

Unterrichtseinheit Sprachdialogsysteme

Lehrerfortbildung I*Teach

29. November 2006

Inhaltlicher Ablauf

- Sachanalyse
- mögliche Einordnung in den Lehrplan
- Didaktische Analyse
- Ziele der Unterrichtseinheit
- Vorstellung der Unterrichtseinheit
mit integriertem Miniteaching

- **Idee**

 - Thematisierung in Form eines Projekt des Schnupperstudiums 2006

- **Mitwirkende**

 - Prof. Dr. Wolfgang Menzel

 - Dr. Cristina Vertan

 - Fabian Greuel (stud. Hilfskraft)

 - Sabrina Hilger (stud. Hilfskraft)

- **Weiterentwicklung**

 - Projekt → Unterrichtseinheit

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen

Hören

Kikala brint tovoluti?

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen



Winter kochtest ganz Blatt?

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen



Wann steigt die Party?

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen



Wann steigt die Party?

Sintu högafi notsi!

Quelle: Wolfgang Menzel, Schnupperstudium 2006

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen



Wann steigt die Party?

Sonderbar werfen die Wellen hinab!

Quelle: Wolfgang Menzel, Schnupperstudium 2006

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen



Wann steigt die Party?

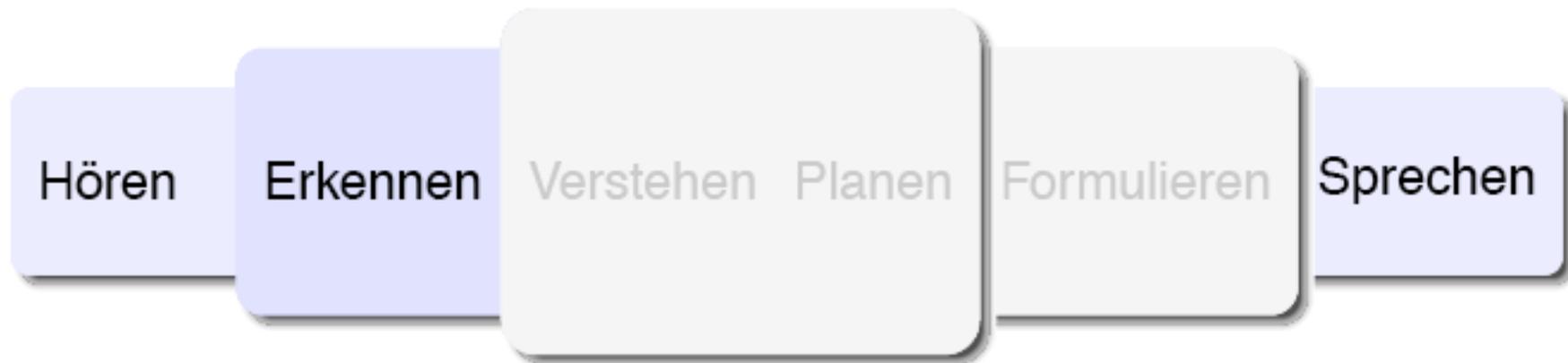
Am Freitag, im Hirsch!

Quelle: Wolfgang Menzel, Schnupperstudium 2006

Fachliche Klärung des Unterrichtsgegenstandes

Motivation

Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen

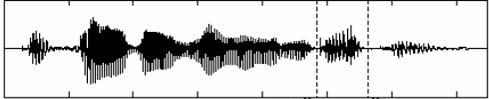


Wann steigt die Party?

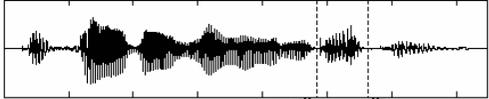
Am Freitag, im Hirsch!

