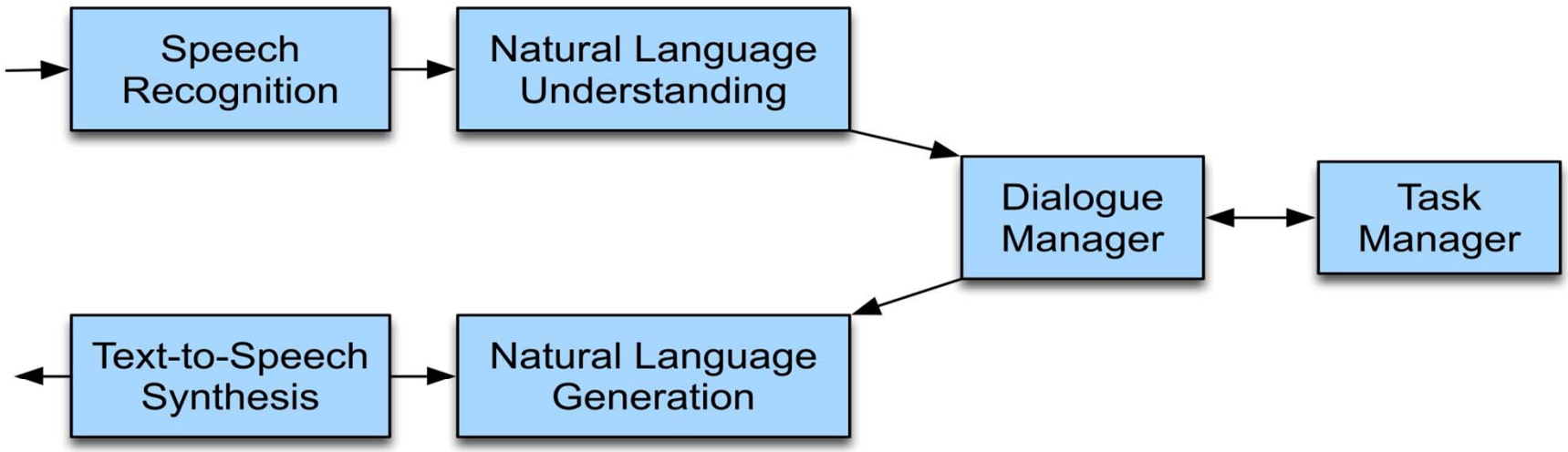
Dialog Systems

Section 24.1 – 24.2 in Textbook

hannes@ru.is

T-725-MALV

Basic Dialogue System



Speech Recognition

- **Optimized** for use in conversational agents
  - Hand-written finite-state or context free grammar
  - Dialogue-state dependent → Restrictive grammar
- Needs to provide **real-time** response
- May provide **confidence level** for agent to decide whether to ask human for confirmation

Natural Language Understanding

- Produces **semantic representation**
- Many based on **frame-and-slot filling**

"Show me morning flights from Boston to SF on Tuesday"

SHOW:
    FLIGHTS:
        ORIGIN:
            CITY: Boston
            DATE: Tuesday
            TIME: morning
        DEST:
            CITY: San Francisco

Natural Language Understanding

- A variety of semantic analysis methods used
- Simple domain-specific **semantic grammars** can fill slots directly

LIST FLIGHTS ORIGIN
Show me flights from Boston
DESTINATION DEPARTDATE
to San Francisco on Tuesday
DEPARTTIME
morning

Natural Language Understanding

- **Semantic Grammars**: CFG in which the LHS of rules is a semantic category

SHOW -> show me | I want | can I see|…

DEPARTTIME -> (after|around|before) HOUR
| morning | afternoon | evening

HOUR -> one|two|three…|twelve (am|pm)

