

12.0 Computer-Assisted Language Learning (CALL)

References:

1. “An Overview of Spoken Language Technology for Education”, Speech Communications, 51, pp.832-844, 2009
2. “Computer-assisted Language Learning (CALL) Systems”, Tutorial, Interspeech 2012
3. “A Recursive Dialogue Game for Personalized Computer-Aided Pronunciation Training”, IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 23, No. 1, Jan 2015, pp. 127-141.
4. “Supervised Detection and Unsupervised Discovery of Pronunciation Error Patterns for Computer-Assisted Language Learning”, IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 23, No. 3, Mar 2015, pp. 564-579.

Computer-Assisted Language Learning (CALL)

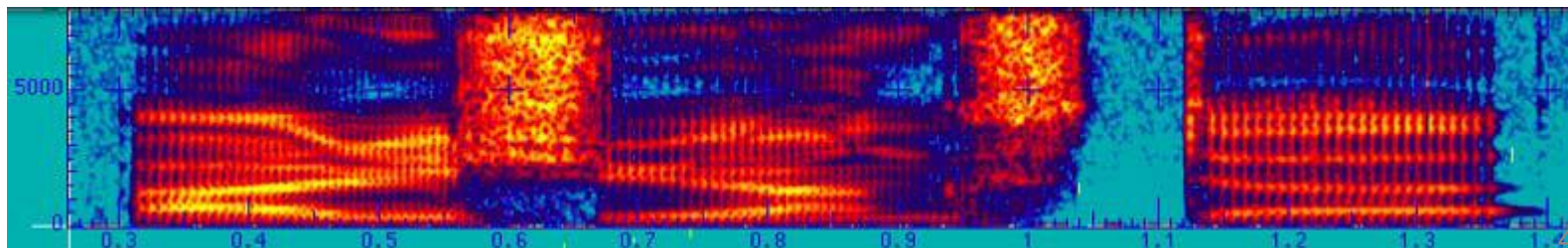
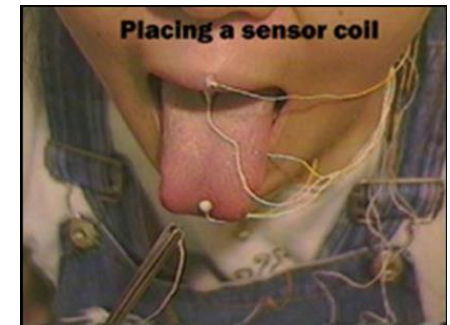
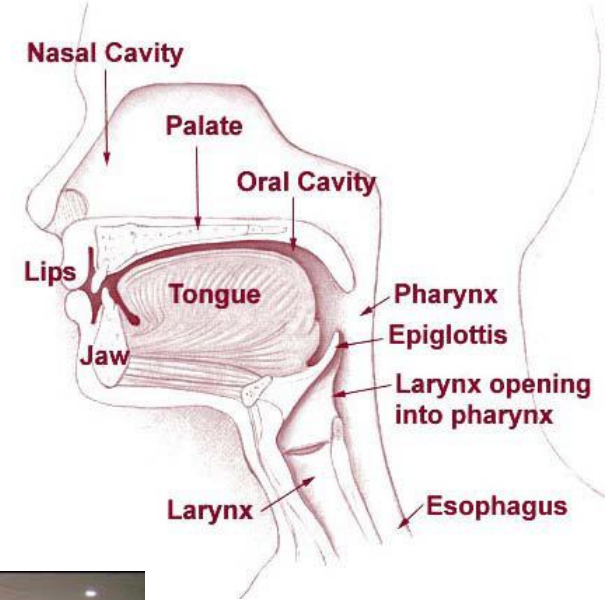
- **Globalized World**
 - every one needs to learn one or more languages in addition to the native language
- **Language Learning**
 - one-to-one tutoring most effective but with high cost
- **Computers not as good as Human Tutors**
 - software reproduced easily
 - used repeatedly any time, anywhere
 - never get tired or bored

Target Skills of CALL

- **Components and sentence composition**
 - Phoneme set, Vocabulary, Grammar
- **Pronunciation: Phonetic and Prosodic**
 - Phoneme
 - Word + tones, stress, etc.
 - Sentence + intonation, etc.
 - Paragraph + prominence, etc.
 - Computer-aided Pronunciation Training (CAPT)
- **Reading**
- **Writing (Chinese characters, etc.)**
- **Listening**
- **Speaking**
- **Dialogues and Communications**
 - travel/shopping, business/negotiation, etc.

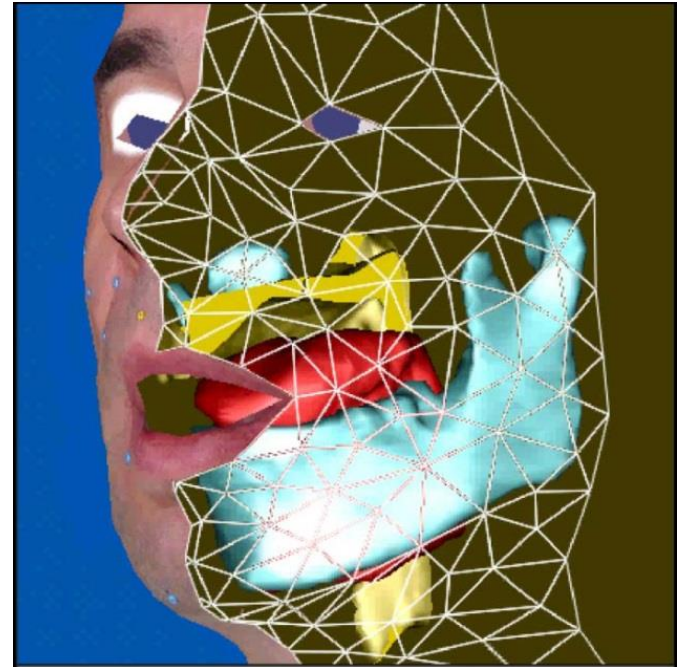
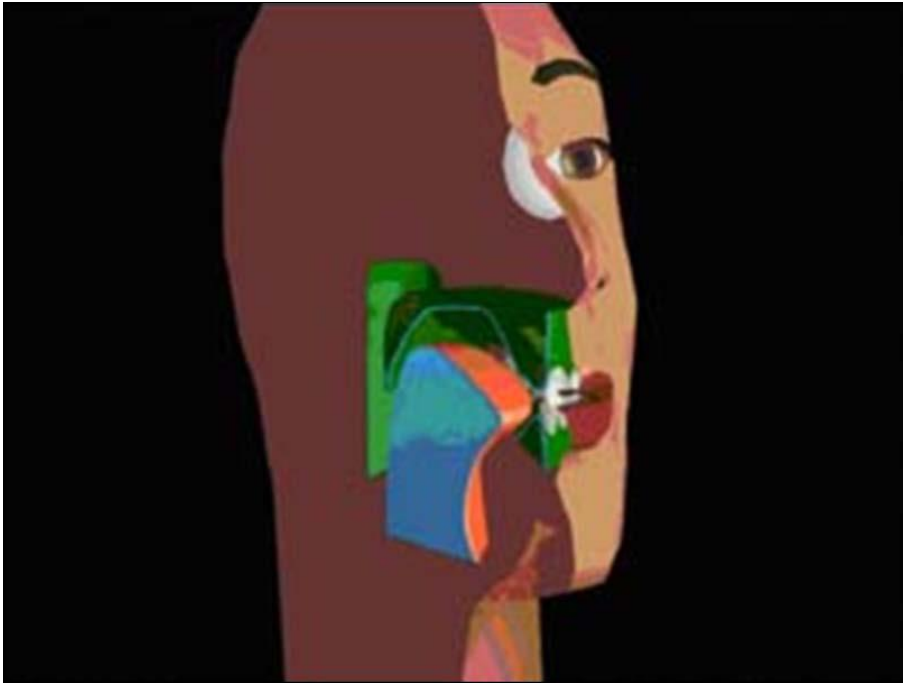
From Articulation to Speech

- Learners are supposed to learn how to control the articulators (vocal tract)
- But the movement of these organs is not easy to observe
- Observation from signals is feasible, but not easy to learn based on signals either



Visual Presentation of Articulation

- Talking Head showing correct articulation
- Acoustic-to-articulatory inversion to estimate the articulatory movements



- Still difficult for learners

Commonly Used Approaches

- **Computer-Aided Pronunciation Training (CAPT)**
 - Qualitative assessment of pronunciation for learners
 - Error pattern detection



Commonly Used Approaches

- **Computer-Aided Pronunciation Training (CAPT)**
 - Qualitative assessment of pronunciation for learners
 - Error pattern detection
- **Spoken Dialogue System**
 - Immersive interactive environment

