



Institute of Automation
Chinese Academy of Sciences

中国科学院自动化研究所

A Spoken Dialogue System based on FST and DBN

Lichun Fan, Dong Yu, Xingyuan Peng,
Shixiang Lu, Bo Xu

Institute of Automation Chinese Academy of Sciences

Outline

- Task
- Problems in NLU
- FST in NLU
- Problems in DM
- DBN in DM
- The whole system
- Conclusion

Outline

- Task
- Problems in NLU
- FST in NLU
- Problems in DM
- DBN in DM
- The whole system
- Conclusion

What is Spoken Dialogue System



What is Spoken Dialogue System



Spoken Dialogue System

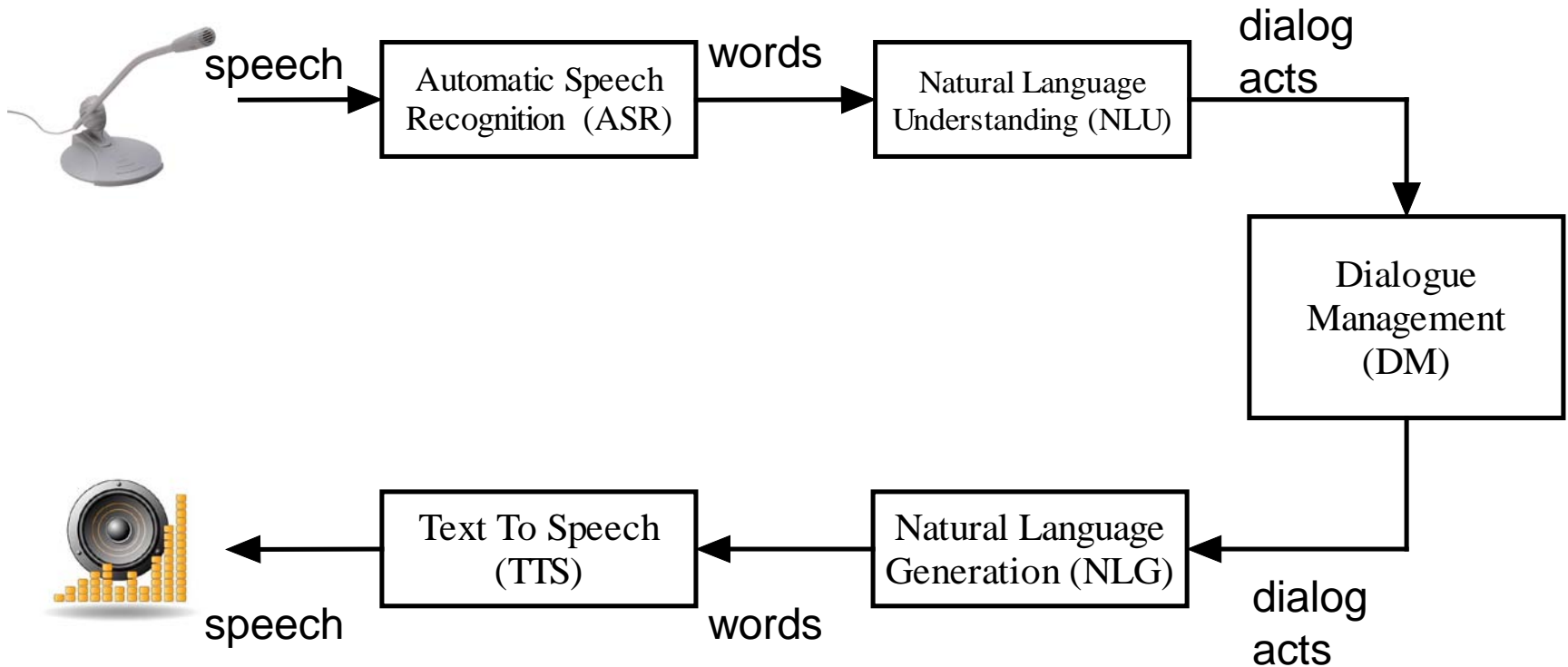
- The Spoken Dialogue System
 - ▶ The dialogue has more than one round
 - ▶ The dialogue is on the same topic
 - ▶ The dialogue context is interconnected