Quelle: Wolfgang Menzel, Schnupperstudium 2006

Fachliche Klärung des Unterrichtsgegenstandes

Sprechen	Audio-Karte Boxen oder Kopfhörer
Formulieren	Vollsynthese: Ton- und Rauschgeneratoren erzeugen Sprachsignal Reproduzierende Synthese: menschliche Sprache wird aufgenommen und wiedergegeben
Planen	Aufbau der Basiselemente: Laute und ganze Phrasen
Hören	Mikrophon Audio-Karte
Erkennen	Frequenzwellen 
Verstehen	Spracherkenner: benutzt Wörterbuch und Dialogmodell

Fachliche Klärung des Unterrichtsgegenstandes

Sprechen	Audio-Karte Boxen oder Kopfhörer
Formulieren	Vollsynthese: Ton- und Rauschgeneratoren erzeugen Sprachsignal Reproduzierende Synthese: menschliche Sprache wird aufgenommen und wiedergegeben
Planen	Aufbau der Basiselemente: Laute und ganze Phrasen
Hören	Mikrophon Audio-Karte
Erkennen	Frequenzwellen 
Verstehen	Spracherkenner: benutzt Wörterbuch und Dialogmodell

Dialogmodellierung

- dynamische Einschränkung des Erkennerswortschatzes in Abhängigkeit vom Dialogzustand
- Wozu braucht man das?
 - Erkennungssicherheit erhöhen
 - ähnliche Aussprache: *Mai* oder *drei*?
 - verschiedene Sprecher
 - schlechte Übertragungsqualität
 - semantische Interpretation erleichtern
 - Was wird von der Maschine erwartet?
 - Zahl: Geldbetrag, Uhrzeit, Datum, Kontonummer, ...
 - Mehrdeutigkeit: *Wann werden S/sie kommen?*
 - Referenzauflösung: Was wird durch *sie* bezeichnet?
- eine Maschine hat keinen gesunden Menschenverstand!

Dialogmodellierung mit Automaten

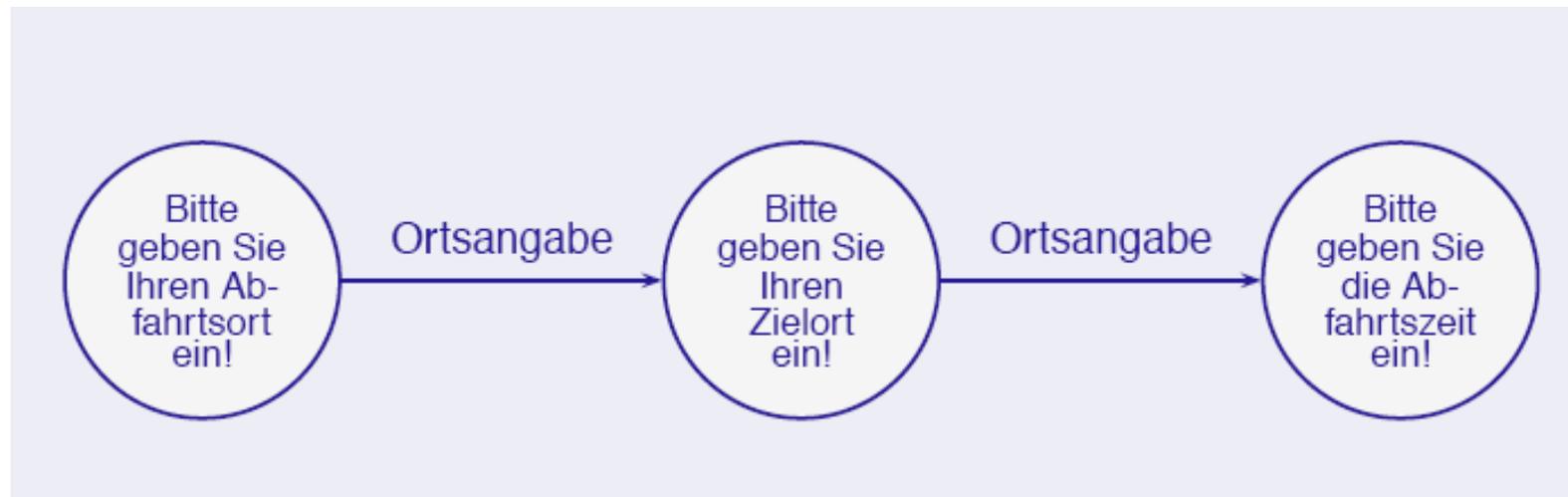
- theoretische Grundlage: deterministischer endlicher Automat
- einfachstes Automatenmodell der Informatik
- effiziente Implementierung
- gute Vorhersagefähigkeit: starke Einschränkung des aktiven Wortschatzes



