

Ein Goal-basierter Ansatz für adaptive multimodale Systeme

Frank Honold, Felix Schüssel, Michael Weber [†]
Gregor Bertrand, Florian Nothdurft, Wolfgang Minker [‡]

Universität Ulm

[†] Institut für Medieninformatik, [‡] Institut für Informationstechnik

Zusammenfassung

Zukünftige adaptive und multimodale interaktive Systeme stellen besondere Anforderungen an die Modellierung der Interaktion sowie an die Systemarchitektur. Wir präsentieren ein modalitätsunabhängiges, flexibles und hierarchisches Dialogmodell, das mittels Dialog-Goals nicht nur eine dynamische Arbitrierung von Modalitäten, sondern auch eine Adaption des Dialogverlaufs zur Laufzeit ermöglicht. Abschließend stellen wir eine Architektur vor, welche die Realisierung eines solchen Dialogmodells durch ein adaptives Dialogmanagement und flexible Fusions- und Fissions-Komponenten erlaubt.

1 Einleitung

Bestehende Systeme sind oftmals für einen spezifischen Einsatzbereich konzipiert und beschränken sich in der Regel auf ein statisches UI mit einer rigiden Interaktionsabfolge. Angelehnt an Normans *Execution and Evaluation Cycle* (Norman 2002) lässt sich die Interaktion eines Menschen mit einem System in vier Phasen einteilen: *Decision*, *Action*, *Perception* und *Interpretation* (Dumas 2009). Zukünftige adaptive Systeme sollen es ermöglichen, diese Aufgaben flexibel im Dialog mit dem Benutzer zu realisieren. Der Dialog sollte dabei multimodal realisiert werden können, um mögliche Vorteile einer solchen Interaktion nutzbar zu machen (vgl. Oviatt 2006). Dies stellt besondere Anforderungen, gerade was die Modellierung des Dialogs und die für die Realisierung nötige Systemarchitektur betrifft.

2 Anforderungen an Dialog und System

Das Dialog-Modell für ein adaptives, multimodales System sollte verschiedene Anforderungen erfüllen. Einzelne Dialogakte (z.B. Äußerungen und Fragen) oder größere Dialoge (z.B. Flugbuchung, Hotelreservierung) sollen als Bestandteile eines Gesamtdialogs *hierarchisch*

strukturierbar sein. So können einzelne Akte zu komplexeren Komponenten zusammengefasst und in einem Gesamtsystem als komplexer Dialog algorithmisch geplant werden. Um eine dynamische Integration unterschiedlicher Modalitäten für die System Ein- und Ausgabe zu unterstützen, wird ein *modalitätsunabhängiges Modell* der zu kommunizierenden Inhalte auf Dialogebene benötigt. Ein modellierter Dialog soll außerdem eine gewisse *Flexibilität hinsichtlich des Dialogverlaufs* erlauben, so dass bestimmte Dialogakte in ihrer zeitlichen Abfolge beliebig strukturierbar sind, solange sie sich nicht gegenseitig in ihrer Reihenfolge bedingen. Es sollte des Weiteren möglich sein, auf einfache Art und Weise einzelne kooperative Dialogfragmente (z.B. das Anzeigen einer zusätzlichen Information) ad-hoc in einen bestehenden Dialog zu integrieren und so den *Dialog dynamisch zu erweitern*.

Die Realisierung eines solchen Dialogmodells erfordert eine Architektur, die nicht nur die Verwaltung des Modells bewerkstelligt, sondern dieses mittels *multimodaler Ein- und Ausgabemöglichkeiten* realisiert. Der Benutzer soll wahlfrei bezüglich seiner Eingabemöglichkeiten sein. Analog dazu, können systemseitig (je nach Benutzer, Kontext und Information) ein oder mehrere Ausgabemodalitäten in unterschiedlichen Kombinationen realisiert werden.

3 Modellierung der Dialog-Goal-Hierarchie

Die hierarchische Dialog-Modellierung des hier vorgestellten Ansatzes integriert diese Anforderungen in ein Modell (vgl. Abbildung 1). Dabei wird immer von einer Intention des Benutzers ausgegangen, die (theoretisch) durch beliebig viele Goals seitens des Systems realisiert werden kann. Diese Goals können über beliebig viele Hierarchiestufen mittels *Sub-Goals* verfeinert werden. Die unterste Ebene, also die Blätter des Dialogbaumes, entsprechen dann *modalitätsunabhängigen, kommunikativen Akten*, angelehnt an klassische Sprechakttheorie (Searle 1969, Jurafsky & Martin 2000, Harris 2005).