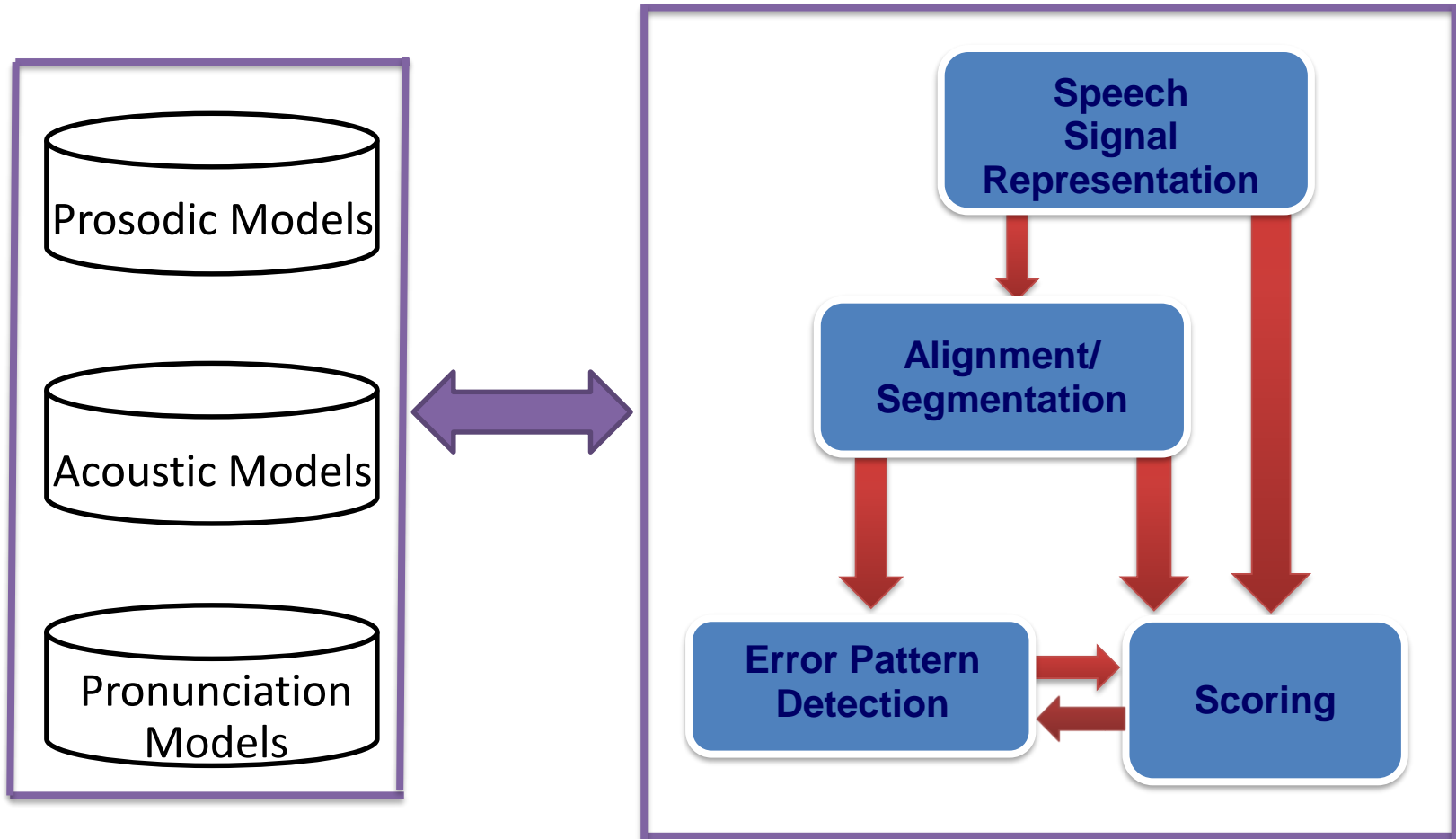


Commonly Used Approaches

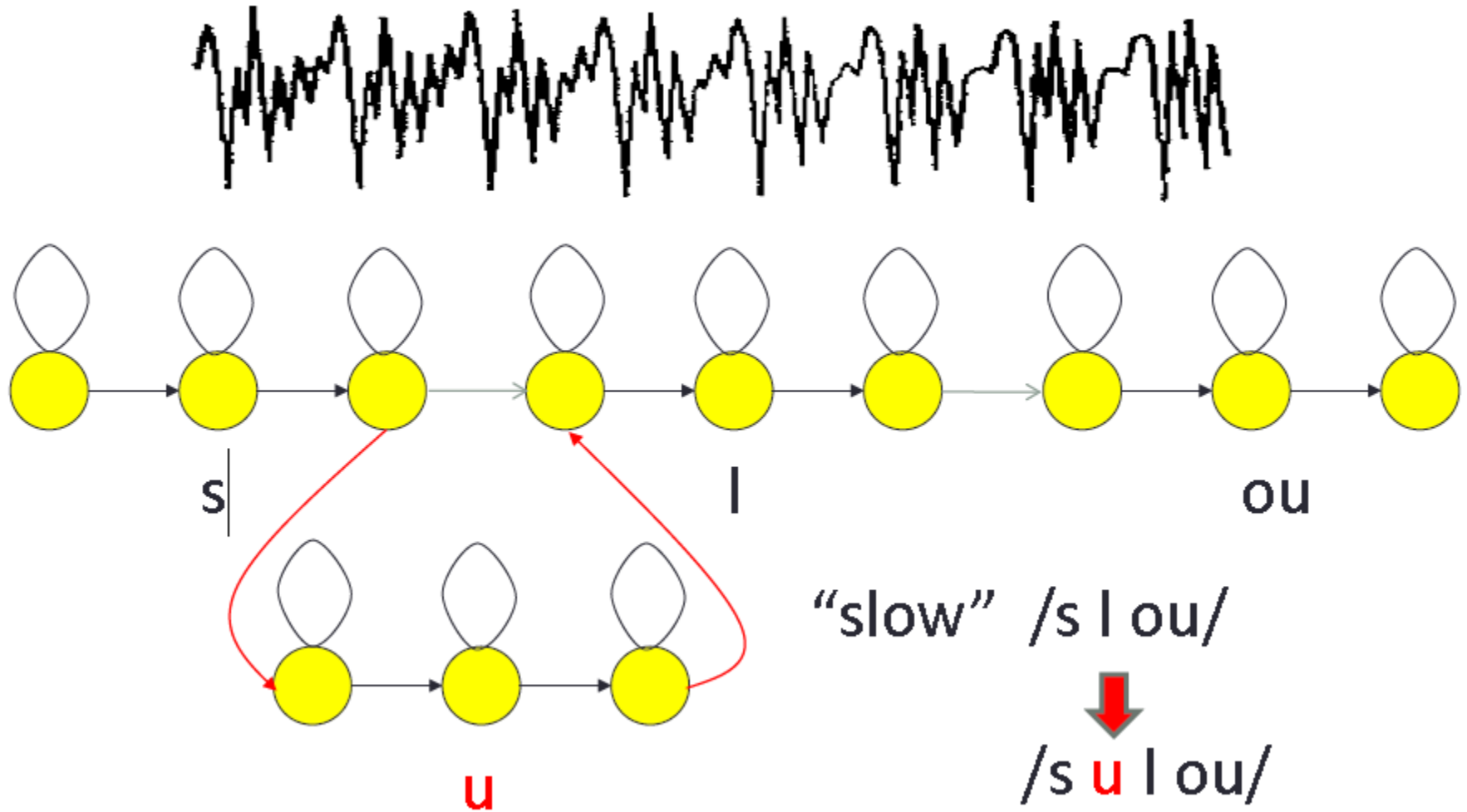
- **Computer-Aided Pronunciation Training (CAPT)**
 - Qualitative assessment of pronunciation for learners
 - Error pattern detection
- **Spoken Dialogue System**
 - Immersive interactive environment
 - Corrective feedback during interaction