Quelle: Wolfgang Menzel, Schnupperstudium 2006

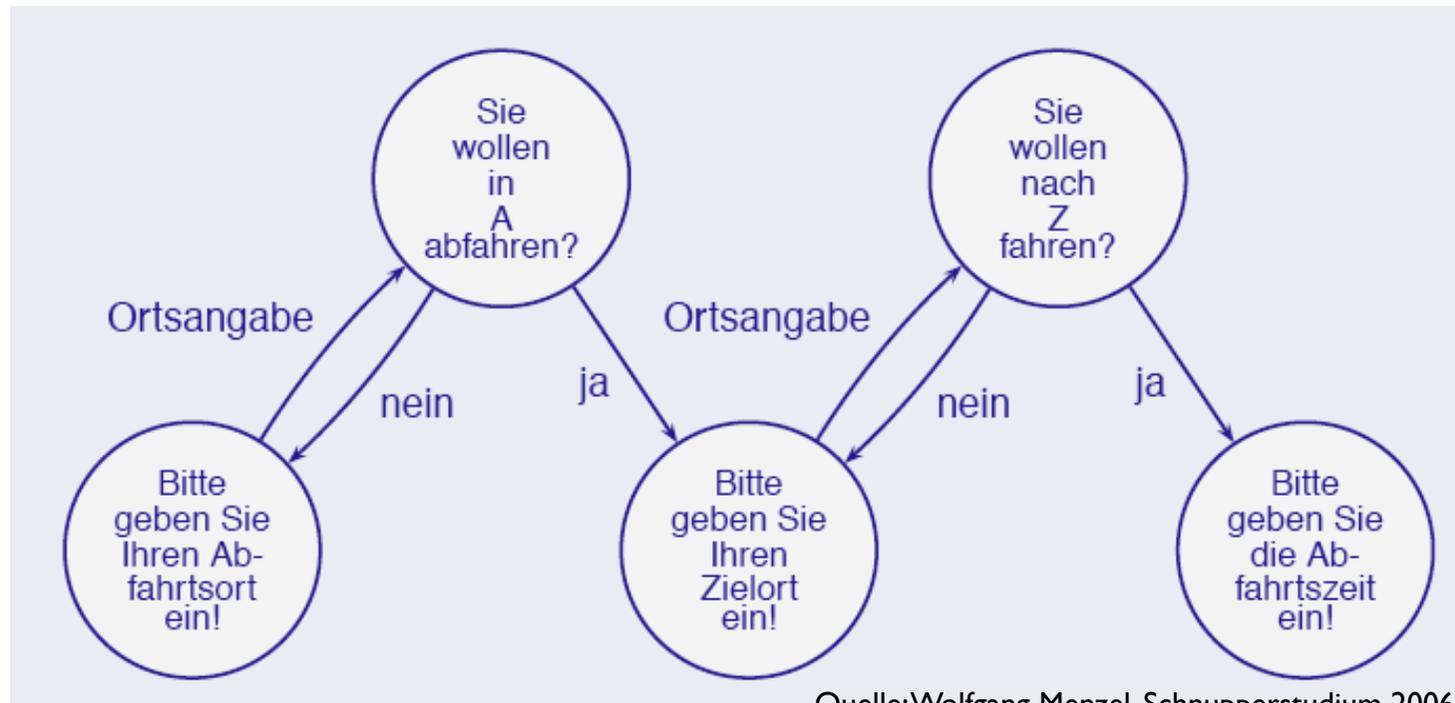
Dialogmodellierung mit Automaten

- theoretische Grundlage: deterministischer endlicher Automat
- einfachstes Automatenmodell der Informatik
- effiziente Implementierung
- gute Vorhersagefähigkeit: starke Einschränkung des aktiven Wortschatzes