Ein Goal kann vertikal mehrere Sub-Goals referenzieren. Die horizontale Struktur kann durch so genannte *Next-Goal Links* vorgegeben werden, um eine Abfolge von Dialogschritten (eine Dialogstrategie) zu definieren. Zusätzlich kann jedes Goal als *system-initiative* oder *user-initiative* gekennzeichnet sein. Durch entsprechende Kennzeichnung der Goals kann so ein *mixed-initiative* Dialog modelliert werden. *Outputs* und *Inputs* dieser kommunikativen Akte definieren modalitätsunabhängig Aus- und Eingaben des Systems.

Jedem Goal können als Vorbedingung sogenannte *Guard*-Variablen, bzw. als Resultat sogenannte *Effect*-Variablen zugeschrieben werden, die festlegen, welche Variablen-Belegungen für die Aktivierung eines Goals gelten müssen, bzw. welche Variablen nach der Ausführung belegt werden. Dadurch wird u.a. die geforderte Flexibilität des Dialogverlaufs gewährleistet. Zusätzlich ermöglicht das Betrachten von *Guards* und benötigten *Effects* zur Laufzeit die Integration zusätzlicher Goals oder Hierarchien und somit die dynamische Erweiterbarkeit des Dialogs. Details und Beispiele zur Dialogmodellierung finden sich in (Bertrand 2011).

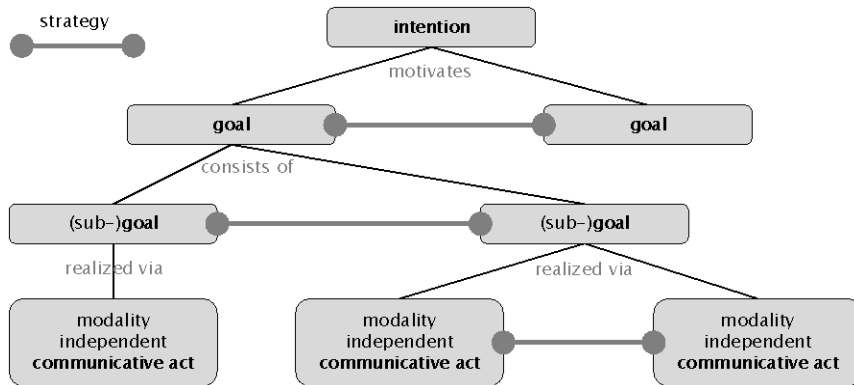


Abbildung 1: Konzept eines hierarchischen, Goal-basierten Dialogs. Realisierung siehe (Bertrand 2011)

4 Systemarchitektur

Die für die Realisierung eines so modellierten Dialogs notwendige Systemarchitektur (Abbildung 2) ist an die eingangs erwähnte Mensch-Computer Interaktionsschleife angelehnt. Der Ansatz lässt sich systemseitig in drei funktionale Bereiche einteilen. Das *Dialog Management* verwaltet das hierarchische Dialogmodell und steuert die zeitliche Abfolge der aus den Goals resultierenden kommunikativen Akte. Das *Interaction Management* nimmt die kommunikativen Akte entgegen, leitet daraus entsprechende multimodal realisierte Interaktionen ab und analysiert Benutzereingaben. Das *Device Management* ist für die Anbindung der Geräte sowie für die zeitlich geordnete Ein- und Ausgabe verantwortlich.