Spoken Dialogue System

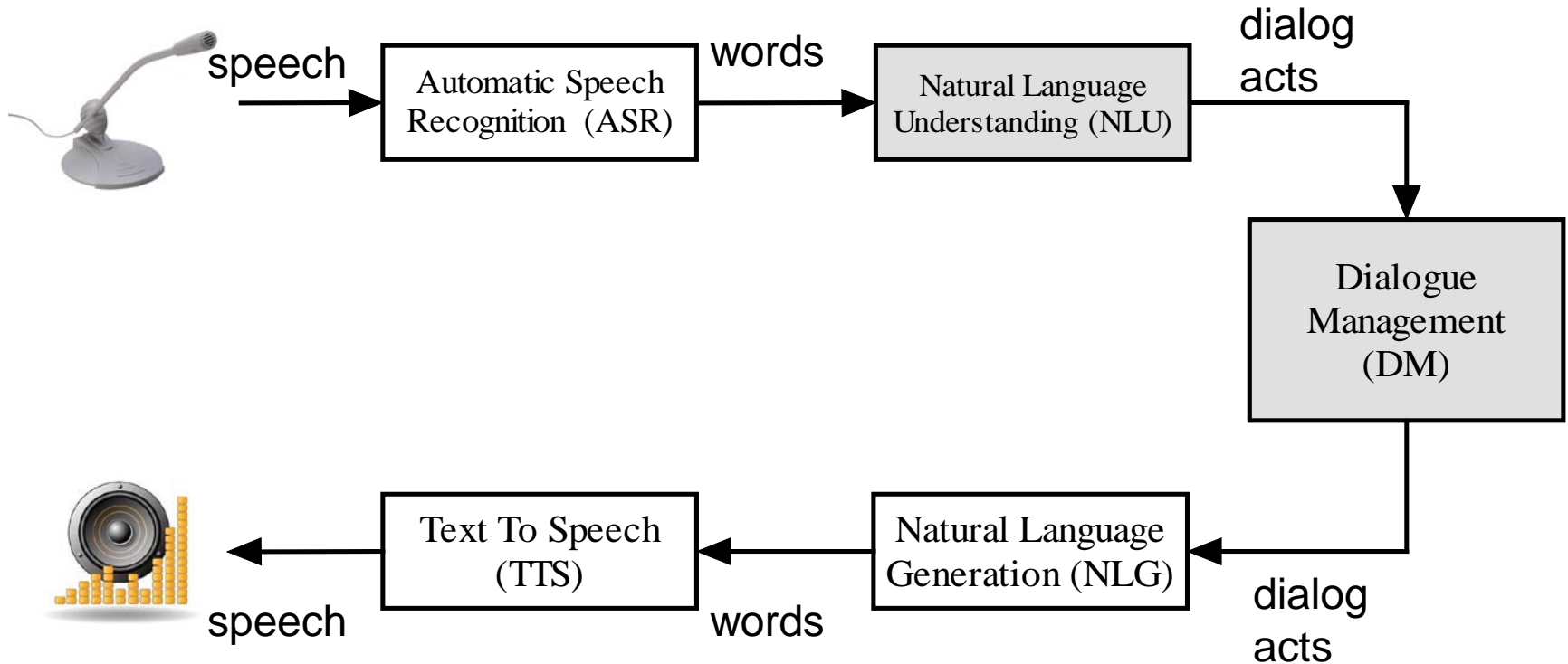
- The Spoken Dialogue System
 - ▶ The dialogue has more than one round
 - ▶ The dialogue is on the same topic
 - ▶ The dialogue context is interconnected
- **An Example**
 - ▶ Ticket reservation system