Pronunciation Scoring and Error Pattern Detection

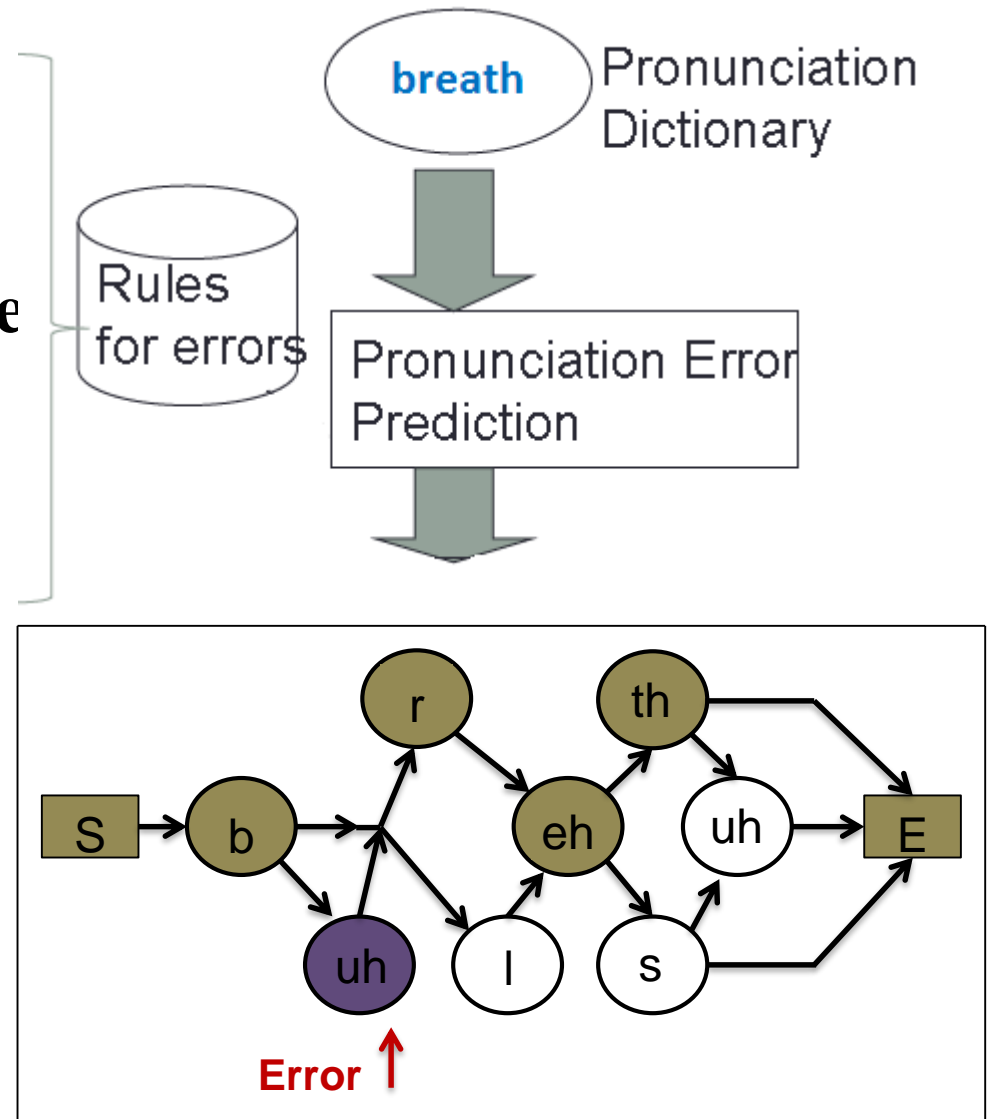


Alignment Problem – Insertion Error



Error Prediction in Pronunciation Modeling

- **No corresponding syllable in L1**
(ex.) sea → she
- **No corresponding phoneme in L1**
(ex.) r → l, v → b
- **Vowel insertions**
(ex.) b-r → b-uh-r



Pronunciation Scoring / Error Pattern Detection

- **Native-likeness**
 - How close to golden native speakers?
 - who are the “golden” speakers?
 - Models trained with a group of “good” speakers
- **Intelligibility**
 - How distinguishable (less confusable) from other phonemes?
- **Learning from Human Language Teachers**
 - Trained to offer scores or error patterns close to scores or patterns given by human language teachers

Example :
Dialogue Game for Pronunciation Learning

Goal of Dialogue Game (1/3)

- **CALL – CAPT**

- **NTU Chinese** offers a score and multi-faceted corrective feedbacks to each pronunciation unit



發音

單字	發音對照	音標	分數
你	[Green play button] [Orange play button]	n	100
		I [Speaker icon]	59
真	[Green play button] [Orange play button]	zh	100
		ēn	100
是	[Green play button] [Orange play button]	shì	70
一	[Green play button] [Orange play button]	yí	50
個	[Green play button] [Orange play button]	gè	100

發音總分 82

發音診斷

您的母音 [ɪ]，發音不是很正確呢！
You are having trouble pronouncing the vowel [ɪ].

發音方法： 舌頭前端向上升起，兩側輕輕頂住上排兩個牙齒，上下嘴唇兩側分開，吐氣振動聲帶發出聲音。

Tips: Rise the front blade of the tongue toward the hard palate, with the mouth slightly open and the lips flat, and let the breath come out between the hard palate and the blade of the tongue.

總分 67

發音 82

聽讀 32

流利度 57

音量 82

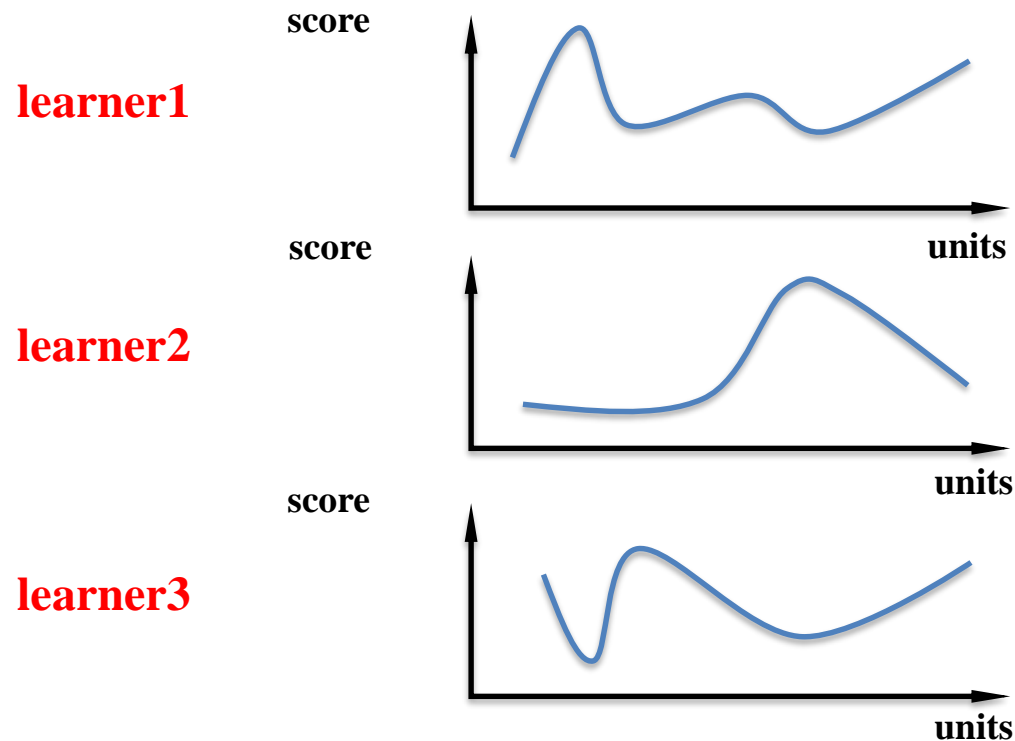
老師 [Waveform]

學生 [Waveform]

- ✓ learner/ref comparison
- ✓ Qualitative assessment
- ✓ Corrective feedback

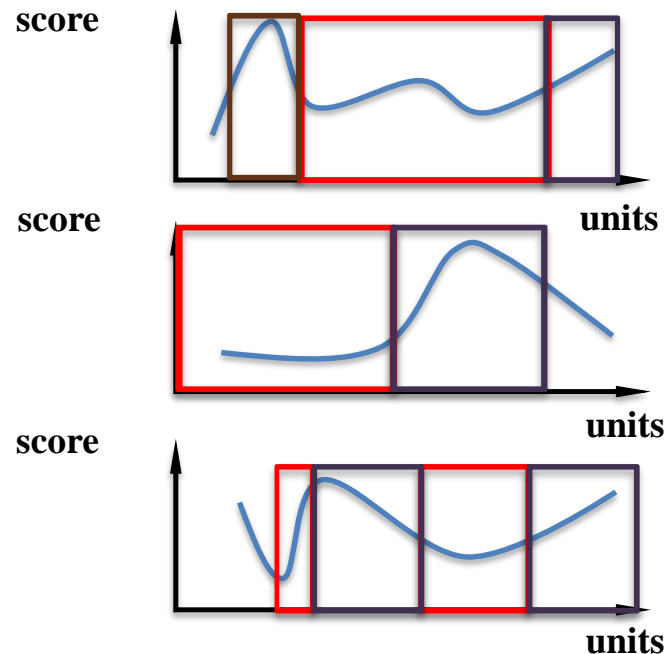
Goal of Dialogue Game (1/3)

- **CALL – CAPT**
 - **NTU Chinese** offers a score and multi-faceted corrective feedbacks to each pronunciation unit
 - Different learners have very different **performance distributions** over different pronunciation units



Goal of Dialogue Game (2/3)

- We wish for each individual learner and each pronunciation unit
 - The worse the score is, the more practice ■
 - The higher the score is, the less care ■
 - No need for repeated practice on the same sentence, but participating in a more interesting **dialogue game**
 - The needed practice opportunities automatically appears **along the dialogue**



Goal of Dialogue Game (3/3)

- **Personalized learning materials based on learning status dynamically obtained on-line along the dialogue game**
- **To achieve this goal**
 - Recursive tree-structured dialogue script
 - Best path within the dialogue script for each individual learner found by MDP

Dialogue Game Script (1/2)

- Tree-structured turn-taking dialogue
 - Restaurant scenario: seating and meal ordering

