

The Virtual Interactive Presenter (VIP)

The Virtual Interactive Presenter



Centrum für Informations- und Sprachverarbeitung

The Virtual Interactive Presenter (VIP)

Was ist VIP?

- Dialogbasiertes Interface für einen „Electronic Programme Guide“
 - System unterstützt Benutzer bei der Spezifizierung der Programmbeschreibung
 - Arbeit mit verschiedenen Auswahlkriterien z.B. Filmgenre, Schauspieler, Sender usw.
- ➔ Vorwissen über das TV-Programm sowie über die Auswahlkriterien ist nicht nötig



The Virtual Interactive Presenter (VIP)

Was ist VIP?

- Benutzer kommuniziert mit einem „human-like character“ auf der Basis von Spracherkennung
(Animiertes Interface basiert auf Microsoft Agent™ System)
- Dialoginitiative ist wechselseitig



The Virtual Interactive Presenter (VIP)

Dialogbeispiel



- U1: Do you have any sports programmes
- S2: *I have found 5 programmes for this choice. Would you like to watch "Keegan's greatest games"?*
- U3: Can I have some cricket instead?
- S4: *I can suggest this programme with cricket: "Sportstalk"*
- U5: Who's conducting?
- S6: *"Sportstalk" features Donald*
- U7: Anything else?
- S8: *I have no other programmes for that selection*
- U9: Is there anything funny, then?
- S10: *I can recommend the following comedy: "Starting Out"*
- U11: I do not want a sitcom
- S12: *What would you like to see, then?*
- U13: A funny movie

Aufbau des Programms

Die 3 wichtigsten Komponenten:

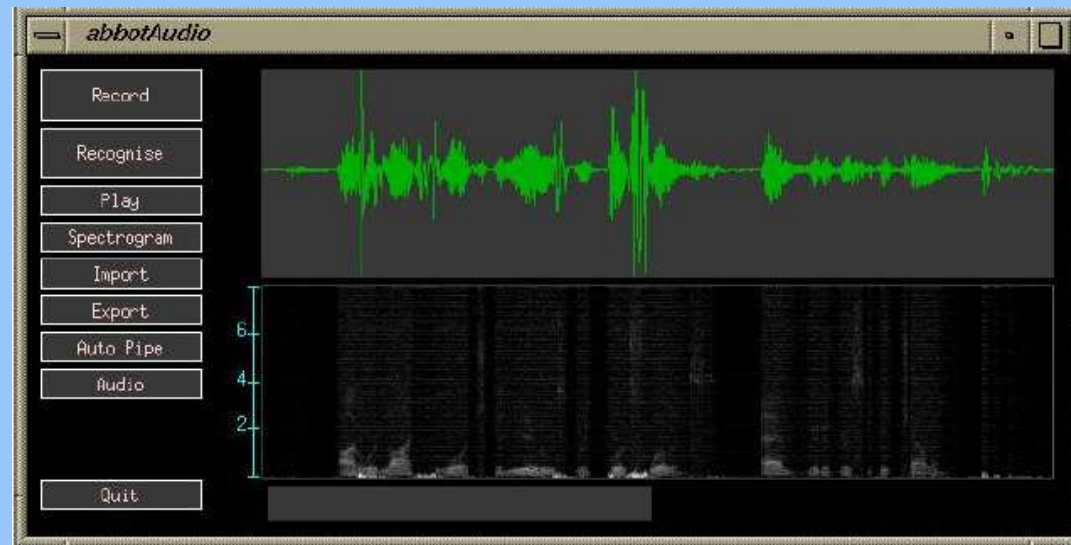
- Spracherkennung
- Parsing
- Dialog



The Virtual Interactive Presenter (VIP)

1 - Spracherkennung

- ABBOT system → Hybrides System
- VIP-ABBOT:
 - ◆ Tri-Gramm-Modell
 - ◆ Trainiert anhand von 200 Fragen und Antworten
 - ◆ 6 Sprecher (je 12 Min. Aufnahmezeit)



The Virtual Interactive Presenter (VIP)

1 - Spracherkennung

- Test
 - Korpus von 500 Äußerungen (4 Muttersprachler, 1 Nicht-Muttersprachler)
- Ergebnis
 - ca. 50 % der verarbeiteten Äußerungen enthalten mindestens einen Erkennungsfehler



The Virtual Interactive Presenter (VIP)

Aufbau des Programms

Die 3 wichtigsten Komponenten:

- Spracherkennung
- Parsing
- Dialog



The Virtual Interactive Presenter (VIP)

2 - Parsing

- Aufgrund der hohen Fehlerrate → kein komplettes Parsing
- Vereinfachter Parser: Tree-Adjoining Grammar (TAG)

Vorgehensweise TAG-Parser

- Generierung aller möglichen Syntaxbäume → „forest“
- Parsing Algorithmus „scannt“ von links nach rechts, vergleicht die Bäume paarweise und hängt, wenn möglich, die Bäume aneinander an
- Wird solange wiederholt, bis keine „Adjunktionen“ mehr möglich sind



The Virtual Interactive Presenter (VIP)