Architecture of a Spoken Dialogue System



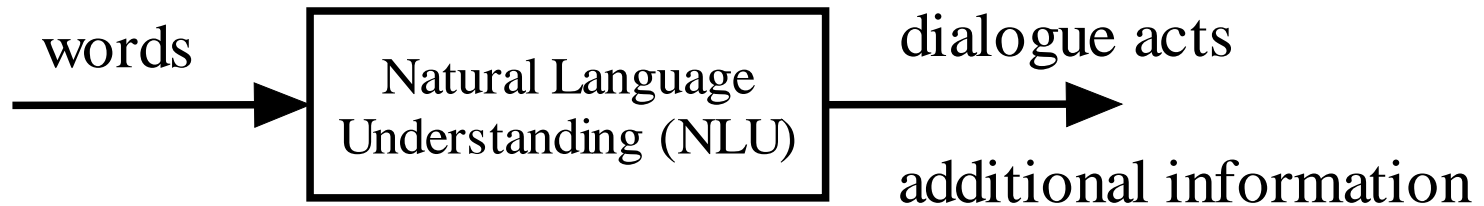
The modules we focus



Outline

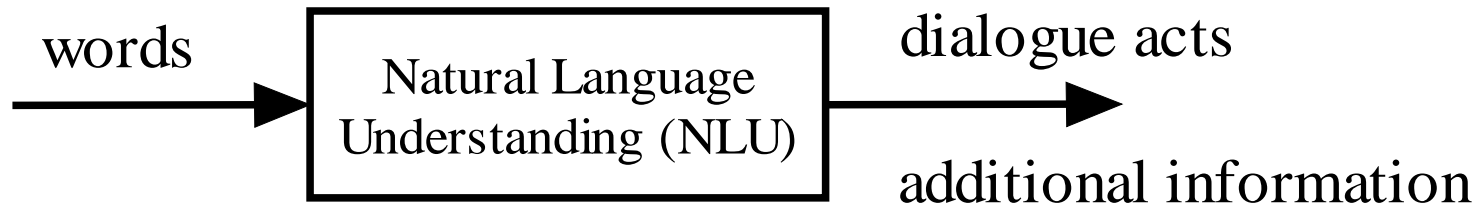
- Task
- **Problems in NLU**
- FST in NLU
- Problems in DM
- DBN in DM
- The whole system
- Conclusion

NLU module function



- map user utterance into Dialogue Act
- extract the additional information in user utterances

NLU module function



- map user utterance into Dialogue Act
- extract the additional information in user utterances
- **Example**
 - ▶ “what time does the next train left to shanghai? ”
 - ▶ **Dialogue Act** : request (time=?)
 - ▶ **Additional information**: To_city=shanghai
From_city= DEFAULT
Vehicle=train

Problems in NLU

- Inevitable recognition errors of the ASR
 - ▶ Not concern in this paper

- Omission is serious in spoken language

- Traditional methods are unsuitable for ASR transcriptions

Why FST

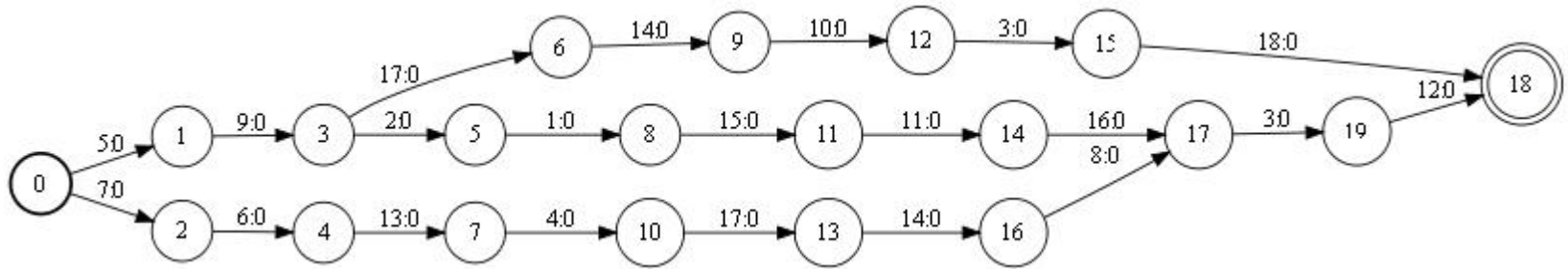
- Powerful in delivering the sequence information
 - ▶ Words sequence

