

Inkrementelle und zustandsbasierte Verarbeitung in einem hybriden Sprachdialogsystem

Sprachdialogsysteme sind Computersysteme, die gesprochene Sprache für die Mensch-Computer-Interaktion nutzen. Ein Hauptproblem der Systeme ist ihre oft mangelhafte Natürlichkeit. Die inkrementelle (schritt haltende) Verarbeitung gesprochener Sprache ermöglicht es Dialogsysteme natürlicher und intuitiver zu gestalten [1]. Systemreaktionen können früher erfolgen und Nutzerreaktionen lassen sich rasch in die laufende Verarbeitung integrieren. Da die inkrementelle Verarbeitung eine Neugestaltung der Systemarchitektur erfordert, stehen vollständige inkrementelle Dialogsysteme für Anwendungen mittlerer Komplexität noch aus. Wir präsentieren einen Hybridansatz, in dem ein traditionelles, zustandsbasiertes Dialogsystem den Rahmen des Dialogablaufs absteckt und einzelne, gut herauslösbare Teilaufgaben mithilfe eines inkrementellen Subsystems gelöst werden. Der Übergang der Kontrolle zwischen den Systemen erfolgt dabei ohne jeglichen Bruch und in einer für den Nutzer transparenten Weise.

Das Anwendungsszenario ist ein geometrisches Puzzlespiel. Die Interaktion lässt sich gut durch abwechselnde Teileauswahl und -positionierung modellieren. Die Identifikation eines Puzzleteils ist durch die Interpretation von referierenden Ausdrücken, Nachfragen nach bestimmten Merkmalen des Teils, oder notfalls durch Ausprobieren gut lösbar („Meinten Sie dieses Teil?“). Ein vordefinierter Unterdialog zur Zielauswahl ist hingegen für die Positionierung der Teile ungünstig, weil für jedes zu legenden Muster eine Vielzahl von Positionen und damit Ausdrücken möglich ist – sowohl für die Beschreibung durch den Nutzer als auch für Nachfragen durch das System. Wir wählen deshalb eine inkrementelle, proaktive Strategie: Das System initiiert die Bewegung des Puzzleteils, wodurch der Nutzer veranlasst wird diese zu steuern anstatt eine problematische Beschreibung der Zielposition zu liefern. Durch diese Gestaltung nach dem Affordance-Prinzip [4] sind die zu erwartenden Nutzeräußerungen und ihr Vokabular deutlich eingeschränkt. Durch die inkrementelle Verarbeitung haben die Befehle eine unmittelbare Wirkung, was ersten Untersuchungen nach zu einem sehr natürlichen Systemverhalten führt.

Dieses hybride Verfahren ist immer dann geeignet, wenn bereits die Inkrementalisierung nur eines kleinen, eng umrissenen Unterdialogs eine spürbare Verbesserung der Benutzbarkeit erwarten lässt, während der übrige Dialog auf traditionelle Weise modelliert werden kann.

[1]: Aist, G., Allen, J., Gallo, C., Stoness, S., Swift, M. & Tanenhaus, M.: „Incremental dialogue system faster than and preferred to its nonincremental counterpart,“ in *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, 2007, Nashville, USA.

[2]: Bobbert, D. & Wolska, M.: „Dialog OS: an extensible platform for teaching spoken dialogue systems,“ in: *Proceedings of Decalog 2007*, Trento, Italy.

[3]: Baumann, T. & Schlangen, D.: „The InproTK 2012 release: a toolkit for incremental spoken dialogue processing,“ in: *Proceedings of the SDCTD Workshop*, 2012, Montréal, Canada.

[4]: Jenkins, H.: „Gibson's Affordances: Evolution of a Pivotal Concept,“ in: *Journal of Scientific Psychology* **3**, pp. 34-45, 2008.