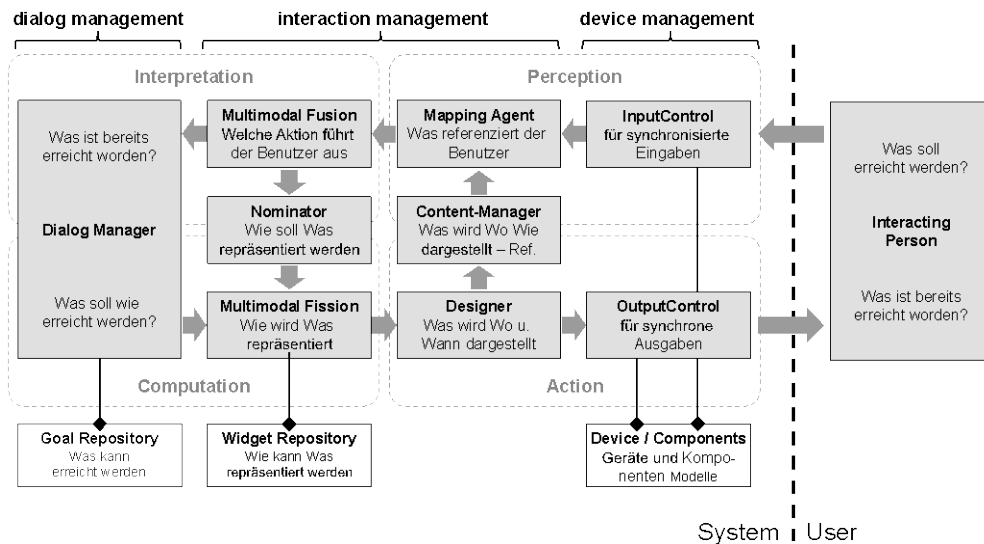


Abbildung 2: Architekturansatz für Goal-basierte adaptive multimodale Systeme

5 Schlussbemerkung

Lange Zeit folgte die Entwicklung interaktiver Systeme der Annahme, dass sich der Mensch an jedwede Benutzungsschnittstelle anpassen kann, welche ihm durch einen Designer allgemeingültig vorgegeben wurde. Mit dem hier vorgestellten Ansatz können Interaktionsdialoge modalitätsunabhängig und flexibel modelliert werden. Für zukünftige Systeme bietet dies den Vorteil, dass sie sich bezüglich ihrer Interaktionsmöglichkeiten an einen individuellen Benutzer adaptieren können, sowohl in ihrem Verhalten, als auch in ihrer Erscheinung.

Das vorgestellte Dialogmodell ist durch seine mit Vor- und Nachbedingungen annotierbaren Goals auch durch einen KI-Prozess bedingt planbar (Biundo 2011). Verändern sich bspw. bestimmte Variablen eines Weltmodells und verhindern das Aktivieren eines Goals durch einen nicht erfüllten Guard, so kann zur Laufzeit ein zusätzliches Goal aus einem vorhandenen Repository eingefügt werden, welches als Effect die entsprechende Guard-Variable ändert. Dies erlaubt es zur Laufzeit, den Dialogverlauf je nach Situation zu ändern, ohne ihn vollständig a priori modelliert zu haben. Aktuell ist das Dialogmodell komplett spezifiziert und es existiert eine Tool-Unterstützung zur Erstellung von Dialogen (Bertrand 2011). Die vorgestellte Systemarchitektur und die enthaltenen Komponenten befinden sich noch in der Konzeptphase.

Literaturverzeichnis

- Bertrand, G., Nothdurft, F., Honold, F. & Schlüssel, F. (2011). CALIGRAPHI – Creation of Adaptive Dialogues Using a Graphical Interface. *COMPSAC 2011: IEEE Computer Software and Applications Conference*.
- Biundo, S., Bercher, P., Geier, T., Müller, F. & Schattenberg, B. (2011). Advanced user assistance based on AI planning. *Cognitive Systems Research*. Volume 12, Issues 3–4, S. 219–236.
- Dumas, B., Lalanne, D. & Oviatt, S. (2009). Chapter 1: Multimodal Interfaces: A Survey of Principles, Models and Frameworks. *Human Machine Interaction – Research Results of the MMI Program. Lecture Notes in Computer Science 5440/2009*. Berlin: Springer-Verlag. S. 3-26.
- Harris, R. A. (2005). *Voice interaction design: crafting the new conversational speech systems*. San Francisco: Morgan Kaufmann Publishers, an imprint of Elsevier Science.
- Jurafsky, D. & Martin, J.H. (2000). *Speech and Language Processing*. London: Pearson-Prentice Hall.
- Norman, D. A. (2002). *The Design of Everyday Things*. New York: Basic Books.
- Oviatt, S. (2006). Human-centered design meets cognitive load theory: designing interfaces that help people think. *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*. New York: ACM. S. 871-880.
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press