- powerful in tolerance of ASR errs
 - ▶ Whole sentence

Outline

- Task
- Problems in NLU
- **FST in NLU**
- Problems in DM
- DBN in DM
- The whole system
- Conclusion

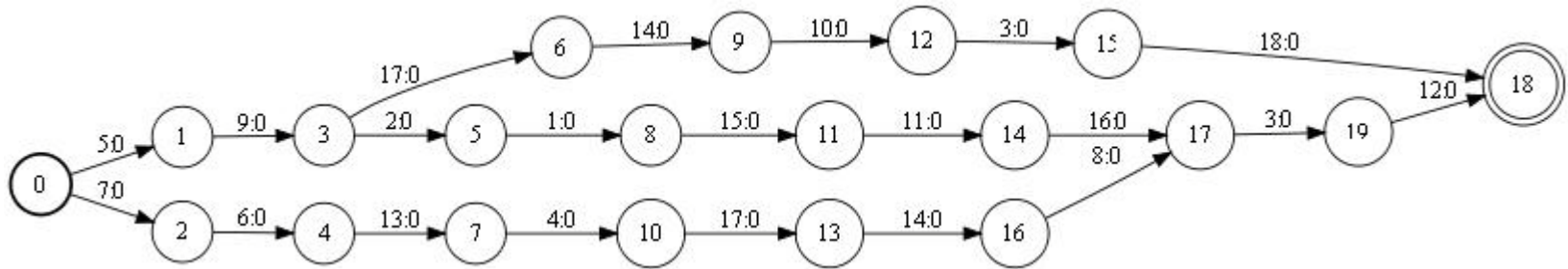
FST



Corpus:

我能兑换美元么
 请问可以兑换欧元吗
 我能在这购买港元吗

FST



Corpus:

