

# Homework 3: Dialog

## ◆ *Part 1*

- *Call TellMe and get two sets of driving directions*
- *Call CMU's Let's Go*
- *Call Amtrak*

## ◆ *Part 2*

- *Build your own pizza ordering systems*
  - *Register with Tell Me Studio*
  - *Use VoiceXML to build a system*
- ◆ *Results are due 17<sup>th</sup> November 3:30pm*



# Speech Processing 15-492/18-492

---

Spoken Dialog Systems

Beyond VoiceXML:

the *Olympus Spoken Dialog Framework*

# Spoken Dialog - VoiceXML

- ◆ *Write (several) vxml “pages” and resources*
  - *Your dialog application control*
  - *Provide grammar for understanding*
  - *Define what your system says*
- ◆ *Generally just use provided ASR/TTS*
- ◆ *Great for basic form-filling applications*
  - *What if your application can't be made into a form-filling one?*

# Olympus Spoken Dialog Framework

- ◆ *A general dialog system architecture*
- ◆ *Modular, open source framework*
  - *Provides components needed to build SDS*
    - ⊗ *ASR/TTS, Language Understanding/Generation, Dialog Management, etc.*
  - *Can replace components with other options*
    - ⊗ *e.g., use a different ASR engine*
  - *Tied together via Galaxy message-passing communication infrastructure*
- ◆ *<http://wiki.speech.cs.cmu.edu/olympus>*