Dialogmodellierung mit Automaten

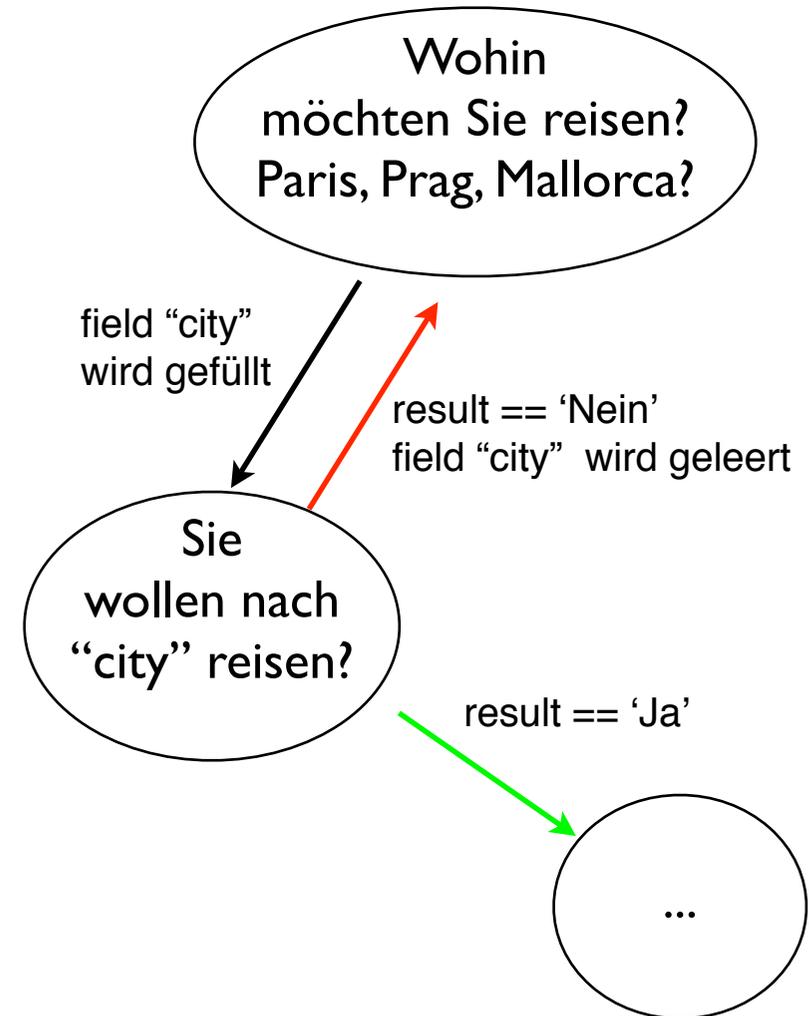
- theoretische Grundlage: deterministischer endlicher Automat
- einfachstes Automatenmodell der Informatik
- effiziente Implementierung
- gute Vorhersagefähigkeit: starke Einschränkung des aktiven Wortschatzes



Quelle: Wolfgang Menzel, Schnupperstudium 2006

Dialogmodellierung mit VXML

```
<?xml version="1.0" encoding="iso-8859-1"?>
<vxml version="2.0" lang="de">
  <form>
    <field name="city">
      <prompt>
        Wohin möchten Sie reisen?
      </prompt>
      <option> Paris </option>
      <option> Prag </option>
      <option> Mallorca </option>
    </field>
    <field name="result">
      <prompt>
        Sie wollen nach <value expr="city"/>
        reisen?
      </prompt>
    </field>
    <filled>
      <if cond="result=='Nein'">
        <clear namelist="city"/>
      </if>
    </filled>
    ...
  </form>
</vxml>
```



Rahmenplan Informatik - Gymnasiale Oberstufe - Hamburg

	Leitlinie	Interaktion mit Informatiksystemen	Wirkprinzipien von Informatiksystemen	Informatische Modellierung	Wechselwirkungen zwischen Informatiksystemen, Individuum und Gesellschaft
S4 Projekt	Roboter-systeme	Visuelle Programmierung, Imperative Programmiersprache, z. B. NQC	Endlicher Automat, parallele Prozesse	Zustandsorientierte Modellierung	Veränderung von Arbeitsplätzen, Veränderung von Gesellschaft
S4 Projekt	Sprachdialog-systeme	Markup-Sprache VXML	Endlicher Automat	Zustandsorientierte Modellierung	Möglichkeiten und Grenzen der Kommunikationsmöglichkeiten

Begründung des Unterrichtsvorhabens

immer häufigere Konfrontation (Auto, Telefonauskünfte, ...)

⇒ **Blick hinter die Kulissen - wie funktioniert das überhaupt?**

Unterschiede zwischen MMK und MCK:

- Wie mächtig ist die MMK überhaupt?
- Wie gut sind die Systeme, die mir begegnen?