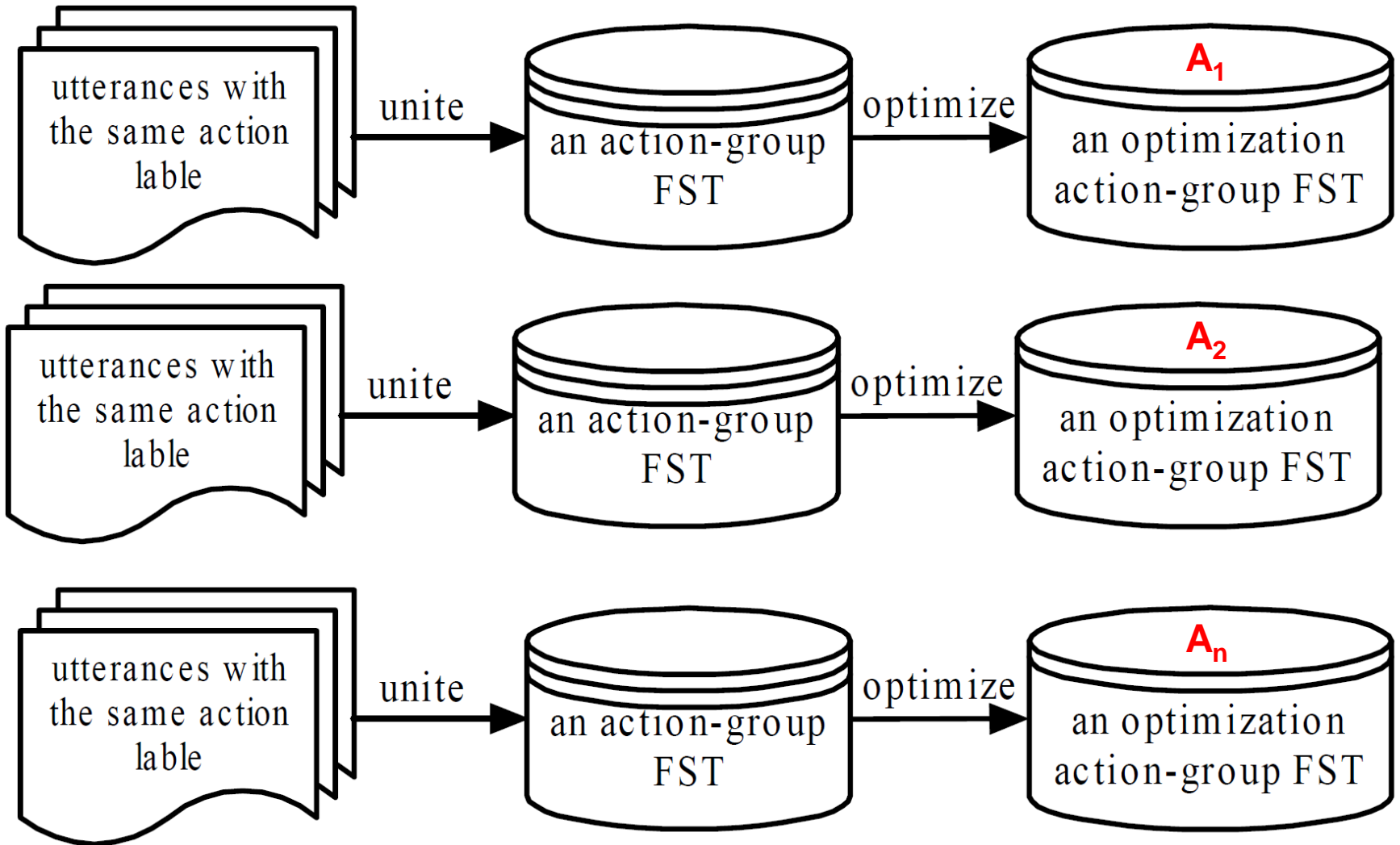
我能兑换美元么
 请问可以兑换欧元吗
 我能在这购买港元吗

Give information (currency_buy_type) and request (buy=?)