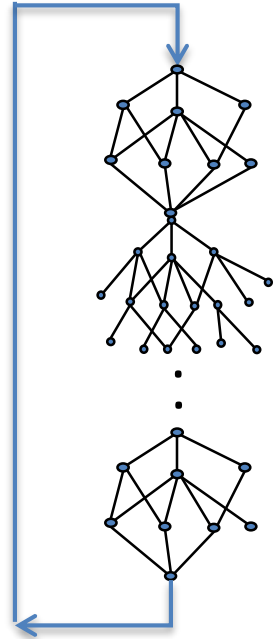
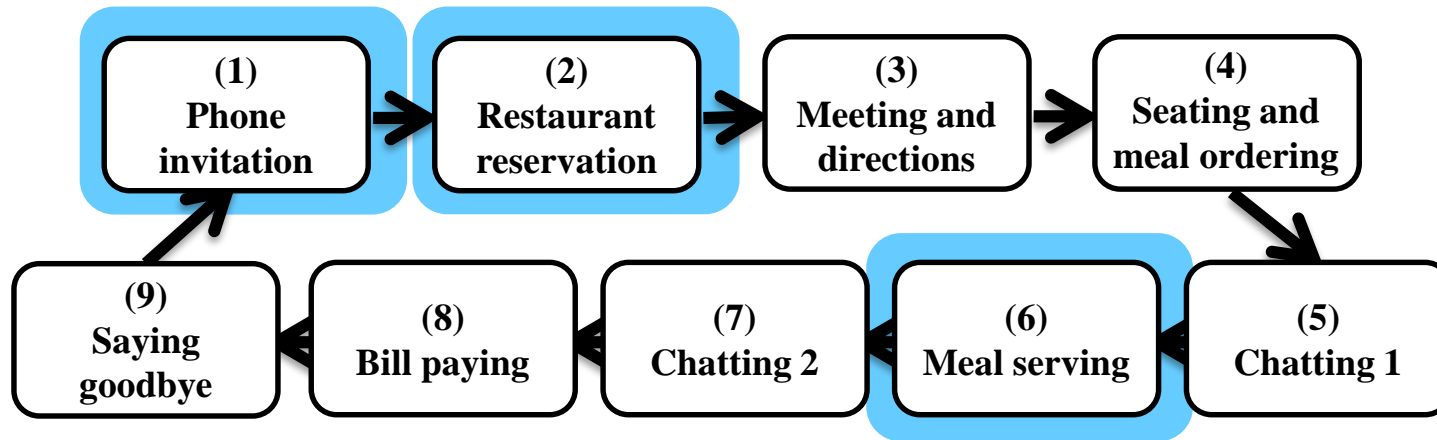
A as Waiter B as Customer

TURN	CONTENT			
A1	歡迎光臨。請問訂位了嗎？ Welcome, do you make any reservation?			
B1	沒有。 No, I don't.		訂了，我姓王。 Yes, My name is Wang.	
A2	現在客滿， 您可能要稍等一下。 It's full now, you may have to wait for a while.	正在整理桌面， 請稍候。 We are cleaning up the table, please wait.	請等一下， 馬上就替您帶位。 Please wait, we will lead you to your seat.	這邊請。 Here please.
B2	謝謝。 Thank you.	好的。 Okay.	下一個就是我們了嗎？謝謝 Are we next? Thanks.	
A3	這個位子可以嗎？ Are the seats okay?	靠窗的桌子， 好不好？ The table near window, is it okay?	對不起， 位子有點兒擠。 Sorry, the seats are small.	我再幫你加 一張椅子。 I'll add a chair for you.

Dialogue Game Script (2/2)

- **Tree-structured turn-taking dialogue**

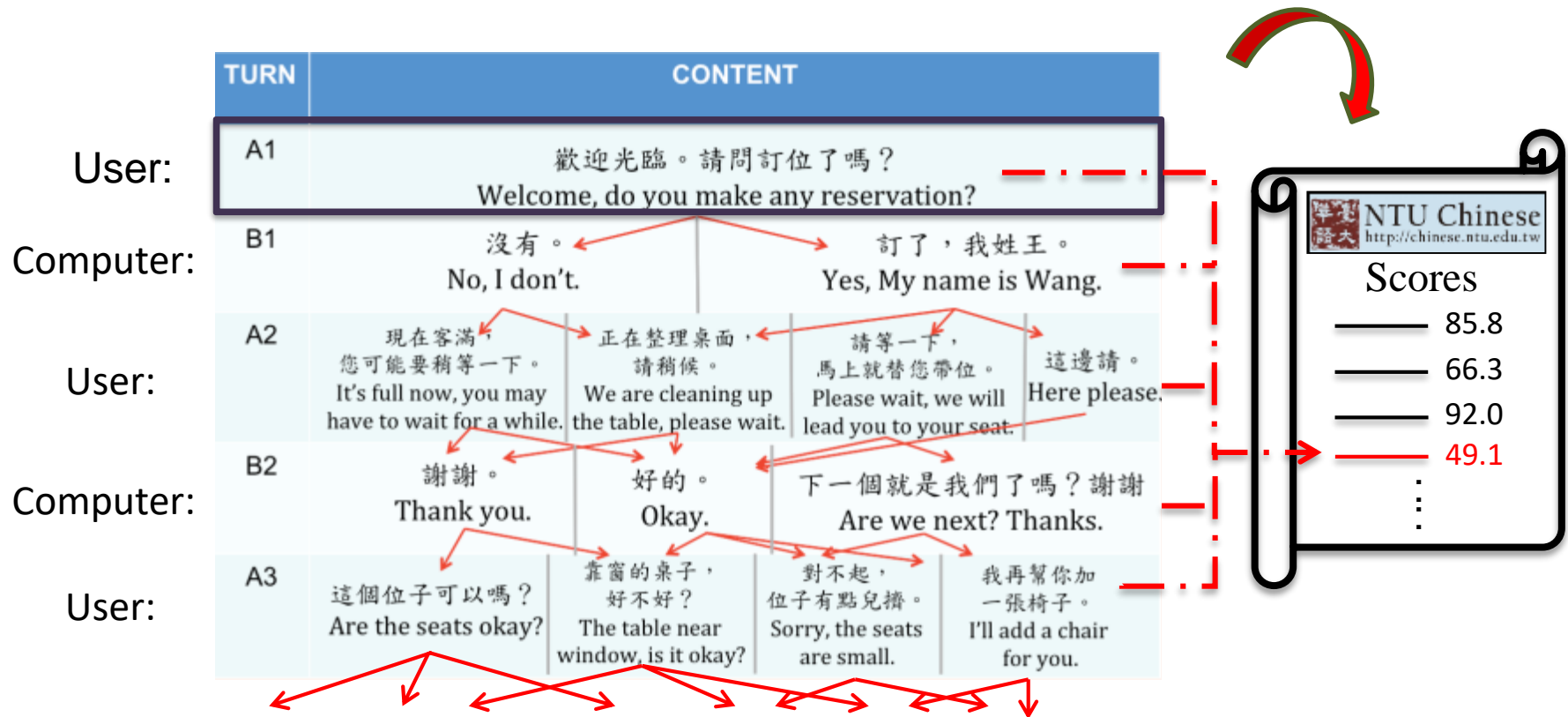
- Restaurant scenario
- 9 sub-dialogues linked recursively



- Almost infinite number of paths within the recursive trees
- Different paths contain different distributions of the pronunciation units for practice
- Different paths good for different learners

System Objective (1/2)

- Based on the recursive dialogue script, the system provides **personalized learning materials** for each individual learner considering his **learning status**

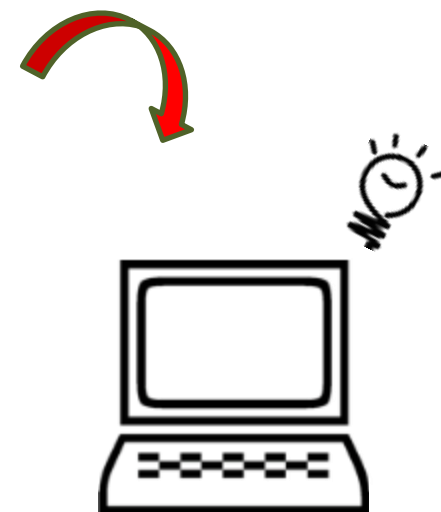


Learning Status: Scores of each unit evaluated by NTU Chinese

System Objective (2/2)

- The system selects on-line the path with the most practice for the lower-scored units for the learner so far, and return the corresponding next sentence to practice

	TURN	CONTENT
User:	A1	歡迎光臨。請問訂位了嗎？ Welcome, do you make any reservation?
Computer:	B1	沒有。 No, I don't.
		訂了，我姓王。 Yes, My name is Wang.
User:	A2	現在客滿，您可能要稍等一下。 It's full now, you may have to wait for a while.
		正在整理桌面，請稍候。 We are cleaning up the table, please wait.
		請等一下，馬上就替您帶位。 Please wait, we will lead you to your seat.
		這邊請。 Here please
Computer:	B2	謝謝。 Thank you.
		好的。 Okay.
		下一個就是我們了嗎？謝謝。 Are we next? Thanks.
User:	A3	這個位子可以嗎？ Are the seats okay?
		靠窗的桌位好不好？ The table near window, is it okay?
		對不起，位子有點兒擠。 Sorry, the seats are small.
		我再幫你加一張椅子。 I'll add a chair for you.



