Antwort generiert
- Durch komplettes Konversations-Tracking Rückschlüsse auf die Intention des Nutzers .

AlFresco: Formales

- Wann: 1993
- Wo: Italien
- entwickelt von: Oliviero Stock, Alfresco-Team
- grobes Einsatzgebiet:
Browsen durch Datenbank italienischer Frescen des 14. Jahrhunderts

AlFresco: Grundsätzliche Idee

Integration von Natural Language Processing und Hypermedia.

- Hypertext eignet sich zum Explorieren und Browsen
- NLP zur gezielten Suche
- Durch gegenseitige Ergänzung deutliche Verbesserung

AlFresco: Systembeschreibung

Input: Sprache, Klicken mit Maus auf Hyperlinks: Output: Sprache, Hypertext (Auswahlmenüs, Bilder, Filmsequenzen)

Model of Interests: Benutzerführung wird automatisch angepasst

AlFresco: gewonnene Erkenntnisse

- Deutlich verbesserte Handhabung