⇒ **eigene Kommunikation bewusst hinterfragen**

⇒ **Fenster zur Forschung**

Grobgliederung der Unterrichtsreihe

Phase 1: Vergleich MMK und MCK mit Sprache

- Aufbau der sprachlichen Kommunikation
- Kennenlernen der Darstellung eines Dialoges mit Automaten

Phase 2: Experimentieren mit existierendem Dialogsystem

- Kennenlernen der Software-Umgebung

Benutzen

- Einführung in die Sprache VXML

Analysieren

Grobgliederung der Unterrichtsreihe

Phase 3: Implementierung und Evaluierung eines Dialogsystems

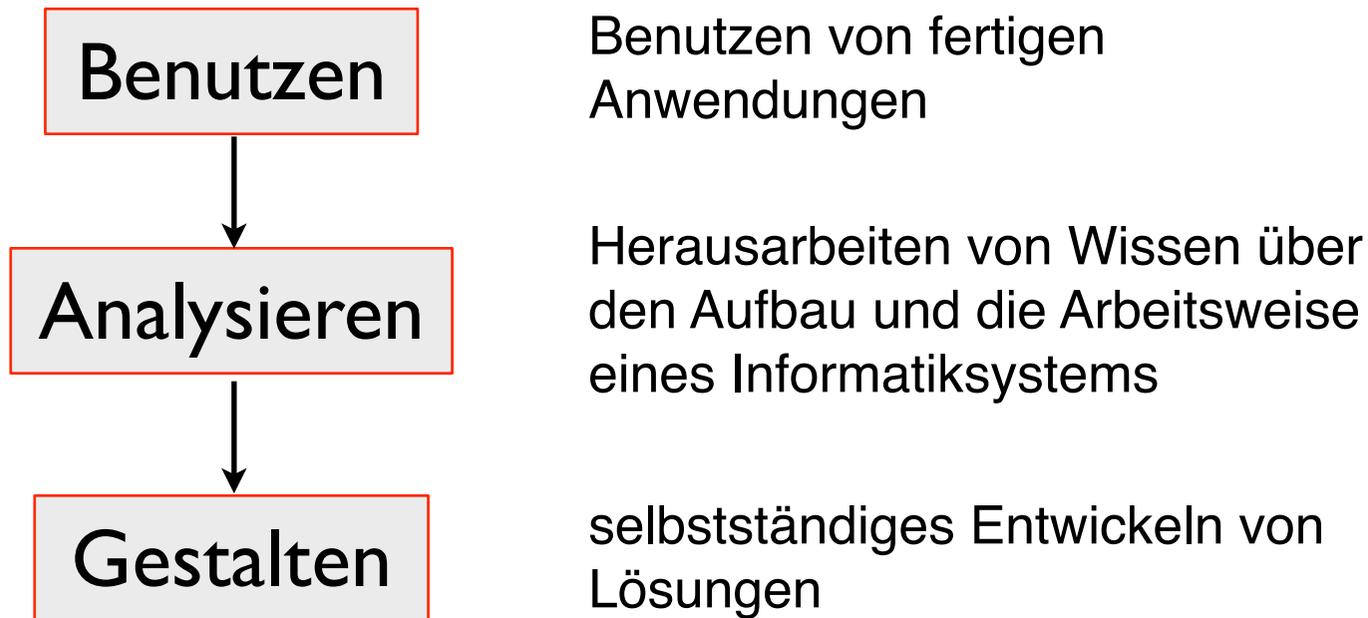
- Ideenskizzierung
 - Programmieren in VXML
 - gegseitiges Evaluieren
- } Verzahnung: Verstehen und Tun

Gestalten

Phase 4: Präsentation

- Vorstellung des eigenen Dialoges

Der didaktische 3-Schritt (nach Norbert Breier)



Fachspezifisches Ziele des Informatikunterrichts

- Kenntnisse von MCK
- Verständnis der grundlegenden Bedeutung von Automatenmodellen für die Informatik
- Modellbildung und Umsetzung mit geeigneter Sprache

Hauptziel der Unterrichtseinheit

- Einblick in die Funktionsweise und den Aufbau eines Sprachdialogsystems
- zustandsorientierte Modellierung und Programmierung

Aufbau der Unterrichtseinheit

Insgesamt sind 14 Unterrichtsstunden für diese Einheit eingeplant.
Sie besteht aus sechs Tasks.