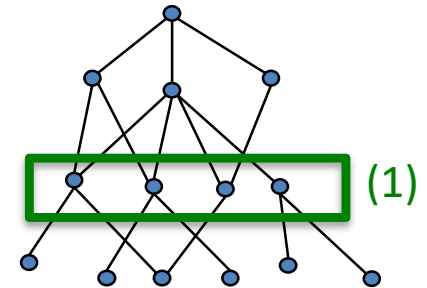
More practice in the **present sentence** doesn't necessarily imply the same for the future sentences along the path

Markov Decision Process (MDP) (1/4)

- **States**
 - learning status of the learner
- **Represented by**
 - Present dialogue turn
 - Learner's average score for every pronunciation unit so far (high-dimensional continuous state space)

(2)

Unit	b	p	m	f	...	45
Score	53.4	89.0	74.7	80.3	...	97.0



Markov Decision Process (MDP) (2/4)

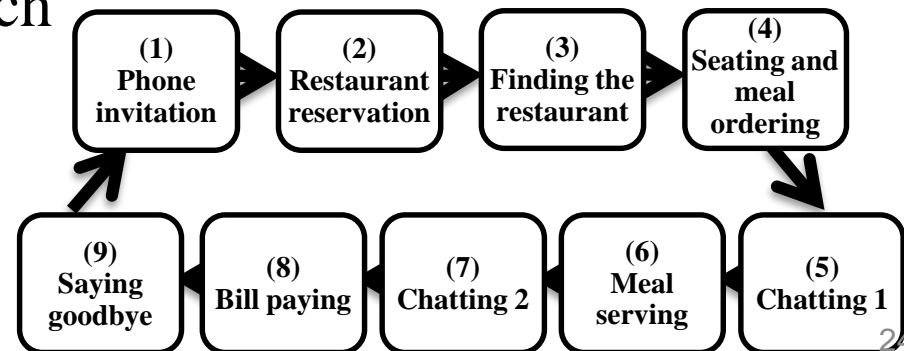
- **Actions**

- The set of sentences to be selected for the learner to practice

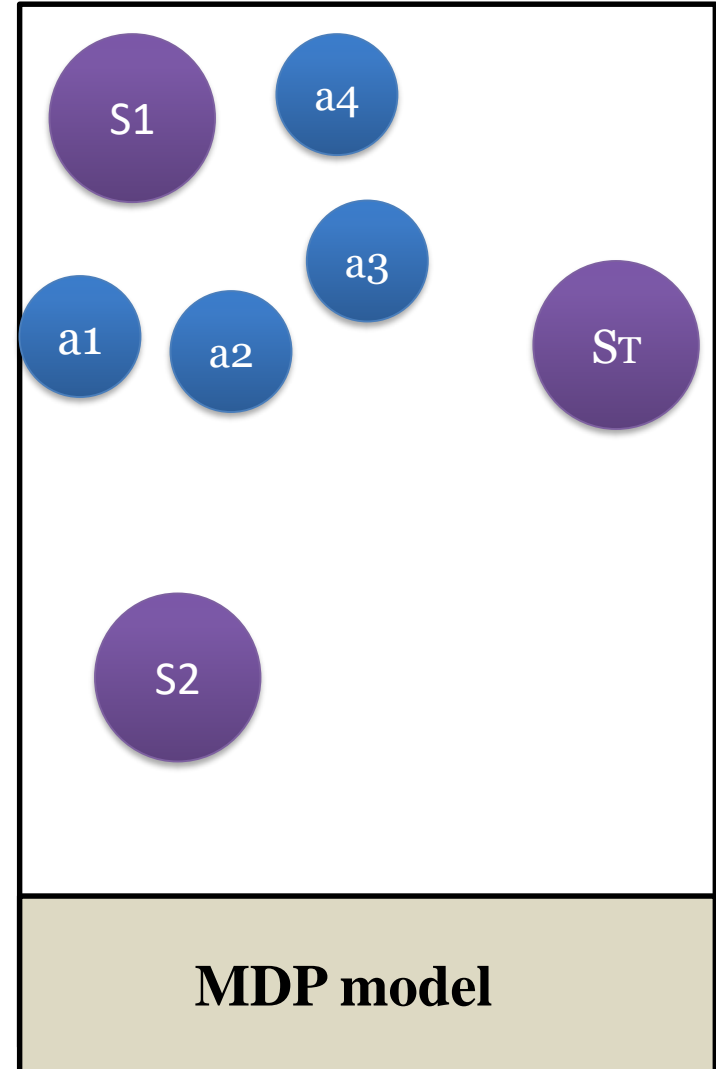
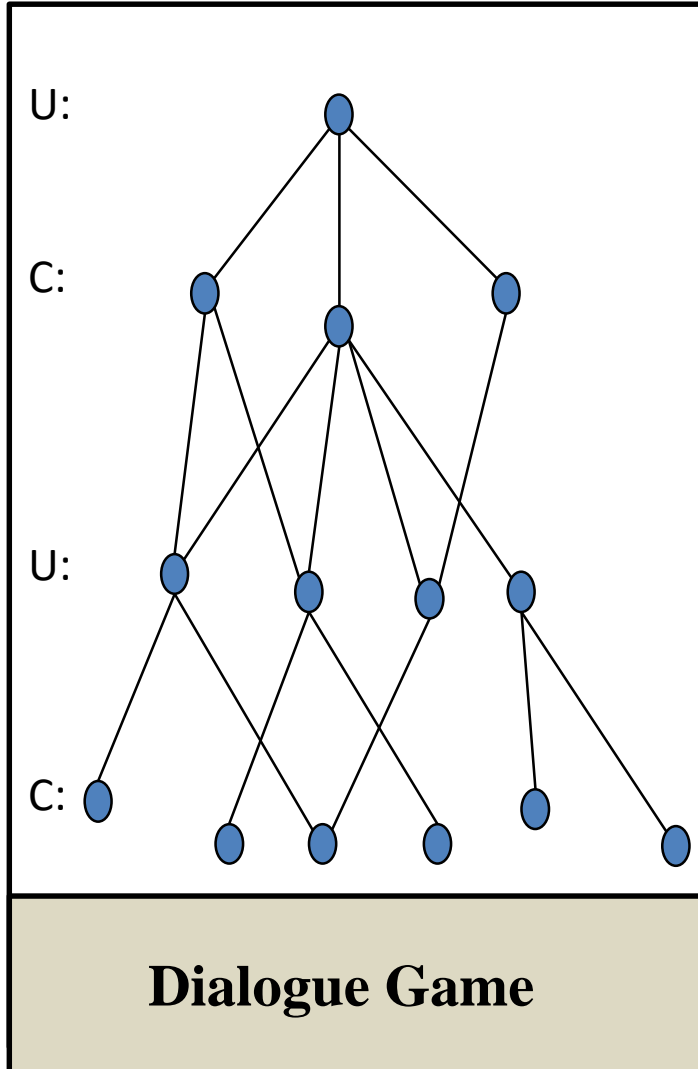
	TURN	CONTENT			
User:	A1	歡迎光臨。請問訂位了嗎？ Welcome. Did you make any reservation?			
Computer:	B1	沒有。 訂了，我姓王。 No, I didn't. Yes, My name is Wang.			
User:	A2	<div>現在客滿， 您可能要稍等一下。 It's full now, you may have to wait for a while.</div> <div>正在整理桌面， 請稍候。 We are cleaning up the table, please wait.</div> <div>請等一下， 馬上就替您帶位。 Please wait, we will lead you to your seat</div> <div>這邊請。 Here please.</div>			
Computer:	B2	謝謝。 好的。 下一個就是我們了嗎？謝謝 Thank you. Okay. Are we next? Thanks.			
User:	A3	這個位子可以嗎？ Are the seats okay?	靠窗的桌子， 好不好？ The table near window, is it okay?	對不起， 位子有點兒擠。 Sorry, the seats are small.	我再幫你加 一張椅子。 I'll add a chair for you.