2 - Parsing

Spezieller VIP-Parser

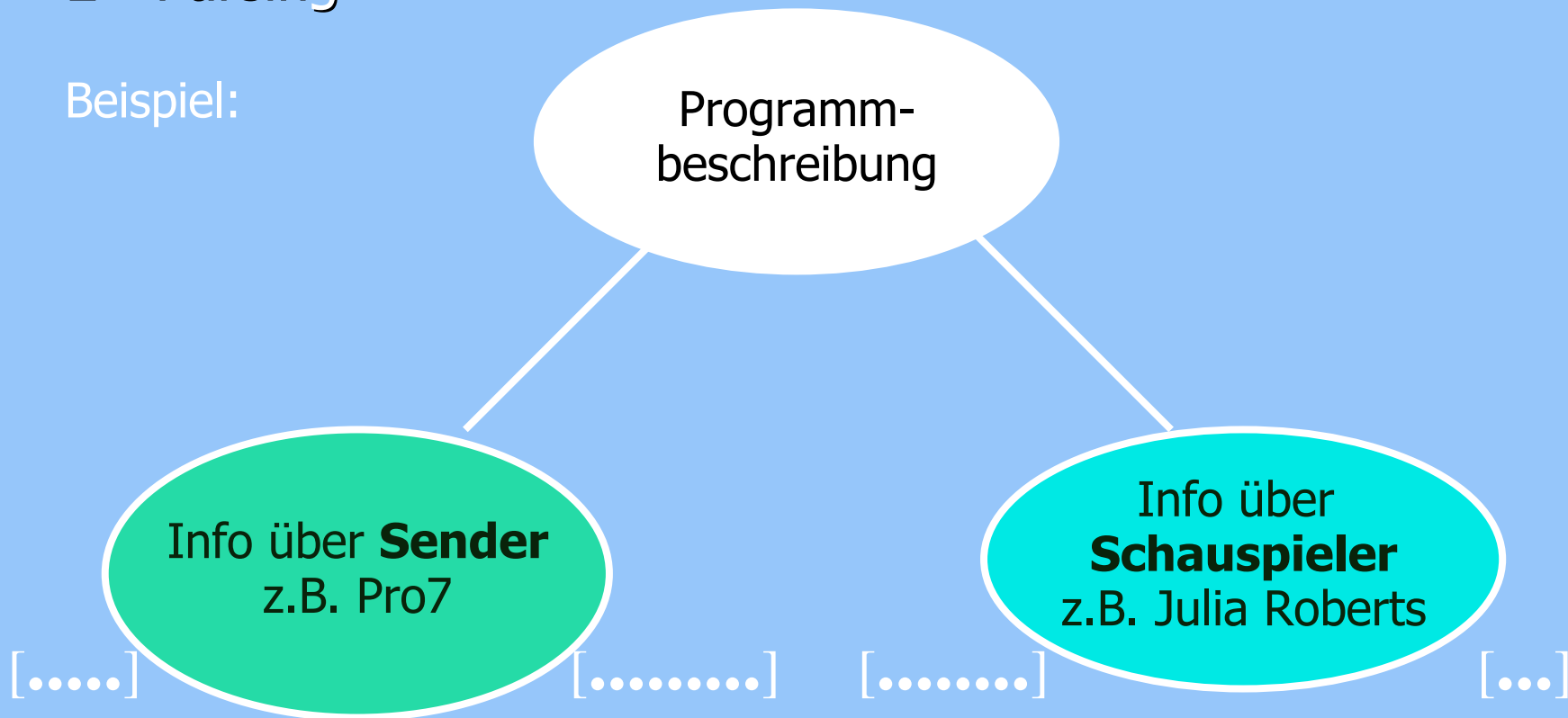
- Während des Parsens werden die semantischen Merkmale der erkannten Satzbausteine angehäuft
- Wenn es nicht gelingt, einen einzelne vollständige Aussage zu parsen, dann werden diejenigen Elemente partieller Parses zusammengefügt, die in einem semantischen Zusammenhang stehen



The Virtual Interactive Presenter (VIP)

2 - Parsing

Beispiel:



The Virtual Interactive Presenter (VIP)

2 - Parsing

- Es entsteht eine **Merkmalsstruktur**, deren Elemente mit den **internen Kategorien des Systems** verglichen werden können
[Kategorien: Genre, Schauspieler, Sender usw.]
- aus den semantischen Strukturen werden **Suchfilter** generiert, die aus der Programmdatenbank die passenden Filme herausfiltern



The Virtual Interactive Presenter (VIP)

2 - Parsing

- ermöglicht Integration von Semantik und Dialogstrategie
- kann relativ gut mit ungrammatischen Äußerungen umgehen

User: IS THERE ANYTHING FUNNY I CAN WATCH
Recognised: THE IS THERE ANYTHING FUNNY I CAN WATCH
Semantics: (((QUESTION) (EXIST) (PROGRAMME ((CHOICE+)) (DET))
((GOOD) (SUB_GENRE COMEDY)) ((AUDIENCE USER)) ((QUESTION))
((VIEW)))
Filter: ((SUB_GENRE COMEDY))
Speech Act: (INITIAL (SUB_GENRE COMEDY) SEARCH)

System: I found 4 programmes corresponding to that selection. What about: "Starting Out"?