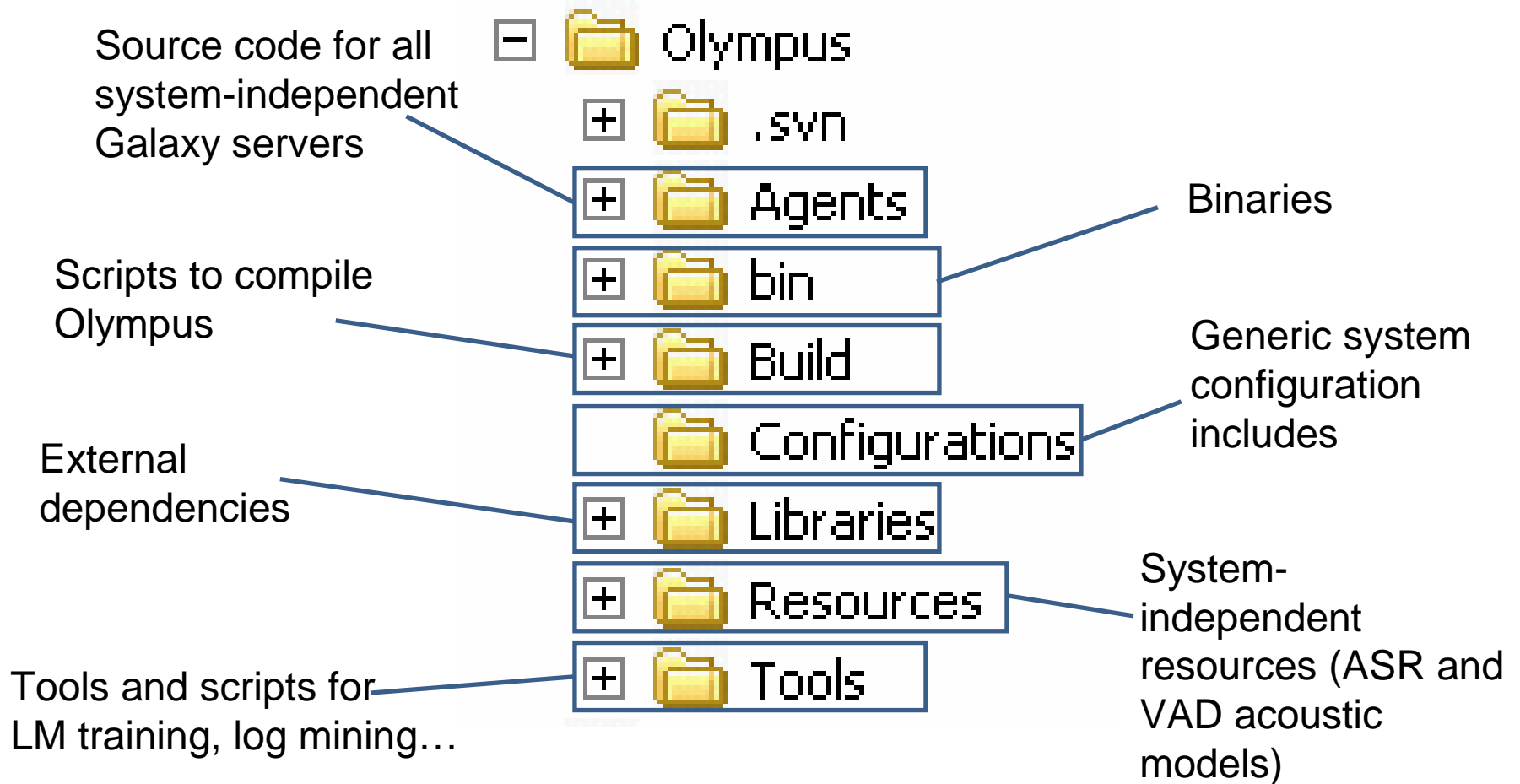
# Example Olympus Systems

- ◆ *Let's Go! (bus information)*
- ◆ *TeamTalk (robot interaction)*
  - *<http://wiki.speech.cs.cmu.edu/teamtalk/>*
- ◆ *Vera*
  - *<http://www.speech.cs.cmu.edu/~awb/vera.wmv>*
- ◆ *Many others*