Markov Decision Process (MDP) (3/4)

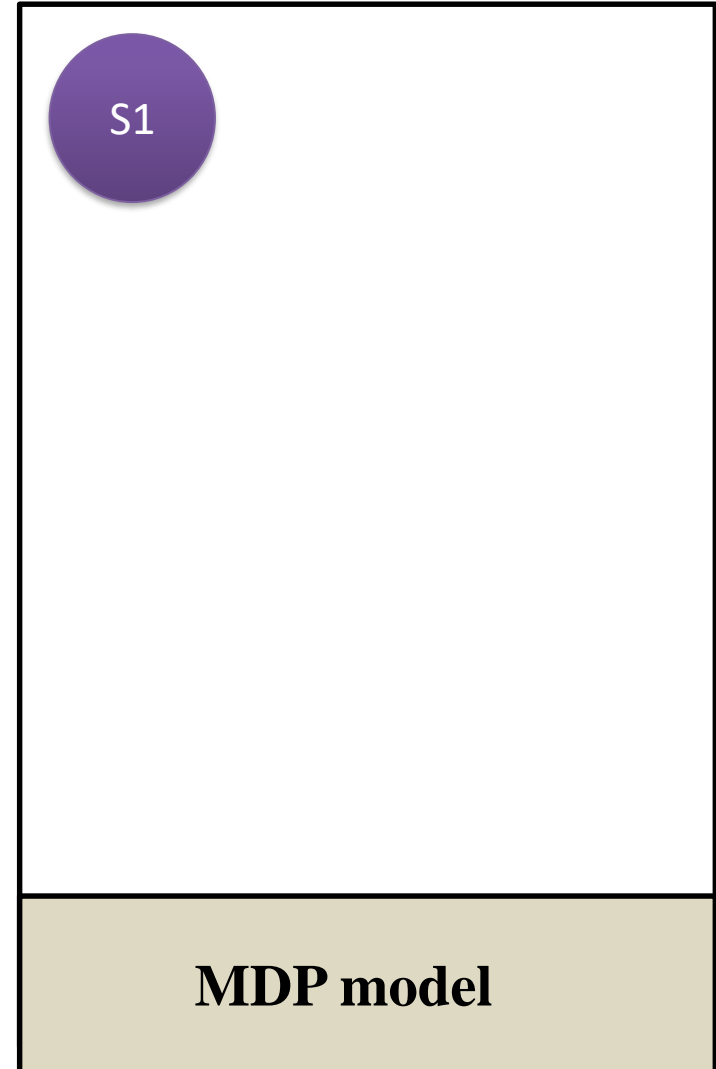
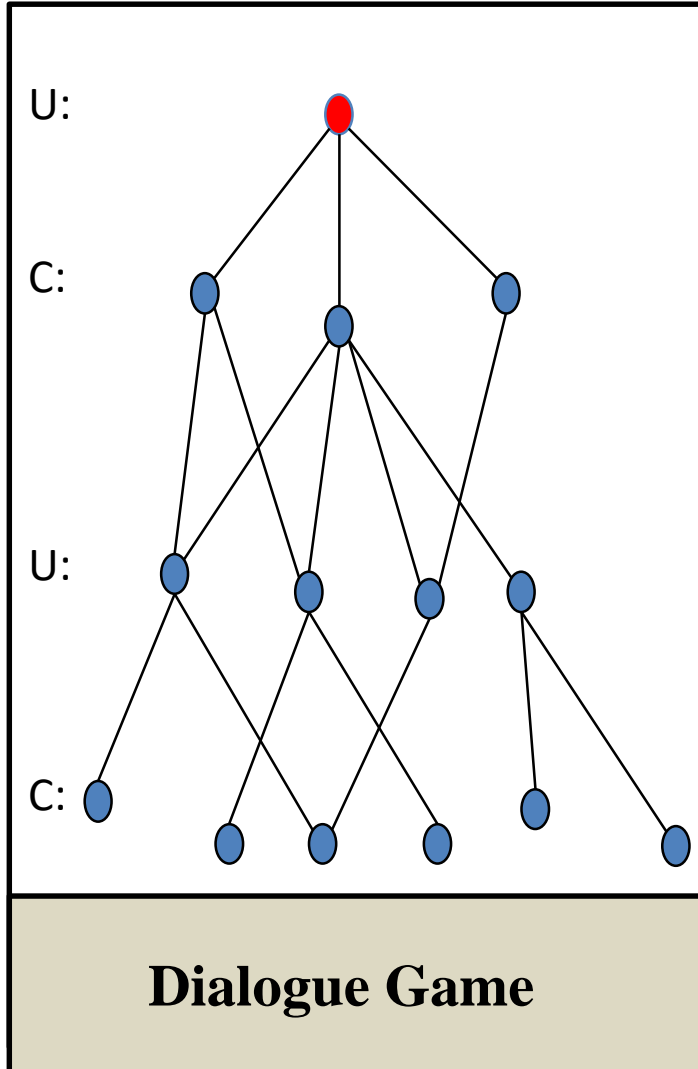
- **System goal setting**
 - all pronunciation units (or a subset of focused units) scored **75 or higher over 7 times** for the learner in minimum number of dialogue turns, etc.
- **Reward**
 - set **cost -1** for every dialogue turn: the less number of turns the better
 - Game ends when system goal reached
- **Policy**
 - Best system action to take at each state
 - to be trained



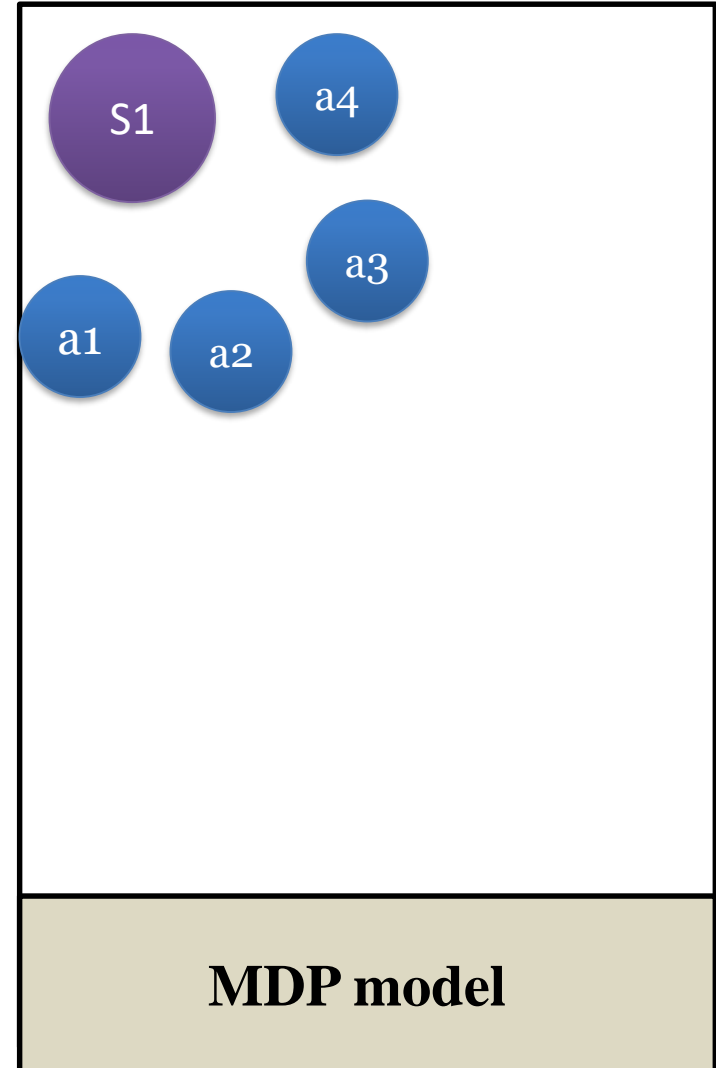
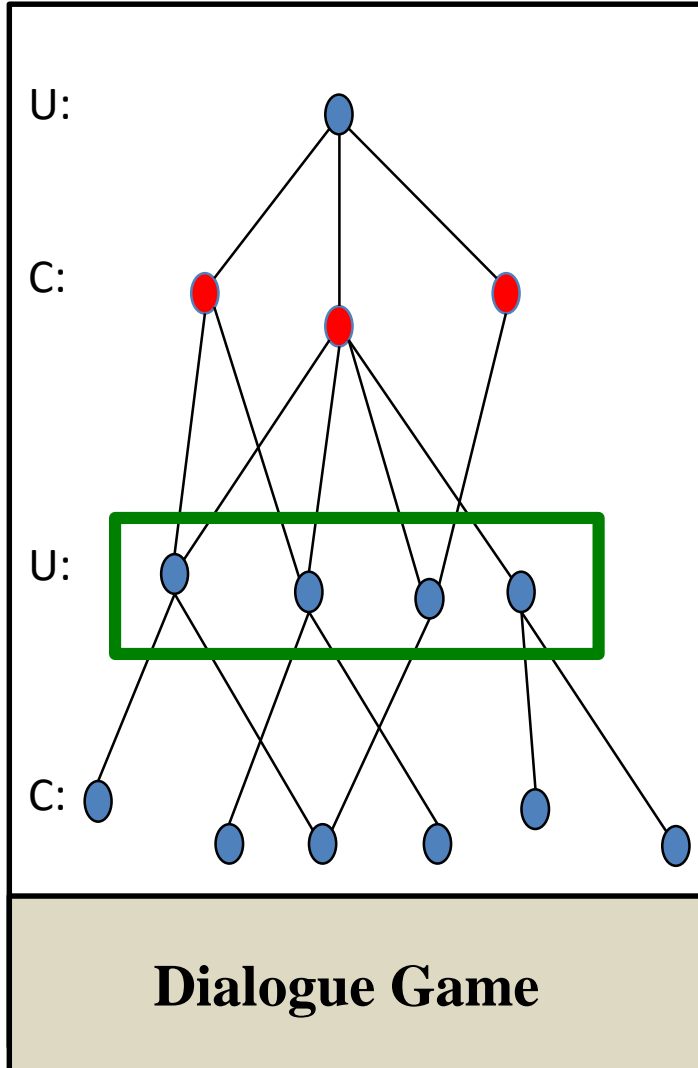
Markov Decision Process (MDP) (4/4)