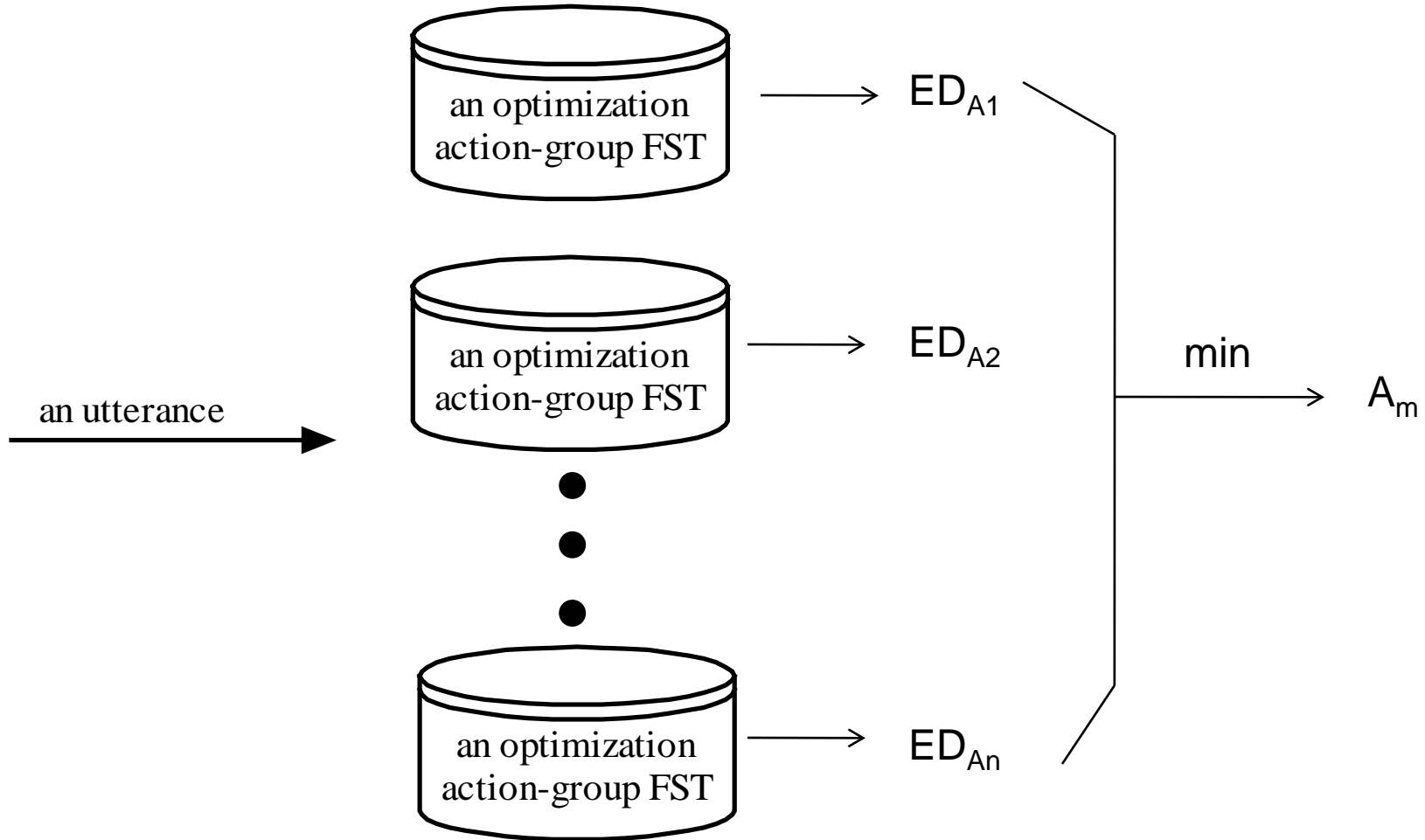
FST model



FST model



Decode



Experiments on FST model

	classification	attribute-value pairs extraction		
	precision	precision	recall	F-measure
action-group FST model	92.75%	92.05%	87.57%	89.75%
keyword matching	90.58%	96.25%	83.24%	89.27%

Experiments on FST model

	classification	attribute-value pairs extraction		
	precision	precision	recall	F-measure
action-group FST model	92.75%	92.05%	87.57%	89.75%
keyword matching	90.58%	96.25%	83.24%	89.27%

Outline

- Task
- Problems in NLU
- FST in NLU
- **Problems in DM**
- DBN in DM
- The whole system
- Conclusion

Problems in DM

- The objective function
 - ▶ Least dialog turns?

- The train data
 - ▶ Need a large amount of labeled data?

Problems in DM

- The objective function
 - ▶ Least dialog turns?

What time will the plan arrive shanghai?

Please give me your identity card and money.