FLIGHTS -> (a) flight|flights

ORIGIN -> from CITY

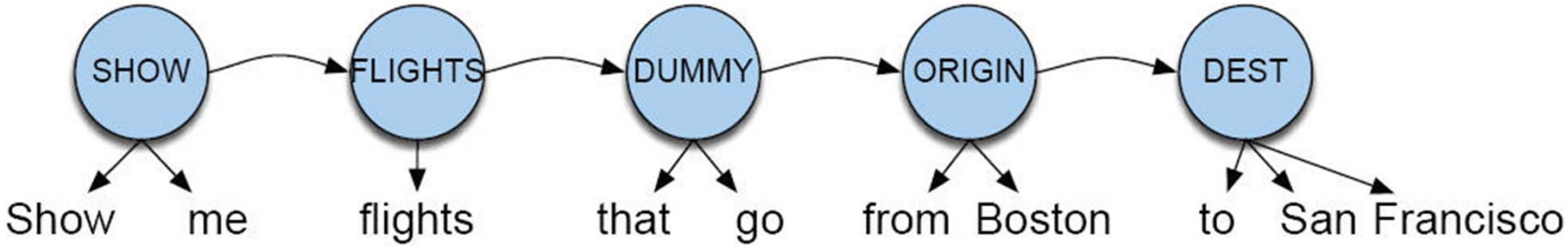
DEST -> to CITY

CITY -> Boston | San Francisco | Denver | Washington

Natural Language Understanding

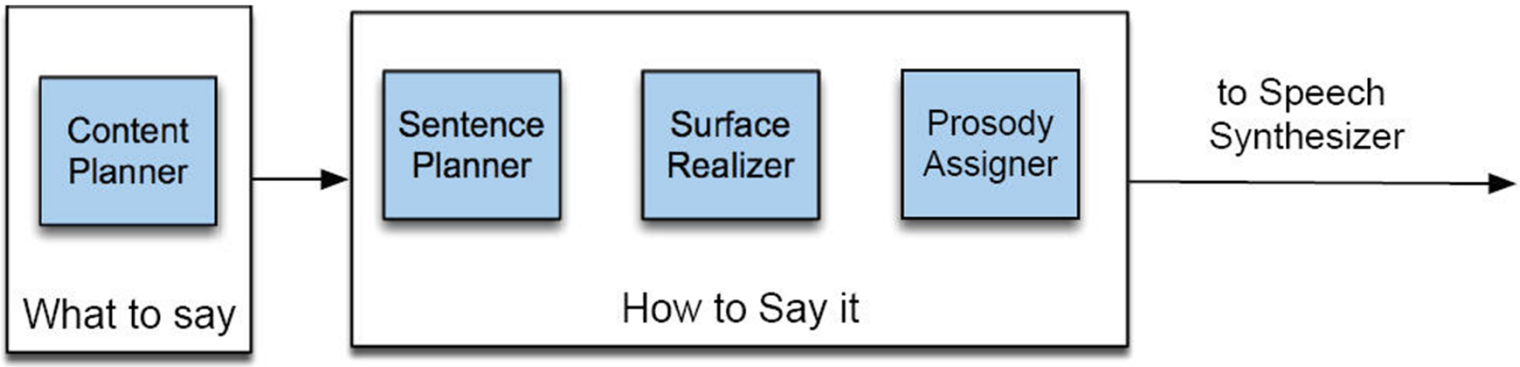
- **Problems** with **semantic grammars**:
  - Hand-coded (lot of manual work)
  - Deal poorly with ambiguity

- Alternatively use **semantic HMMs**



Natural Language Generation

- NLU chooses syntactic structures and words needed to **express a particular meaning**
- Two general approaches:
  - **Template-based** generation
  - Full **natural language generator**



full natural language generator

Text To Speech

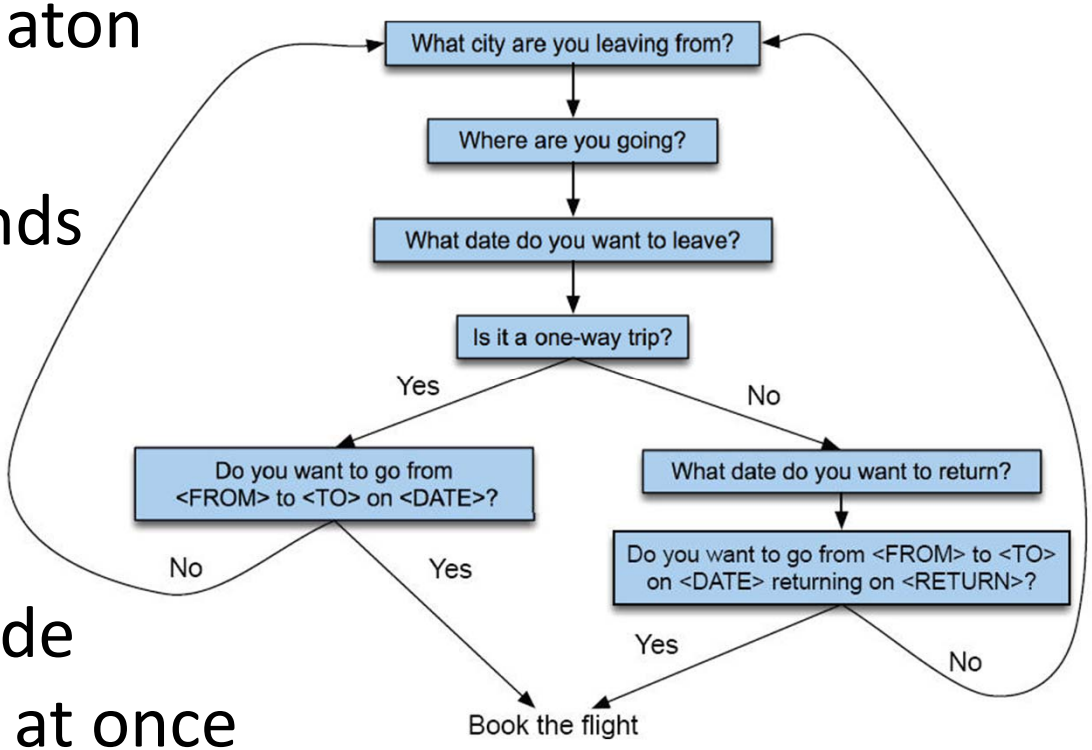
- Takes words and prosodic annotations
- **Synthesizes** a waveform typically using one of two methods:
  - Through **parametric control** of wave generators
  - Through **concatenation** of recorded speech samples

Dialogue Manager

- Four typical approaches:
  - **Finite-state** (system-initiative)
  - **Frame-based** (mixed-initiative)
  - Information-state based
  - Plan-based

Dialogue Manager

- **System-initative finite-state**
  - Completely controls the conversation through a finite-state automaton
  - May allow a few universal commands
  - **Problem:**
    User cannot provide more information at once



Dialogue Manager

- **Mixed-initiative frame-based**
  - Asks questions to **fill slots** in a frame (form)
  - User may also volunteer information for other open slots
  - Picking the right frame may require rules

Slot	Question
ORIGIN CITY	"From what city are you leaving?"
DESTINATION CITY	"Where are you going?"
DEPARTURE TIME	"When would you like to leave?"
ARRIVAL TIME	"When do you want to arrive?"

Questions associated with certain slots

Dialogue Manager

- **Grounding** in a dialogue system is typically done through either explicit or implicit confirmation
  - **Explicit:**
    U "Paris." C "You want to go to Paris?"
  - **Implicit:**
    U "Paris." C "When do you want to go to Paris?
- Also possible to completely **reject** and first give rapid **reprompting** and then add detail