

Lösungsvorschlag für das Übungsblatt 2.

Aufgabe 1

Aus Ihren Beiträgen

b)

Das generelle Dialogverhalten ist akzeptabel. Es ist im Allgemeinen klar welche Aussagen vom Kunden erwartet werden und es werden mehrere ähnliche Schlüsselwörter akzeptiert. So versteht das Programm sowohl "ja", als auch "richtig" um eine Aussage zu bestätigen. Das Grundgerüst des Systems ist relativ simpel aufgebaut. Es werden schematische Aussagen vom Kunden erwartet, die mit Hilfe eines Beispiels erläutert werden. Falls das Programm eine Aussage nicht versteht. So fängt es an das Gesamtkonstrukt in Unterpunkte zu zerlegen und jede Information einzeln abzufragen. Anschließend wiederholt das System die verstandenen Daten und lässt sich die Korrektheit vom Benutzer bestätigen. Erst wenn alle benötigten Daten vollständig Vorhanden sind sucht das System nach einer zutreffenden Verbindung und liest diese vor. Danach erklärt es einem wieder die Auswahlmöglichkeiten. So kommt es, dass das Programm in einem eng abgesteckten Rahmen abläuft, was es dem Kunden einerseits unmöglich macht Informationen in Erfahrung zu bringen, die von dem Systemdesignern nicht vorgesehen wurden, andererseits wird dadurch die Erfolgsquote von richtig erkannten Aussagen des Kunden erheblich erhöht. Probleme treten vor allem bei undeutlicher Aussprache oder Aussagen auf, die nicht dem Erwartungshorizont des Programms zu der jeweiligen Frage entsprechen.

c)

Das Grundgerüst des Programms ist wie schon erwähnt relativ starr und eng abgegrenzt. Die Grundidee scheint zu sein, dass das Programm durch das Gespräch führt. Zu jeder der Auswahloptionen wurden dem Programm eine feste Anzahl von Daten einprogrammiert, welche es vom Benutzer in Erfahrung bringen muss um dann anhand der Daten das gewünschte Resultat zu liefern. In Kurz: Es werden dem Benutzer verschiedene Auswahlmöglichkeiten geboten, der Benutzer entscheidet sich für eine und das Programm fragt ihn die für diesen Menüpunkt vorgesehenen Daten ab. Evtl. Gibt es auch noch Untermenüs ähnlich einer Baumstruktur. Ein ausbrechen aus dieser Struktur von Seiten des Programms oder des Benutzers sind nicht vorgesehen.

d)

System nimmt den Anruf an --> Begrüßungsansage. Dann läuft das Programm immer nach dem selben Prinzip ab: Programm stellt Auswahlmöglichkeiten zur Verfügung --> Benutzer wählt --> Programm fragt Daten ab --> Programm verarbeitet Daten. --> Programm gibt Resultat an Benutzer aus --> Programm stellt Auswahlmöglichkeiten zur Verfügung -> usw. usw. Sollte vom Programm etwas nicht verstanden werden, so fragt es nach oder untergliedert die einzelnen Datensätze in kleinere Einzelaufnahmen.

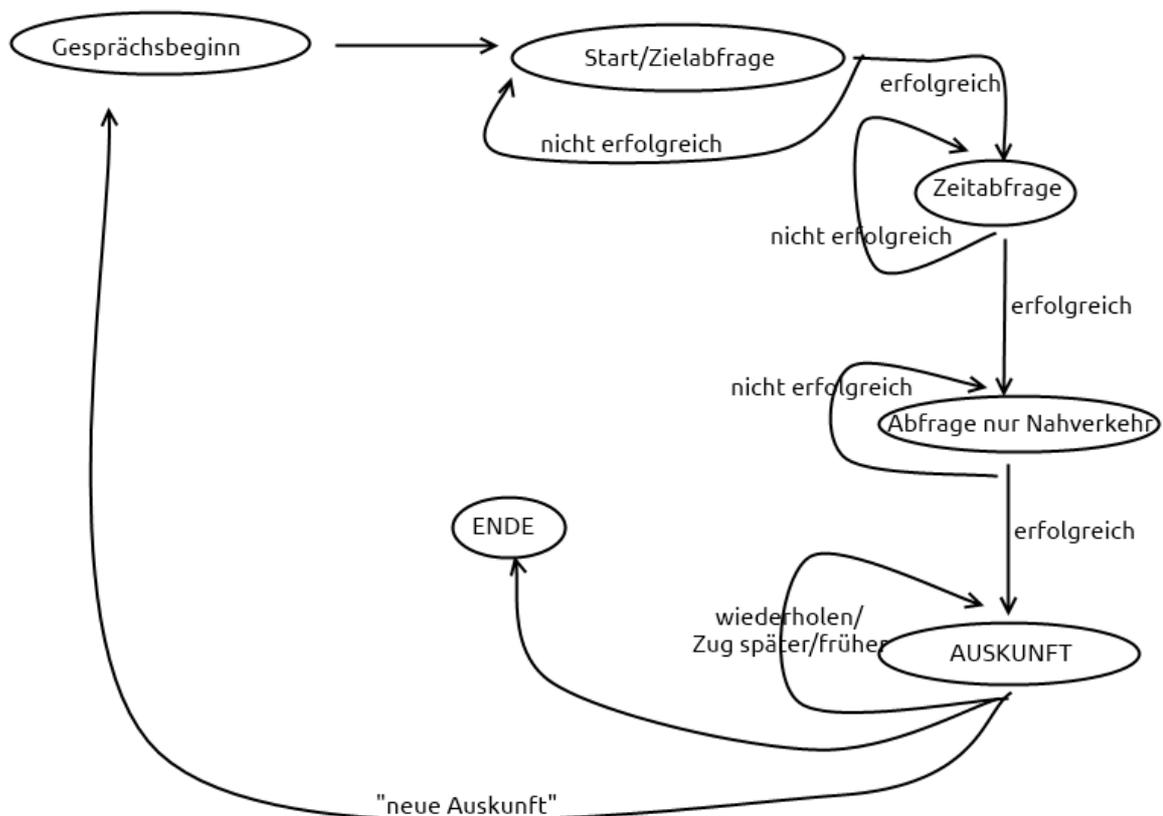
Noch ein Beispiel...