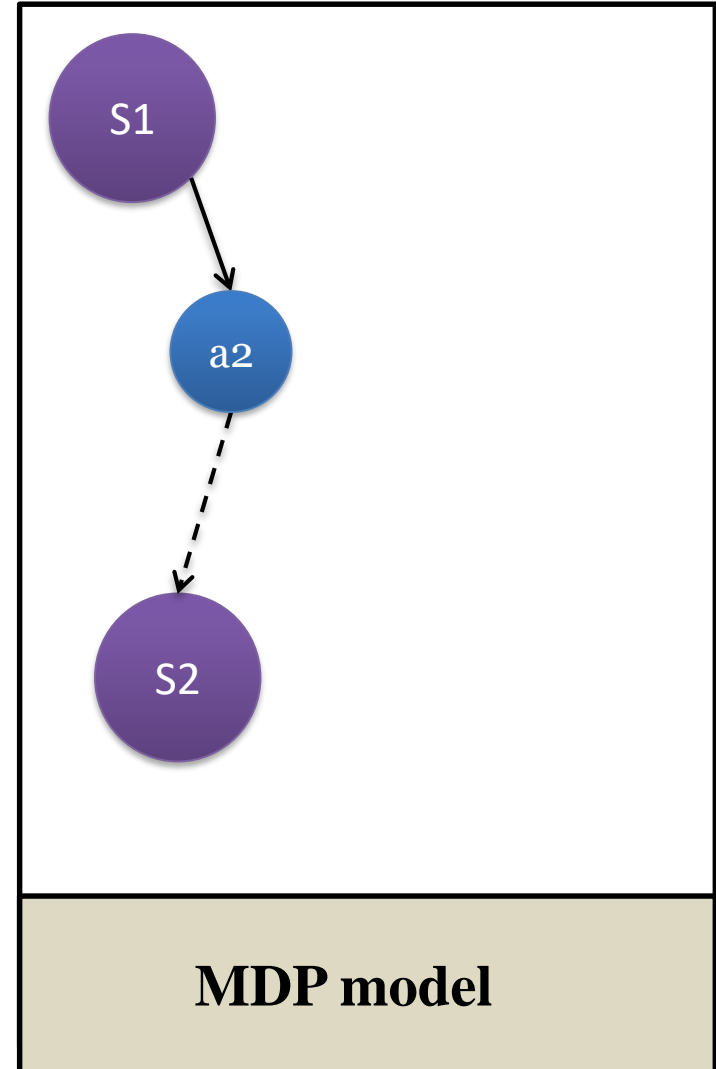
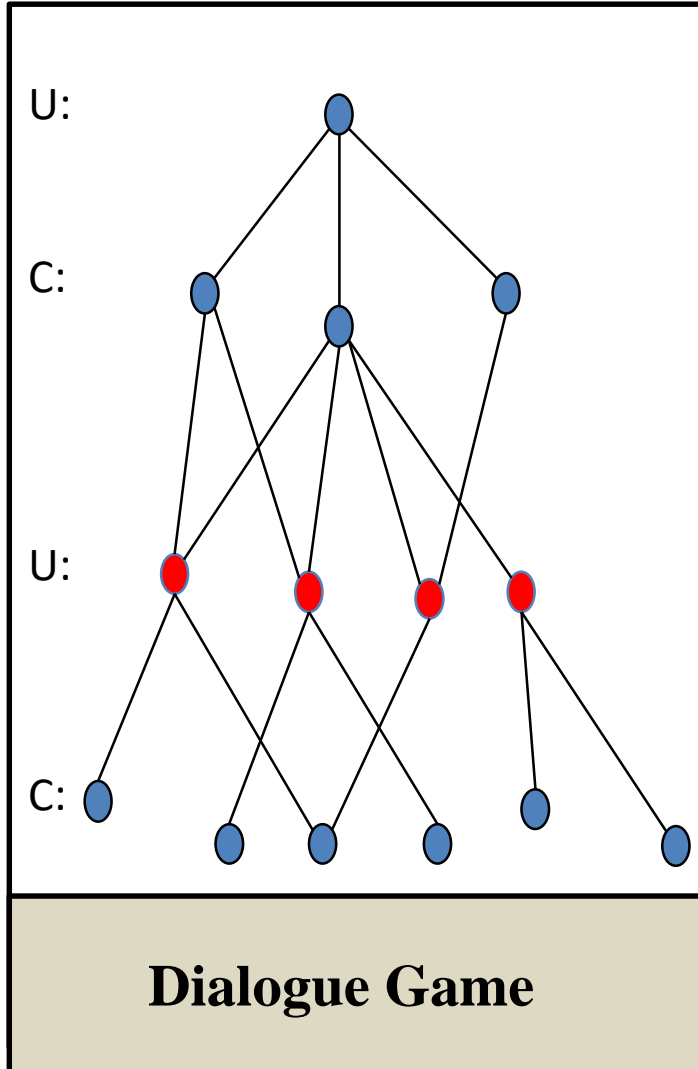
Markov Decision Process (MDP) (4/4)



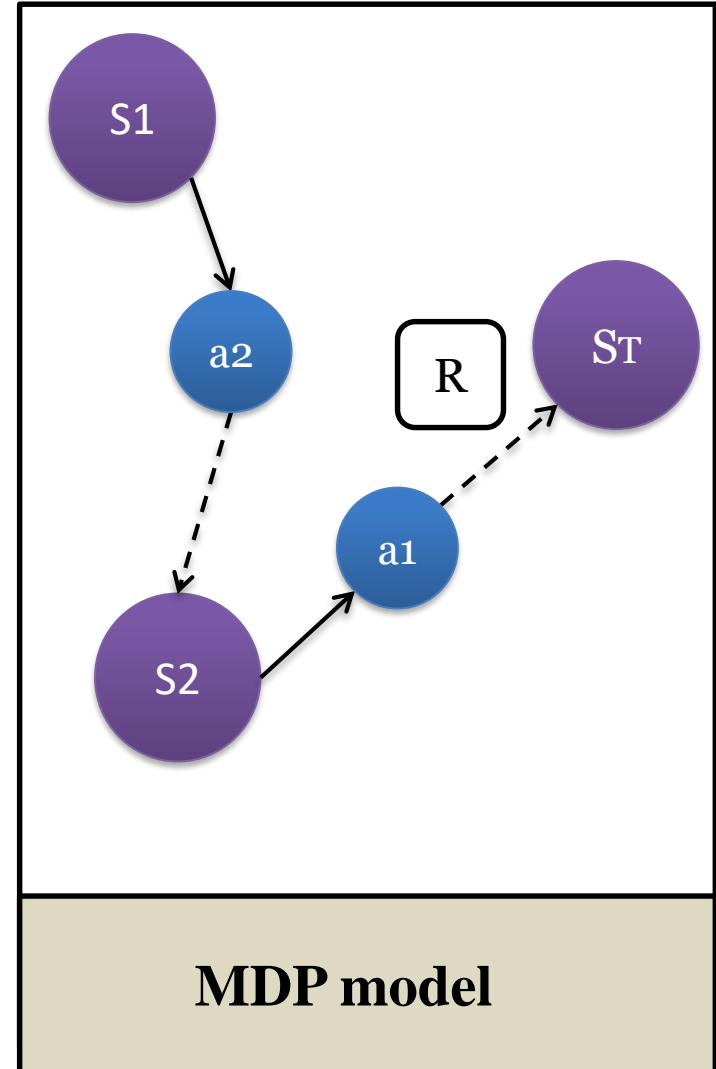
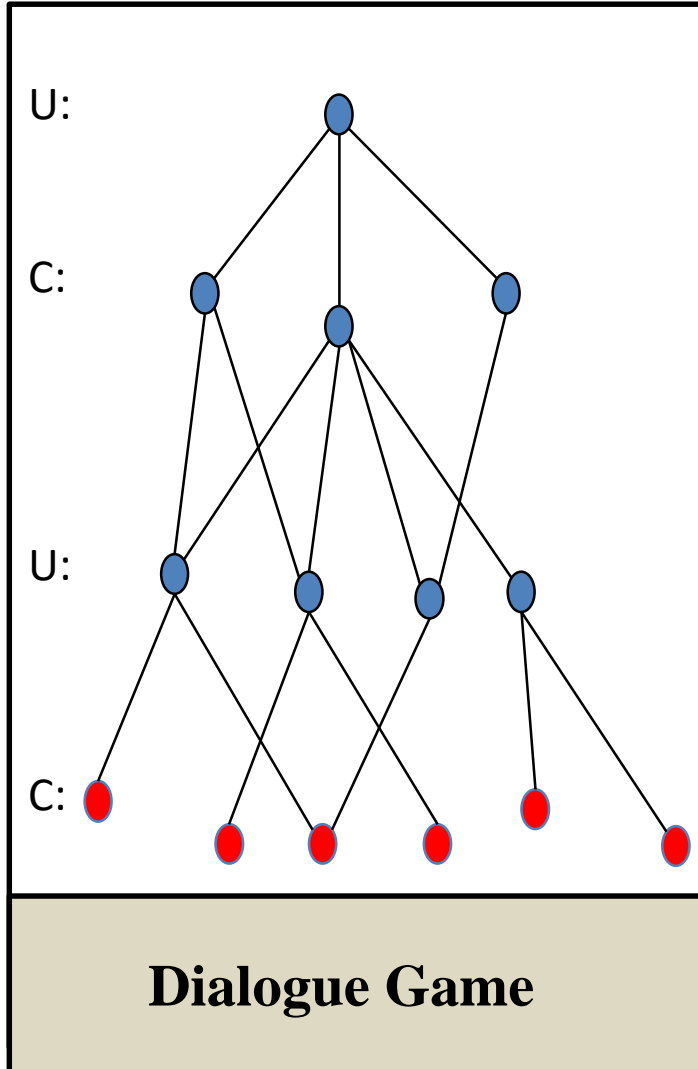
Markov Decision Process (MDP) (4/4)