Task 1:

- Dialogmodellierung mit deterministisch endl. Automaten **2:14**

Task 2:

- Experimentieren mit einem existierenden Dialogsystem **2:14**

Task 3:

- Ideenskizzierung eines geeigneten Dialoges **1:14**

Task 4:

- Implementierung des eigenen Dialoges **6:14**

Task 5:

- Evaluierung und Verbesserung **1:14**

Task 6:

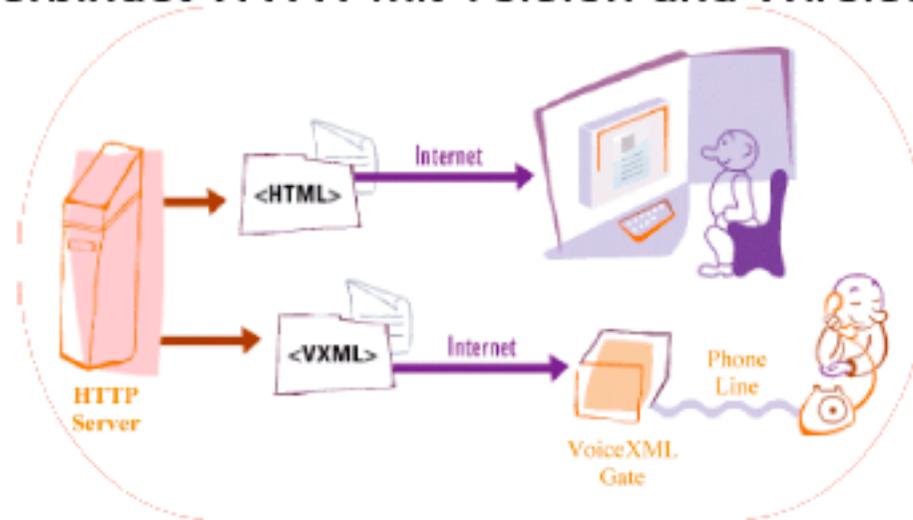
- Aufbau einer Präsentation **2:14**

Inhalte

Phase	Inhalt	Form	Medien
Einstieg	Kommunikation mit gesprochener Sprache ist mehr als nur Hören und Sprechen	LV	PPT
Vertiefung	Wie modelliert man einen Dialog?	LG	PPT
Übung	Modellierung von Dialogen, die als Audio vorliegen	PA	Audio-Files, Audioplayer, Kopfhörer, AB1
Zusammenfassung	Was bleibt unklar?	UG	

Was ist Voice-XML?

- Voice-XML
 - Ist eine XML-basierte Dialog-Markup-Sprache
 - Sie verbindet WWW mit Telefon und Wireless-Geräten



- Wird (normalerweise) von einem Voice-Browser statt einem GUI-Browser, z.B. Explorer, ausgeführt
- Voice-XML-Anwendungen ermöglichen:
 - Benutzereingabe durch Sprachsignale oder Telefontasten
 - Sprachausgabe durch Sprachsynthese oder Audiodateien

VoiceXML Elemente /Tags

- `<vxml>` erstes Element in jedem VXML-Dokument
- `<form>` leitet einen Dialog (zur Informationsdarstellung oder Datensammlung) ein
- `<block>` enthält nicht-interaktiven (=executable) Code
- `<prompt>` markiert Audio-Ausgaben (synthetisiert oder aus Datei)
- `<field>` markiert ein Eingabefeld in einem `<form>`
- `<filled>` nennt in einem Feld die Aktion, die nach Füllen des einbettenden Feldes startet
- `<menu>` ein in `<form>` eingebetteter Dialog zur Auswahl aus Alternativen
- `<choice>` definiert je ein Menu-Item
- `<grammar>` spezifiziert die gültige Grammatik
- `<goto>` Verbindung zu einem anderen Dialog in demselben oder einem anderen Dokument