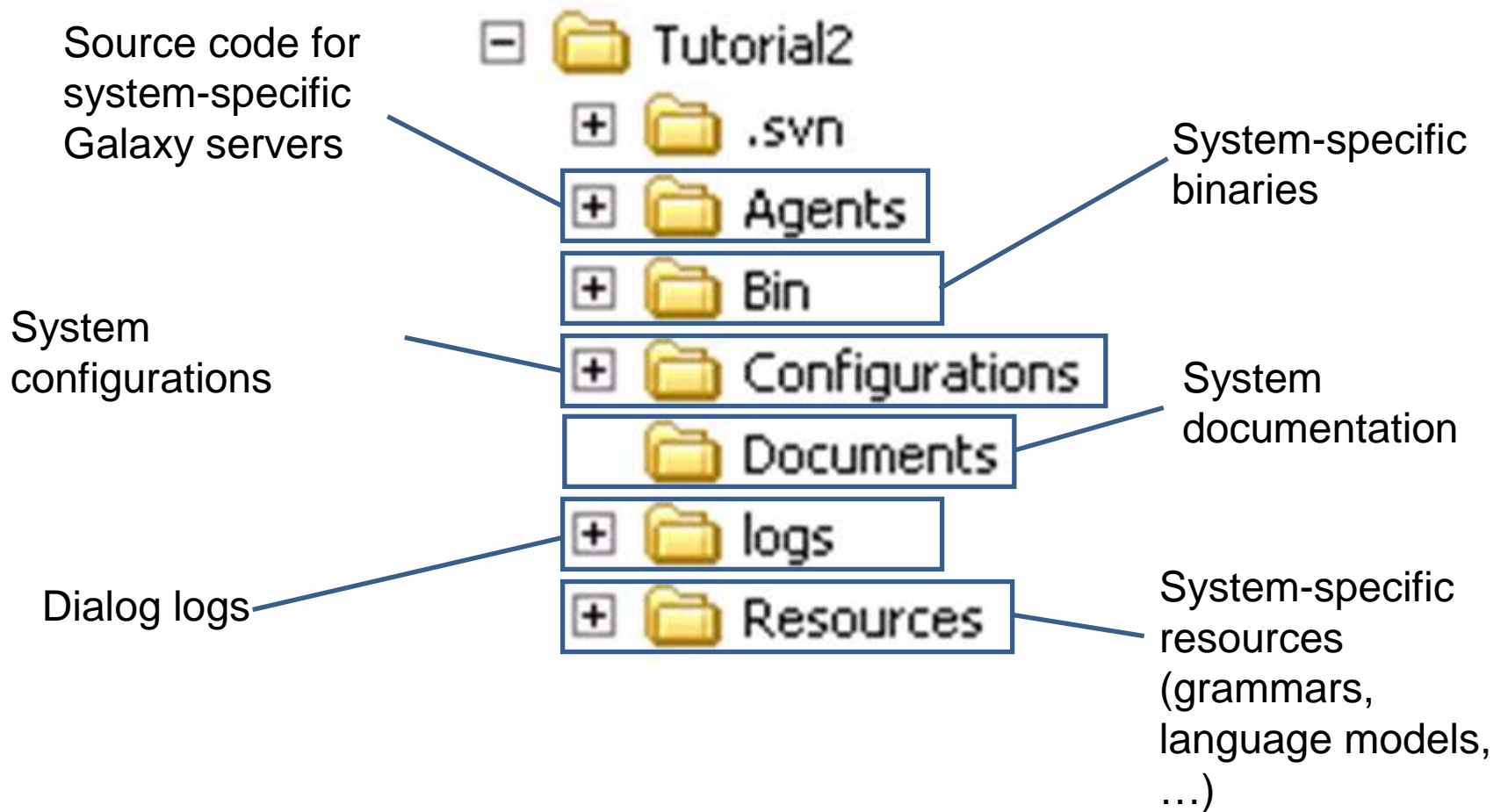
# Organization of Olympus Systems

- ◆ *Core components*
  - *Generic, useful in multiple different systems*
- ◆ *Application components*
  - *System-specific, useful for a single application*