b) Das System wirkt durchaus weit fortgeschritten, mit ein bisschen Geduld bekommt man seine gewünschte Auskunft. Probleme sind mir folgende aufgefallen: Es gibt kein System für Rückfragen (wie bitte?), was bei der ziemlich schlechten Sprachqualität durchaus manchmal hilfreich wäre. Die Erkennung der Spracheingabe funktioniert meistens, allerdings nur, wenn man sich strikt an das vorgegebene Raster hält, Variationen davon überfordern das System schnell – meine Angabe „8 Uhr abends“ wurde wie 8:00 Uhr (morgens) interpretiert.

c) Ich könnte mir denken, dass das System die Eingabe mit einer Datenbank an möglichen Angaben abgleicht. Sofern die Eingabe mit hoher Wahrscheinlichkeit zugeordnet werden konnte, wird sie wiederholt und akzeptiert; ist die Sicherheit zu niedrig, wird (mit Beispielangabe) erklärt, das man sich Wiederholen soll.

Generell läuft jedes Gespräch nach einem vorgefertigten Raster ab, das System kommt nicht damit zurecht, wenn man zuerst die gewünschte Abfahrtszeit sagt oder den Zielbahnhof vor dem Ankunftsbahnhof nennt.

d)



Aufgabe 2.

Fehler bei

Mary:

- Datumsangaben: bei der vorgegebenen Schreibweise „12.3.1998“ werden alle Ziffern und Punkte einzeln vorgelesen. Bei der Eingabe „12. 3. 1998“ (mit Leerzeichen) werden Ordinalzahlen als solche erkannt, 1998 wird als eine Zahl ausgesprochen, jedoch nicht als Datumsangabe erkannt (Ausgabe „eintausendneunhundertachtundneunzig“ anstelle von „neunzehnhundertachtundneunzig“).
- kein Erkennen von englischsprachigen Namen (Blair) → in deutscher Mary nur deutsche Phoneme
- falsche Betonung
- Pausen, wo keine sein sollten, keine Pausen wo welche sein sollten
- Fehlen von genauer Artikulation

AT&TTS

- monotone Stimme
- falsche Aussprache bei falsch vordefinierter Sprache

Folgende Teilsysteme sind aktiv:

Tokenisierung (Zerlegung in einzelne Wörter)

POS-Tagging (Wortartenbestimmung)

Lexikalische Analyse (im Lexikon gibt es zu jedem Eintrag Informationen über die Wortart und andere grammatische Kategorien und außerdem die Aussprache in Form einer phonetischen Transkription .

Analyse des syntaktischen Kontextes (korrekte Lesart)

Prosodische Analyse (Festlegung von Phrasengrenzen und Akzenten).

Phonologische Analyse und Aussprache (die Aussprache eines Wortes ist durch seine Transkription im Lexikon gegeben. Im Eingabetext auftretende Wörter, die nicht im Lexikon enthalten sind, werden durch Ausspracheregeln transkribiert. Deswegen wird der Satz bei der vordefinierten Sprache bei AT&& TTS (z.B. Englisch, wobei ein deutscher Satz eingegeben wurde, nach englischer Art ausgesprochen).

Aufgabe 3.

1. Spracherkennung : kann auf einzelne Sprecher adaptiert und durch Sprachmodell und/oder Spracherkennergrammatik auf Domäne zugeschnitten werden.
2. Sprachinterpretation : verschiedene Parsingtechnologien von Schlüsselworterkennung bis kompletten Parsern.
3. Multimodale Fusion : Erkennung von verschiedenen Modalitäten.
4. Diskursinterpretation : slot-filling (das schrittweise Zusammensetzen der Benutzerintention über mehrere Interaktionen hinweg).
5. Dialogmanager: Berechnung der Systemreaktion, Planung der Ausgabeinhalten.
6. Präsentationsplanung: Festlegung genauer Inhalte der Ausgabe.
7. Multimodale Fission (Medienallokation) : Verteilung der Inhalte auf die verschiedenen Modalitäten.
8. Grounding : Verständniskontrolle durch die Wiederholung (Wiedergabe) der Eingabe.
9. Sprachgenerierung/Sprachsynthese , Synchronisation der parallel erzeugten Ausgaben.