Versionsangaben zu VXML

Kommentar

block erwartet keine Aktion -
hier ist die Begrüßung damit realisiert

enumerate bewirkt, dass die
nachfolgenden choices aufgezählt werden

field wird über den Namen angesprochen

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<vxml xmlns="http://www.w3.org/2001/vxml"
version="2.0">
  <!--Hello World Program -->
  <form>
    <block>
      <prompt> Guten Tag!
    </prompt>
    </block>
    <field name="abteilung">
      <prompt>Mit welcher Abteilung darf ich Sie
        verbinden?<enumerate/>
      </prompt>
      <choice next="abtA">Abteilung A</choice>
      <choice next="abtB" >Abteilung B</choice>
      <prompt>Ich verbinde Sie mit <value expr="abteilung"/>.
        Auf Wiederhören.
    </prompt>
    </field>
  </form>
</vxml>
```

Wer hat die Kontrolle? \Rightarrow das System

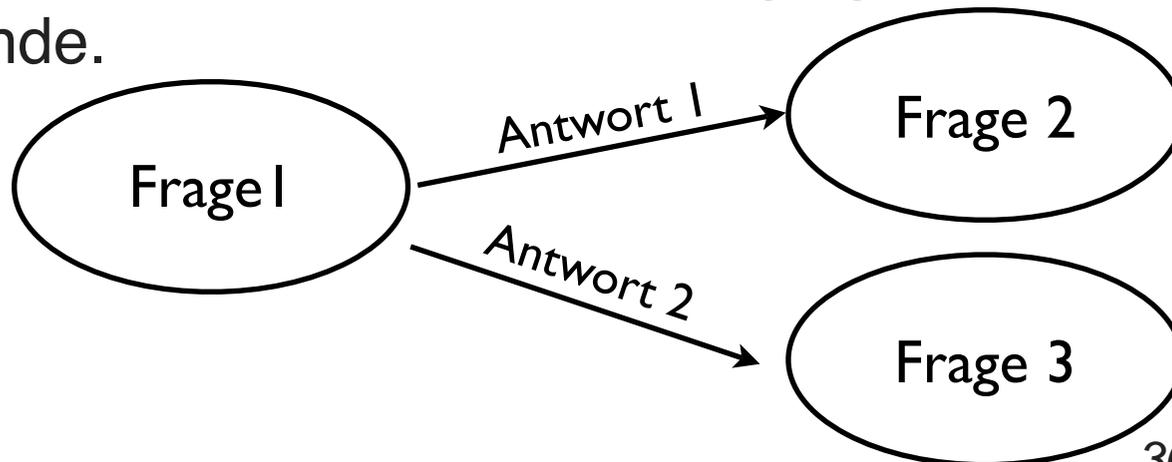
Was macht das System denn? \Rightarrow Fragen stellen, Auswahl bieten



Was macht der Benutzer? \Rightarrow antworten



Danach wird die nächste Frage gestellt, oder der Dialog ist zu Ende.



Aufgabenblatt 1

Gehe auf www.voice-community.de.

Unter “Best Practice” findest Du Hör-Beispiel.

Modelliere den Dialog “1.FC Köln - Poldis Gewinnspiel”.

im Vortrag an dieser Stelle ergänzt:

Moodle

<http://nats-www.informatik.uni-hamburg.de:8080/iteach/moodle/index.php>

NATS-Schnupperstudium

<http://nats-www.informatik.uni-hamburg.de/>

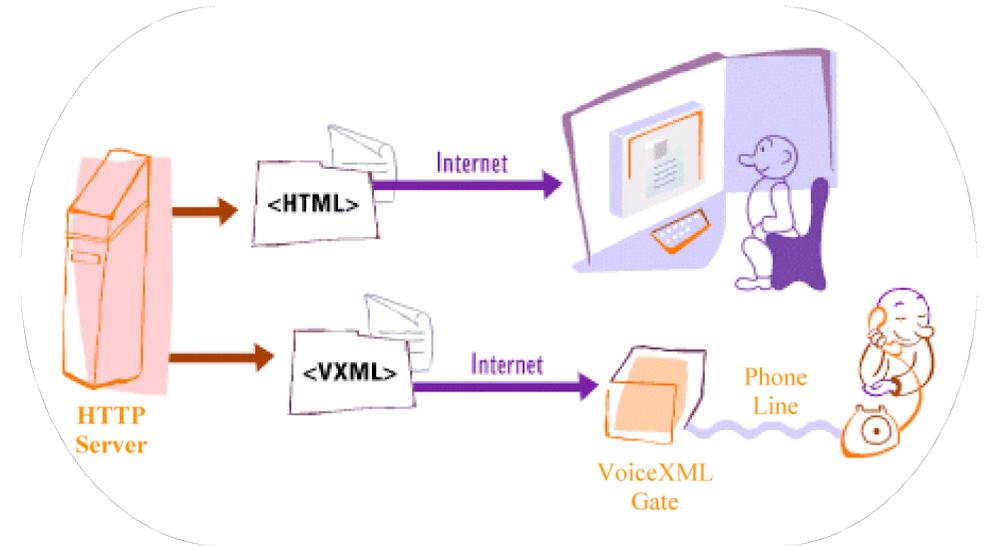
Didaktischer Hintergrund

Vertrautwerden mit dem Aufbau eines Dialoges
Begriffe “endlich” und “deterministisch” wurden selbst
erarbeitet (Information skills)
greifbare Vorstellung über das Thema

Inhalte

Phase	Inhalt	Form	Medien
Erforschung	Ausführen von waehrungsumrechner.vxml über Internet-Telefonie, Kennenlernen von Skype	PA	Skype, Headset
Besprechung	Wie greifen wir auf die Datei zu? Wo liegt sie?	LG	PPT
Einrichtung	Accounteinrichtung bei Voxbuilder, Anlegen des Projektes "Währungsumrechner"	PA	Wiki, AB2 ProjektGuide, Voxbuilder
Vertiefung	Arbeiten mit dem Quelltext des Währungsumrechners	PA	Wiki, AB2 ProjektGuide, Skype, Voxbuilder
