Mehrere Personen

Kernproblematik: welche Person spricht? Fehlerquelle in der Sprachanalyse, der Personenzuweisung. Differentierung nach Voiceprint, wie schnell, wie verlässlich? Wahrscheinlichkeiten weiterreichen → explodierende Komplexität.

Im Dialogfluß - wie beziehen sich Antworten aufeinander? Separate Gespräche zur selben Zeit, wie überschneiden sie sich? Welcher Dialogpartner weiß was? (Dauerndes Nachfragen, auch falscherweise, wenn das System wechselnden Dialogpartner erkennt). Wenn eine Antwort von einer anderen als der vorherigen Person gegeben wird, ist diese Antwort zum Dialog der vorherigen Person oder einem anderen Frame?

Gespräch unter den Personen → fürs System einfach nur Sprache: Diskriminierung zwischen Dialog zwischen Personen und Dialog mit System.

Insgesamt sehr viel mehr Fellerquellen und Systemlogik. Jonglieren mit viel mehr Wahrscheinlichkeiten → Wahrscheinlichkeit, falsch zu liegen, multipliziert sich. System wird viel schneller "viel falscher" liegen(?) - Debugging und Produktreife ist schwerer.

Zusätzliche Sensoren wie Kamera und Gesten/Blick-Tracking wird essentieller.

Mehrsprachigkeit

Ein bisschen wie mehrere Personen. Aber durchaus einfacher.

Problematik - wenn Sprachen gekreuzt werden, explodiert der grammatikraum. Das heißt die Grammatik hat viel mehr Fehlerquellen (läuft gegen das Gebot, Grammatik einzuschränken um Fehler zu kontrollieren). Probability handover - Kompelxitäts-Runaway

Auf Design-Ebenene: andere Wörter, andere Grammatik, andere (kulturelle) Bedeutungen, Kontext, andere Möglichkeiten das selbe (oder ganz andere Dinge) zu sagen. Sprachdialog-Design braucht native Sprecher. Im Dialog-Fluss-design und der Antort-Logik muss mitgedacht werden.

Wenn keine nativen Sprecher und Grammatik mit automatischer Übersetzung - Fehlerquelle von falsch-übersetzungen, andere Semantik und Pragmatik - System kann nicht reagieren, reagiert falsch, etc. Kann auch passieren wenn System nicht perfekt ist.

Wenn zwei Sprachen oder mehr verwendet werden - in welcher Sprache antwortet das Dialogsystem dann?

Was passiert wenn die Sprache von der Person mitten im Satz gewechselt wird?

Evaluation

"Ist das Program korrekt konzipiert"? "Proven Correct Algorithm" - sehr schwierig

Große Problem: wir arbeiten viel mit Wahrscheinlichkeiten - also noch weniger genau spezifizierbare eingaben. Wann ist versagen akzeptabel, wann nicht? Wie setzen wir Standards von akzeptabler Reaktion und was verlangt dem System zu viel ab?

Wie decken wir die Breite an menschlichen Sprechern ab? Unterschiedliche Muttersprache, "physiologische Phonetik(?)", all so etwas beinflusst die Ausprache - wieviel muss abgedeckt sein, mit welcher Zuverlässigkeit?

Welche Sprache und Grammatik muss abgedeckt sein? "Alltagsdeutsch" vs "Hochdeutsch", die Dialekte und Kreole verschiedener Gruppen, Neogolismen und Slang,

Wie gehen wir mit Menschen mit Behinderungen um? Wenn das SDS Kernteil des Interaktionsdesigns ist, sollte nach Value-driven Design eine Inklusion und dann entsprechend würdige Redundanz existieren.

Anwendungen

Die große Problematik mit Sprachdialogsystemen in der Anwendung: vom Design her wollen wir sehr gut gliederbare, geordnete Dialoge. Menschliche Sprache läuft dem entgegen - under Dialog ist mit das dynamischte und ungebundene.

Ein stratifizierter Dialog lässt sich durchaus übersichtlicher gestalten auf einer grafischen Oberfläche.

Die Problematik zu vielen Anwendungen ist der Bedarf für eine fehlerabfangende, "tiefere" Intelligenz um Dialog rund um eine Anwendung im dynamischen Fluss menschlicher sprachlicher Kommunikation zu gestalten. Dies läuft Design-Weisheiten alter Systeme entgegen (halte die Komplexität flach, um Fehlerquellen und Ressourcen-Escalation zu vermeiden)

Die Herausforderung also: den Horizont der Grammatik und internen Logik vorantreiben (am besten High level, damit sich Designer am wenigsten mit dem Detail-Design der Dynamik herumschlagen müssen - solche Fähigkeiten in modulare Module abkapseln die bei vielen SDS genutzt werden können), aber gleichzeitig bewusst sein, wann Komplexität zu stark eskaliert.