The Virtual Interactive Presenter (VIP)

Aufbau des Programms

Die 3 wichtigsten Komponenten:

- Spracherkennung
- Parsing
- Dialog



The Virtual Interactive Presenter (VIP)

3 - Dialog

Aufgabe des Dialogs:

Fortschreitende Verfeinerung einer Programmbeschreibung



Sprechakttheorie: Jedem Sprechakt entspricht eine bestimmte Konstruktionsform, die in Zusammenhang mit einer kommunikativen Funktion steht.

Zurückweisung

Implizite
Zurückweisung

Spezifikation

USW.



The Virtual Interactive Presenter (VIP)

3 - Dialog

Inhaltsbasierter Ansatz: Vergleich der aktuellen semantischen Inhalte mit dem vorherigen Suchfilter.

Spezifikation

z.B. Suchfilter zuerst Genre dann Subgenre

Zurück-
weisung

Bestimmte Wörter als Marker, z.B. nein, nicht, langweilig, usw.

implizite
Zurück-
weisung

Semantische Inhalte der aktuellen Äußerung überschreiben diejenigen der vorherigen



The Virtual Interactive Presenter (VIP)

3 - Dialog

- Erkennungsfehler im Dialog
 - ✓ Fatale Erkennungsfehler
 - ✓ Benutzer korrigiert den Erkennungsfehler
 - ✓ System korrigiert den Erkennungsfehler
 - ✓ Semantische Robustheit
 - ✓ "Zufällige" Robustheit



The Virtual Interactive Presenter (VIP)

3 - Dialog

- Robuster Umgang mit Erkennungsfehlern im Dialog
 - Fatale Erkennungsfehler

U19: Can I have a movie?

S20: *I have found 22 programmes for that selection. I would recommend: "Midnight Fear"*

U21: **I want a movie my good can watch**

(REF: I want a movie my kids can watch)

S22: *Here is "Midnight Fear", then. Goodbye!*



The Virtual Interactive Presenter (VIP)

3 - Dialog

- Erkennungsfehler im Dialog
- Benutzer korrigiert den Erkennungsfehler

U33: Is there a movie my good can watch?
(REF: is there a movie my kids can watch?)
S34: *I have 22 programmes for your request.
Would you like to watch: "Midnight Fear"*
U35: What is its rating?
S36: *Its rating is caution*
U37: No, I want a movie my kids can watch
S38: *I would recommend this programme
corresponding to that "family" rating: "Rio
Bravo"*



The Virtual Interactive Presenter (VIP)

3 - Dialog

- Erkennungsfehler im Dialog
- System korrigiert den Erkennungsfehler

U39: What kind of sports boring you have?
(REF: what kind of sports programmes do you have)

S40: What would you like to see, then?

U41: A sports programme

S42: I could find 5 programmes for this choice. I can recommend: "Row Z"



The Virtual Interactive Presenter (VIP)

3 - Dialog

- Erkennungsfehler im Dialog
- System korrigiert den Erkennungsfehler

*S43: What about the following programme:
“Countryfile”*

**U44: Again want a documentary (REF: I don’t
want a documentary)**

S45: Is this programme all right, then?



The Virtual Interactive Presenter (VIP)

➤ Semantische Robustheit

Beispiel 1

U54: Do you have any documentaries

S55: *I have 2 programmes for your request. My first selection is: "Ground Force"*

U56: **An old gardening is not a little A** (REF: no, gardening is not OK)

S57: *I can recommend this programme, which is not a gardening: "Fangs"*

Beispiel 2

U50: what kind of movies do you have

S51: *I have found 22 programmes for this choice. I would suggest: "Twelve Monkeys"*

U52: **Who is playing the title all?** (REF: who is playing the title role?)

S53: *"Twelve Monkeys" features Bruce Willis*



3 - Dialog

Erkennungsfehler im Dialog

- Semantische Robustheit

Kombination aus einer relativ
oberflächlichen Verarbeitung und einer
natürlichen Redundanz in der Sprache
des Nutzers

The Virtual Interactive Presenter (VIP)

3 - Dialog

- Erkennungsfehler im Dialog
- Semantische Robustheit

Yes, this is fine.

Redundanz

No, I don't like that one.

No, gardening is not ok.



The Virtual Interactive Presenter (VIP)

3 - Dialog

- Erkennungsfehler im Dialog
- Semantische Robustheit

System:	<i>What about „Starting Out“?</i>
User:	NO I DON'T WANT A SITCOM.
Erkannt: SITCOM.	AN OLD I DON'T WANT A
System:	<i>What would you like to see then?</i>



The Virtual Interactive Presenter (VIP)

3 -

• Erkennungsfehler im Dialog

Dialog

➤ Zufällige Robustheit

U58: What movies do you have?

S59: *I found 22 programmes for this choice. I can suggest: "Salvador"*

U60: **Other one like is this to there it is** (REF: I don't want my kids to see this)

S61: *I have 21 other programmes for that selection. What about the following programme: "Casablanca"*