Zusammenfassung	Was bleibt unklar?	UG	

Didaktischer Hintergrund



schrittweises Kennenlernen der benötigten Software

Wiederholung der Realisation von Sprachdialogsystemen
und damit Verständnis über den Sinn von Voxbuilder

Einführung in VXML über den 3-Schritt

Aufgabenblatt 2

Rufe mit Skype bei Voxbuilder an.

ID: 49 29 00 34

PIN: 0400

Inhalte

Phase	Inhalt	Form	Medien
Finden	Finden und Begründen eines geeigneten Dialoges	TA	AB3
Formulieren	Erstellung eines Dialogmodells	TA	AB3
Problemlösen	Bestimmung der Reihenfolge der zu implementierenden Dialogabschnitte	TA	AB3
Rücksprache	Ist das so umsetzbar?	UG	

Didaktischer Hintergrund

Grundelemente sind bekannt

Verstehen des Zusammenhangs zwischen VXML und dem Dialogmodell:

- Kreis als Frage/Auswahl entspricht field
- Antwortmöglichkeiten entsprechen option

Inhalte

Phase	Inhalt	Form	Medien
Programmieren	Ausprogrammierung des Dialoges	TA	Skype, Voxbuilder, Headset, Wiki
Informations- findung	Suchen nach geeigneten VXML-Tags	TA	Wiki, Internet
Problemlösen	Fehlersuche	TA	Voxbuilder
Besprechung	Besprechung häufig zu beobachtender Fehler	UG	

Inhalte

Phase	Inhalt	Form	Medien
Evaluierung	Testen und Beurteilen eines anderen Dialogsystems	TA	Skype, AB4
Feedback geben, Feedback nehmen	Besprechung von Unzulänglichkeiten in den einzelnen Dialogen	UG	AB4
Feedback nutzen	Verbesserung des eigenen Dialoges	TA	Skype, Voxbuilder, AB4, Headset

Aufdecken der Unzulänglichkeiten von Voxbuilder

- erkennt keine Zahlen
- Spracherkenner nicht gut trainiert

Aufdecken der Unzulänglichkeiten des Dialoges:

- Waren die Sprachprompts eindeutig?
- Ist der Dialog benutzerfreundlich gestaltet worden?
- Musste die Eingabe noch einmal bestätigt werden?

Inhalte

Phase	Inhalt	Form	Medien
Besprechung	möglicher Aufbau einer Präsentation	LG	
Übung	Strukturierung und Aufbau einer Präsentation	TA	AB5, PowerPoint o.ä.

Inhalte

Phase	Inhalt	Form	Medien
Vortrag	Vorstellung des aufgebauten Dialoges	SV	Beamer, PowerPoint

im Vortrag an dieser Stelle ergänzt:
www.voxbuilder.com