

Inkrementelle und zustandsbasierte Verarbeitung in einem hybriden Sprachdialogsystem

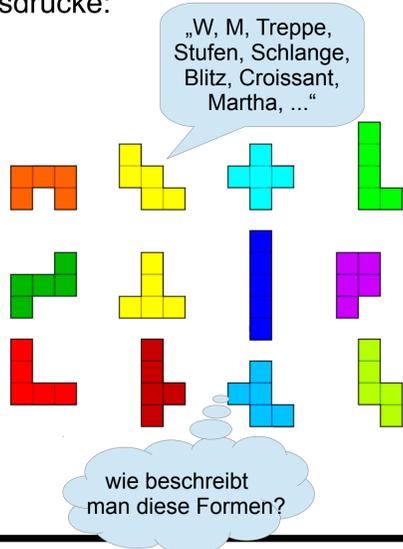
Zustandsbasierte Dialogsysteme sind zwar einfach implementierbar, ihr Ping-Pong-Ansatz der Dialogablaufsteuerung schränkt allerdings die Nutzbarkeit auf statische Domänen (in denen die Veränderungen langsam relativ zur Sprechgeschwindigkeit erfolgen) ein.

Im Gegensatz dazu bieten inkrementelle Dialogsysteme die Möglichkeit, Aktionen auszuführen während der Nutzer spricht, oder Nutzerfeedback in laufenden Systemausgaben zu berücksichtigen; hierfür stehen allerdings noch keine vollständigen, leicht integrierbaren Dialogmanager zur Verfügung.

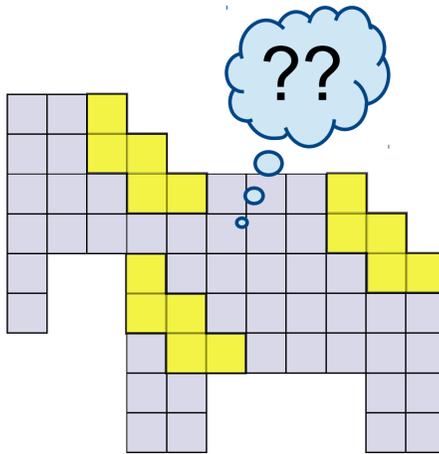
Ziel dieser Arbeit ist die Integration zustandsbasierter und inkrementeller Teilsysteme, welches die jeweiligen Stärken beider Ansätze kombiniert.

Referierende Ausdrücke als Quelle von Ausdrucksvielfalt (selbst in einfachen Domänen, hier Pentomino-Puzzle)

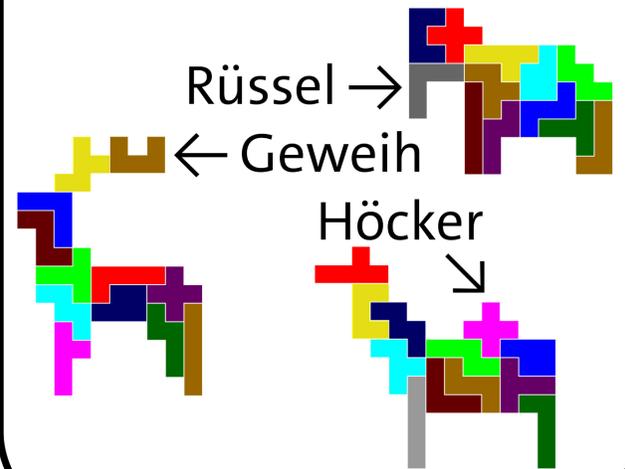
- Geschlossene Domäne (12 Puzzleteile)
- Klar unterscheidbare Eigenschaften (Farbe, Form?, relative Anordnung, ...)
- Dennoch komplexe referierende Ausdrücke:



- Sehr große Anzahl an Zielpositionen
- Keine Unterscheidbarkeit durch Farbe/Form, nur Anordnung



- Aufgabenspezifisches Vokabular für die Zielpositionen
- Schlechte Übertragbarkeit auf neue Puzzlefiguren

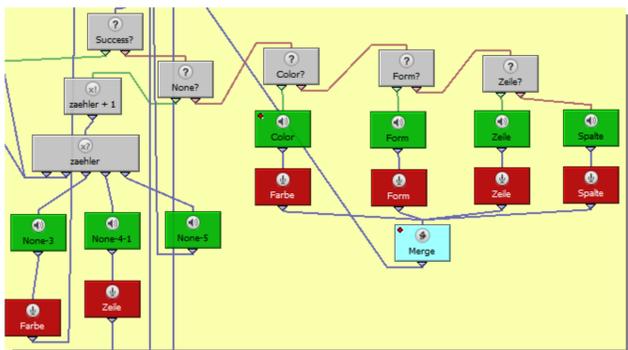


- Problem: Referierende Ausdrücke komplex zu generieren und zu analysieren
- Lösung: Steuerung der Positionierung über Interaktion, bessere Führung des Nutzers

Zustandsbasiertes System mit DialogOS

Teilauswahl und übergeordnete Dialogsteuerung

- Grammatik für (relative) komplexe Äußerungen
- Unterdialoge zur Disambiguierung
- Spätes Feedback (erst nach der Äußerung)

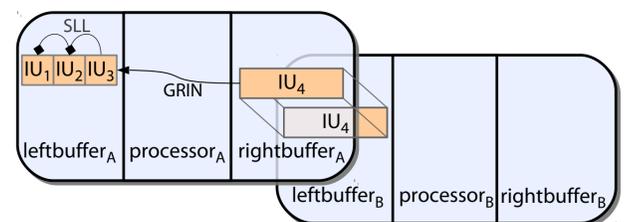


Sys: Welches Teil soll ich jetzt nehmen?
 Usr: Nimm das grüne Teil.
 Sys: Welche Form hat das grüne Teil?
 Usr: Es sieht aus wie ein C.
 Sys: Ok, wo soll ich das Teil plazieren?
 Sys: <beginnt das Teil zu bewegen>
 Usr: Ja, genau, weiter runter, etwas links genau da!
 Sys: Liegt das Teil so richtig?
 Usr: Ja!

Inkrementelles System mit InproTK

Teilpositionierung durch direktes Steuern der Bewegung

- Einfache Spracherkennung mit kleinem Vokabular
- Sechs Kommandos für Richtungsangaben ausreichend
- Direktes Feedback (während der Äußerung)



Steuerung der Bewegung:

- Ermöglicht überhaupt erst Anpassung an andere Puzzlefiguren und „falsches“ Positionieren von Puzzleteilen
- Verschiebung von Komplexität weg von sprachlichen Äußerungen hin zur Interaktion
- Große Robustheit, da Systemfehler direkt korrigiert werden können
- System-initiierte Bewegung (vs. warten auf Nutzeräußerung):
- Wurde von allen Nutzern als Hinweis auf die Steuerfähigkeit des Systems angenommen
- Führt zu signifikant kürzerer Bearbeitungsdauer
- Wird als transparenter und reaktiver bewertet
- System nutzt die „Affordance of Motion“ zur Führung des Nutzers