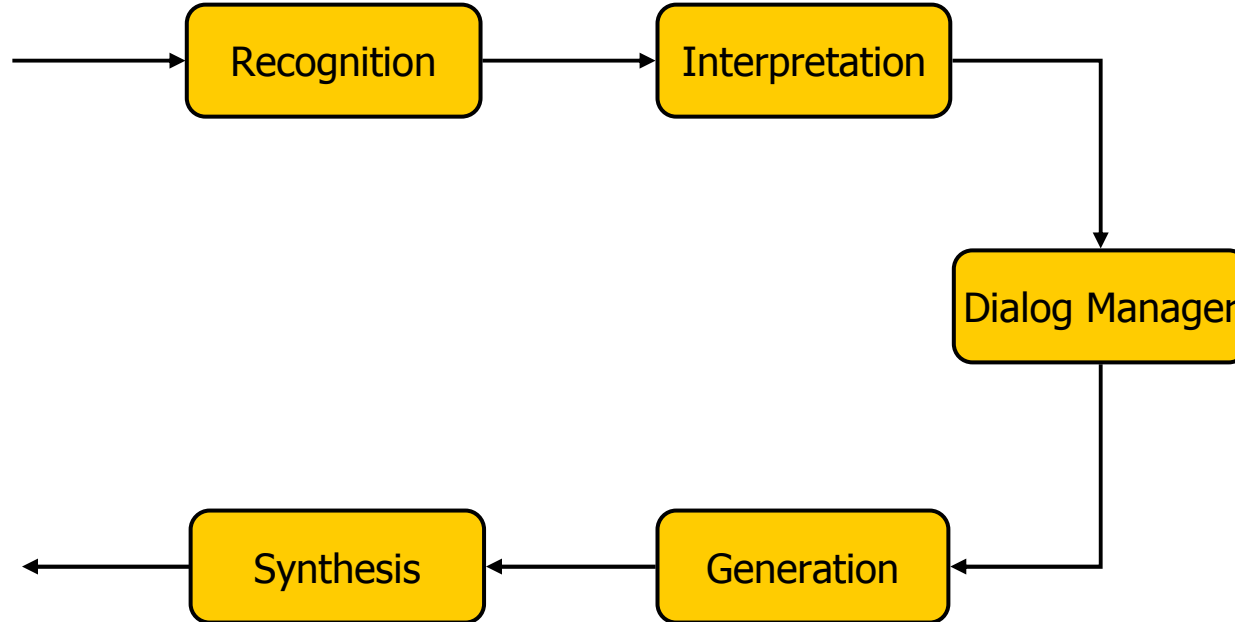
# Olympus Core Directory Structure



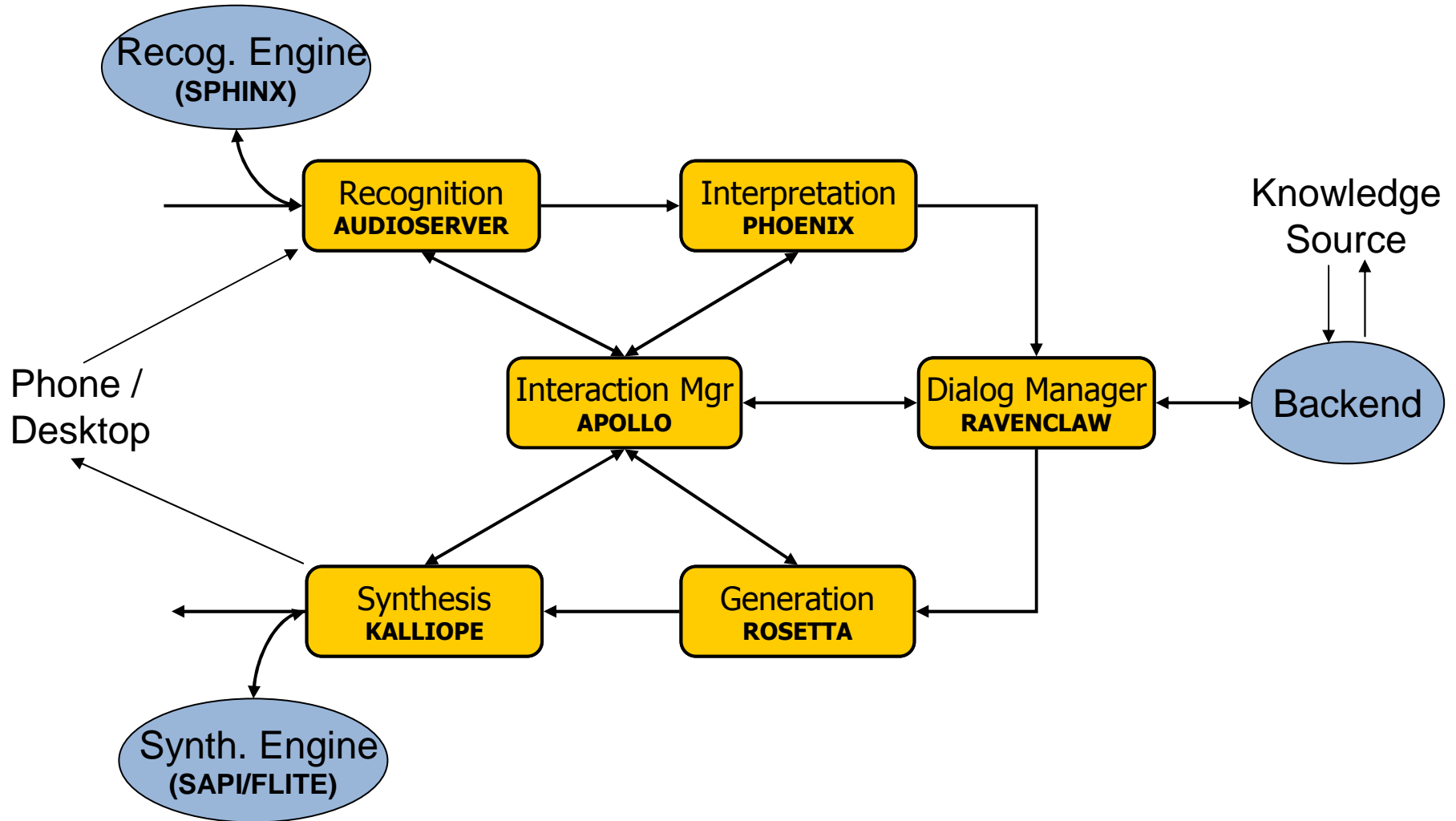
# System Directory Structure



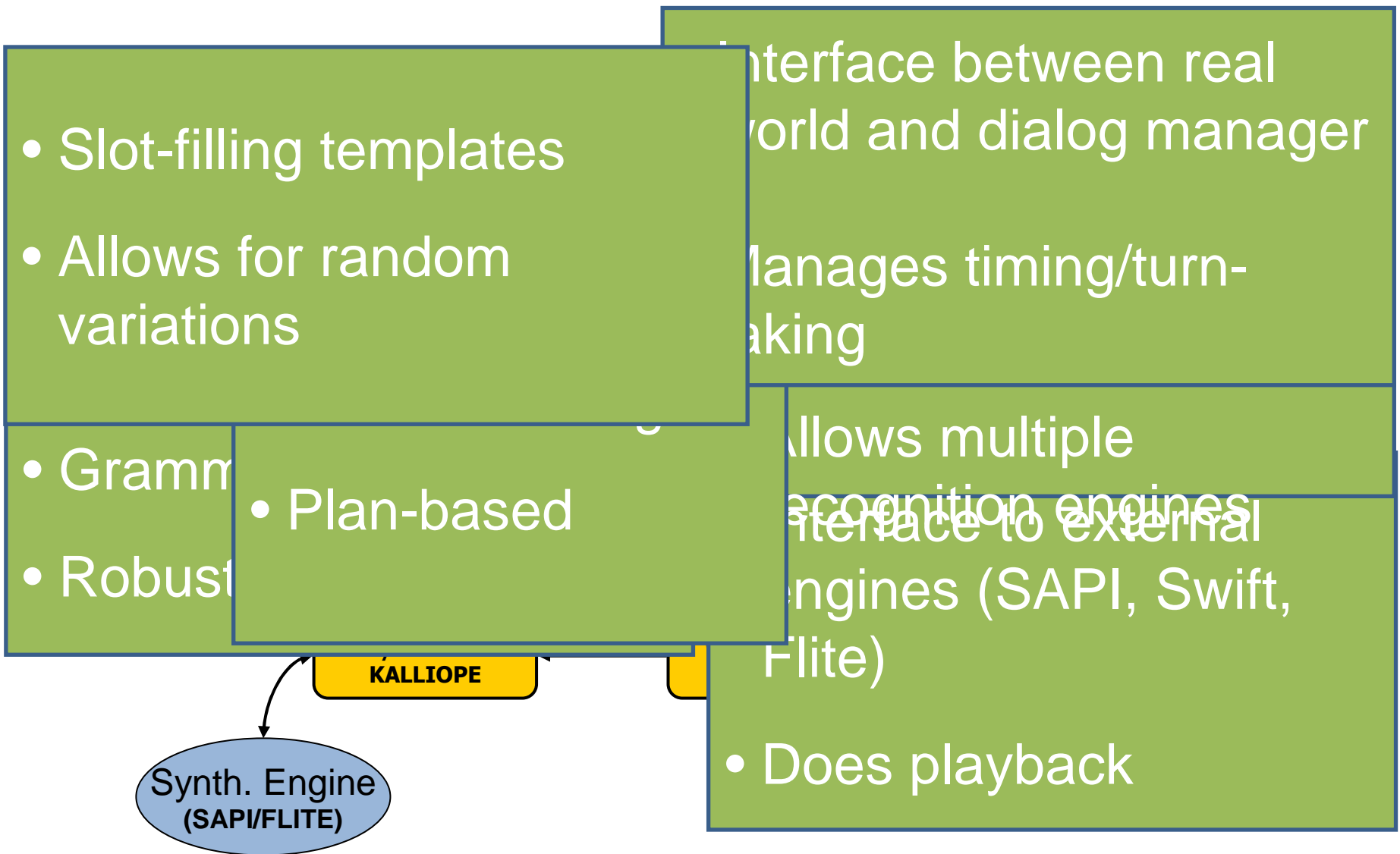
# Typical Pipeline Architecture



# Pipeline Architecture in Olympus



# The Olympus Architecture





# Grammar

- ◆ *Used for two things:*
  - *Parsing*
  - *ASR language model if one isn't available*
- ◆ *The Phoenix Parser*
  - *Context-Free Grammar*
  - *Robust parser*

# Phoenix Parser / Grammar

## ◆ CFG Grammar

- *Manually-generated domain-specific grammar rules*
- *Reusable, generic sub-grammars*
  - ⊗ *[Yes], [No], [Number], [DateTime], [Help], [Repeat], [Suspend], etc...*

```
DO YOU HAVE SOMETHING A BIT LARGER?  
[NeedRoom] (  
  [_i_want] (DO YOU HAVE SOMETHING) )  
[RoomSizeSpec] (  
  [room_size_spec] (  
    [rss_larger] (LARGER)))
```

- ◆ *Parses all incoming hypotheses and passes all parses along...*

```
[room_size_spec]  
  ([rss_large])  
  ([rss_small])  
  ([rss_larger])  
  ([rss_smaller])  
  ([rss_smallest])  
  ([rss_largest])  
;  
[rss_large]  
  (large)  
  (big)  
  (huge)  
;  
[rss_larger]  
  (*the larger)  
  (*the bigger)  
  (too small)  
;  
[rss_largest]  
  (*the largest)  
  (*the biggest)  
;  
[rss_small]  
  (small)  
  (little)  
;
```

# Example Phoenix Grammar

[Place]

(carnegie mellon university)

(downtown)

(robinson towne center)

(the airport)

(south hills junction)

(mount oliver)

(the south side)

(oakland)

(bloomfield)

(polish hill)

(the strip district)

(the north side)

;

[NextBus]

(\*WHEN\_IS \*the next \*BUS)

(\*WHEN\_IS \*the BUS after that \*BUS)

WHEN\_IS

(when is)

(when's)

BUS

(bus)

(one)

;

# Confidence Annotation - Helios

- ◆ *Builds accurate confidence scores using features from 3 sources of knowledge:*
  - *Speech recognition*
  - *Language understanding*
  - *Dialog management*
- ◆ *Selects hypothesis with maximum confidence score*