Markov Decision Process (MDP) (4/4)



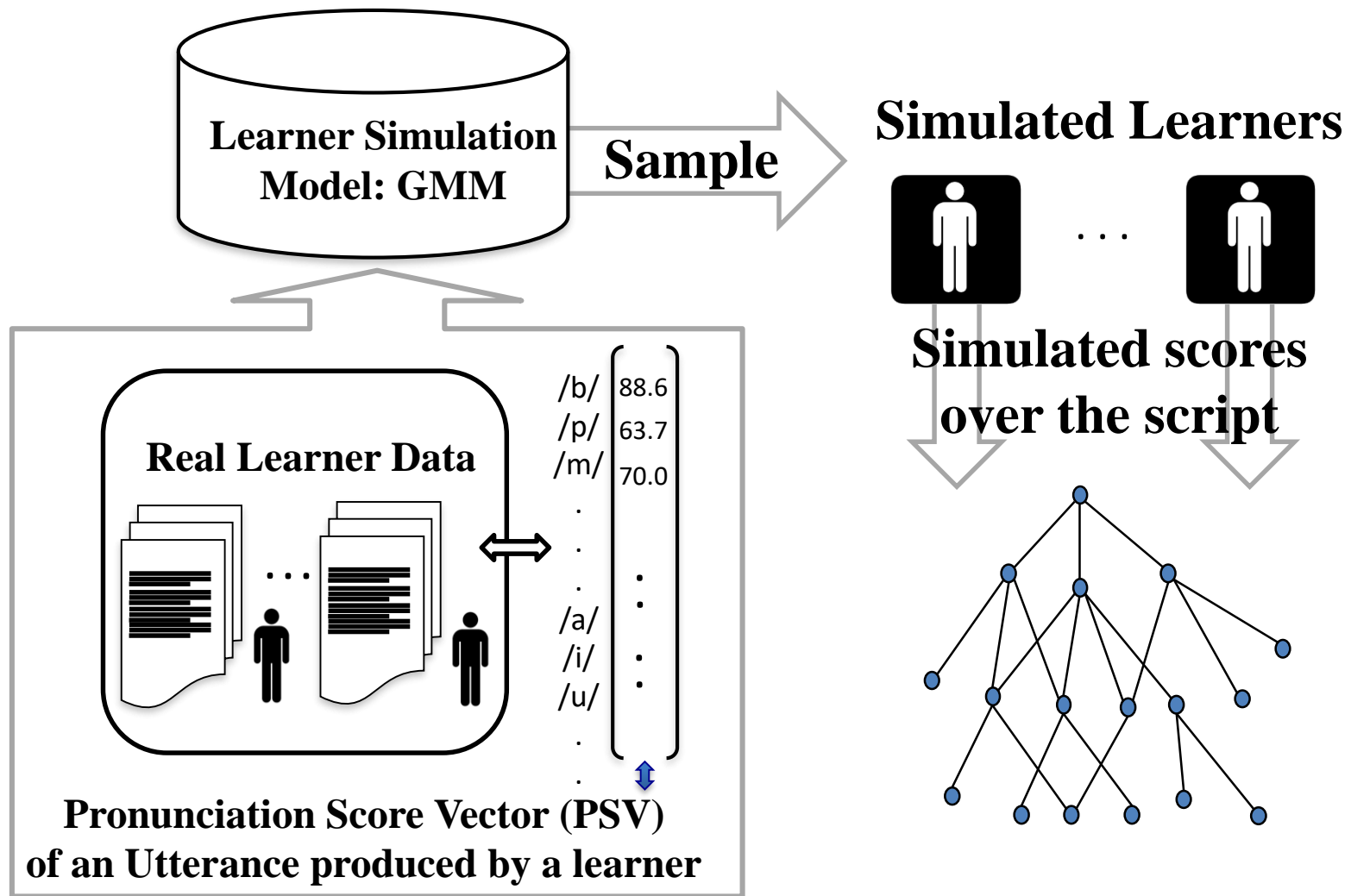
Markov Decision Process (MDP) (4/4)



Learner Simulation (1/3)

- Policy training needs **sufficient** training data
- Since we need real learner's language learning behavior
 - It is not easily available
- **Learner Simulation Model** is developed for generating a large number of training data
- Real learner data
 - 278 learners from 36 countries (balanced gender)
 - Each learner recorded 30 phonetically balanced sentences

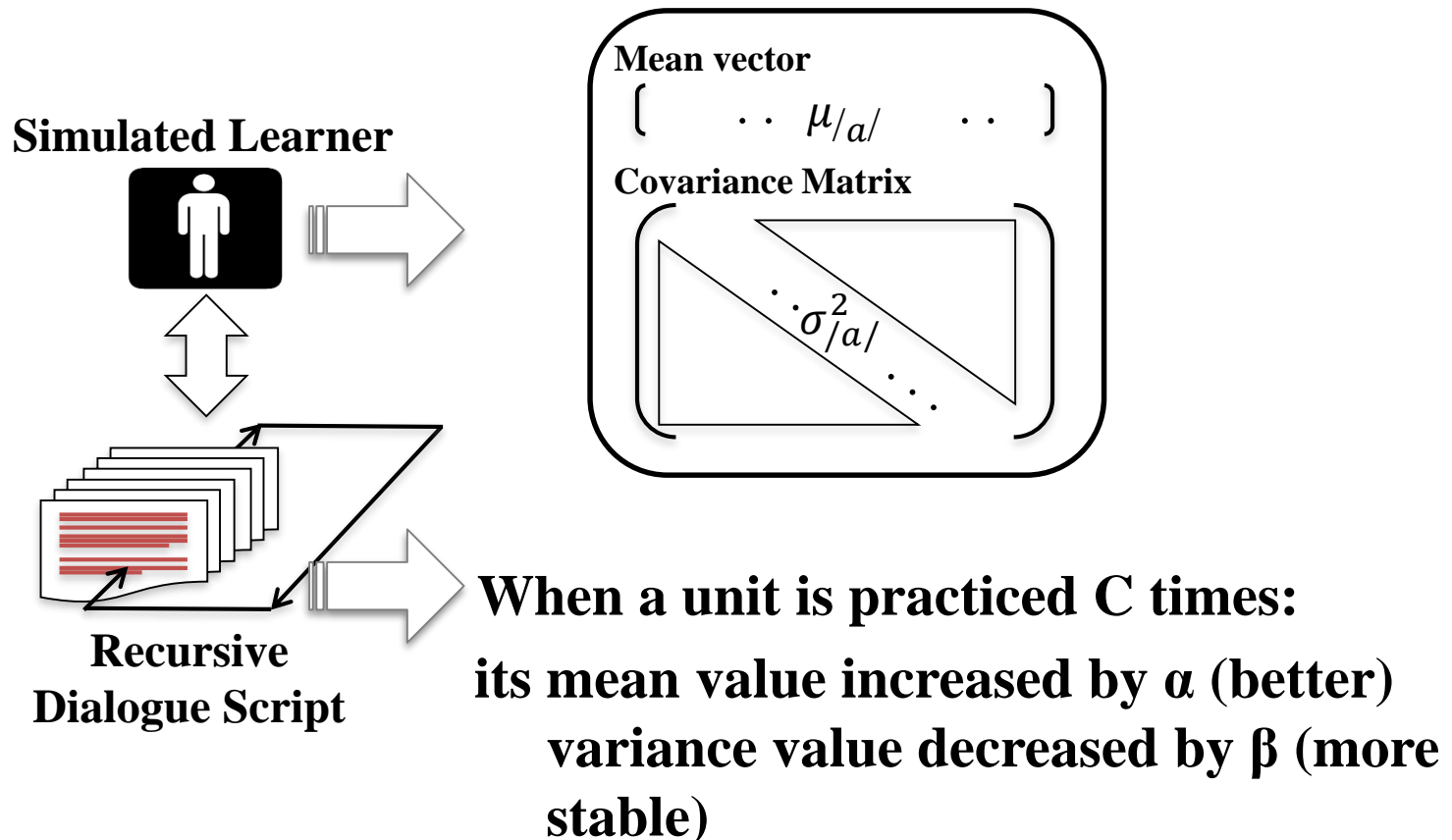
Learner Simulation (2/3)



Each mixture represents a group of learners with a **specific pronunciation characteristics**

Learner Simulation (3/3)

- **Simulation of Incremental Pronunciation Improvement**
 - Gaussian random variables: $C, \alpha, \beta \sim N(\cdot, \cdot)$
(learning rate increases when overall performance is better)



Complete Framework