Problems in DM

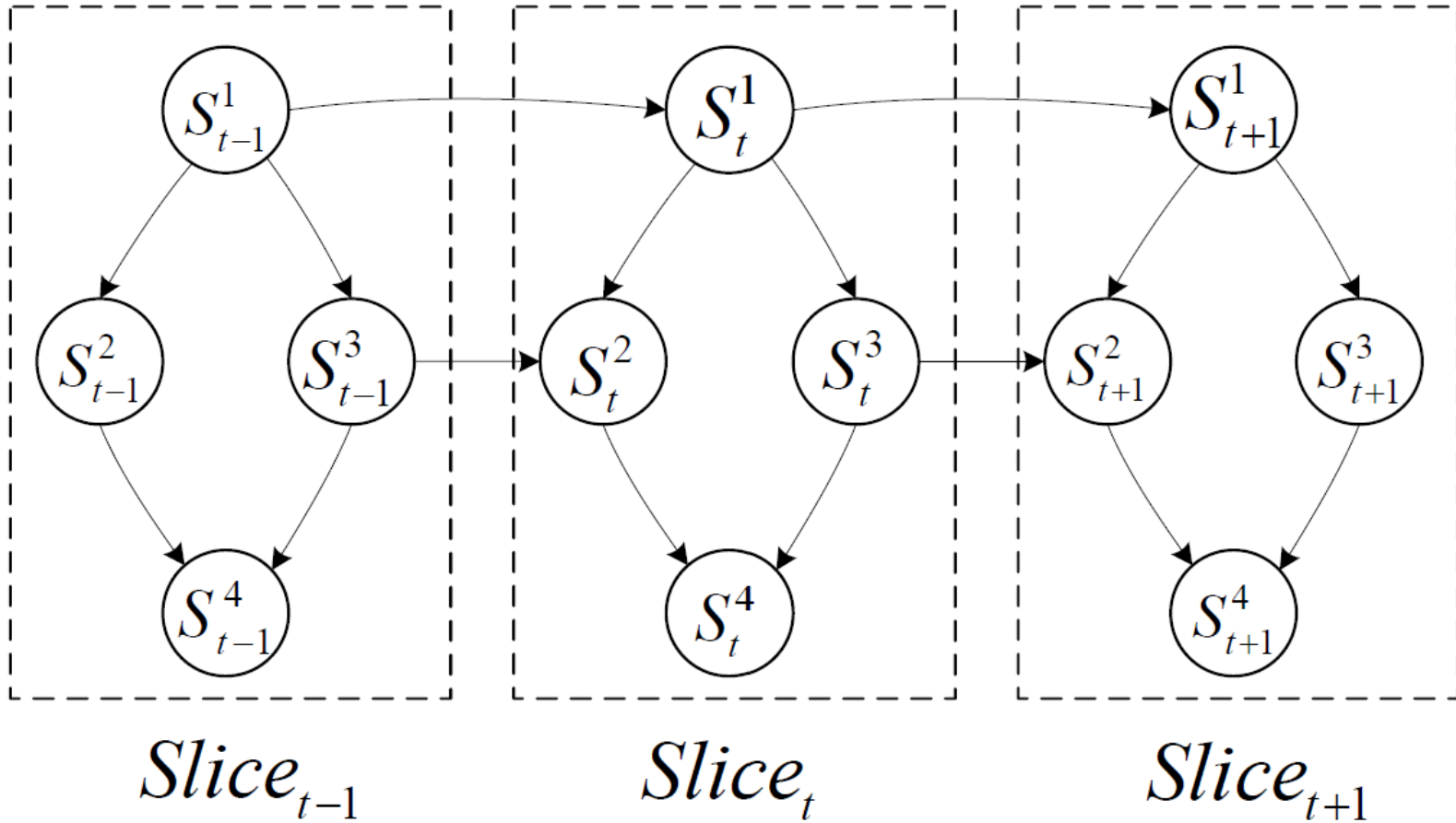
- The objective function
 - ▶ Least dialog turns?

- The train data
 - ▶ Need a large amount of labeled data?

Why DBN

- Need small amount of data
- Inference dependent on the training data
- Inference can rely on the dialogue history

The topology of DBN



Outline

- Task
- Problems in NLU
- FST in NLU
- Problems in DM
- **DBN in DM**
- The whole system
- Conclusion

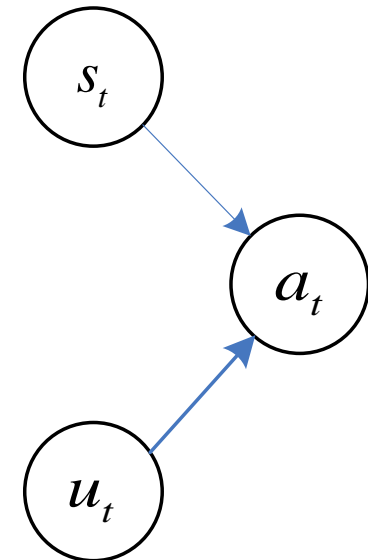
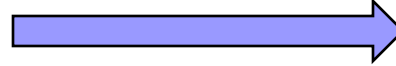
Dialogue management

- Object
 - ▶ The system response (system act)
- Factors should be considered
 - ▶ The dialog history (system state)
 - ▶ The user utterance (user act)

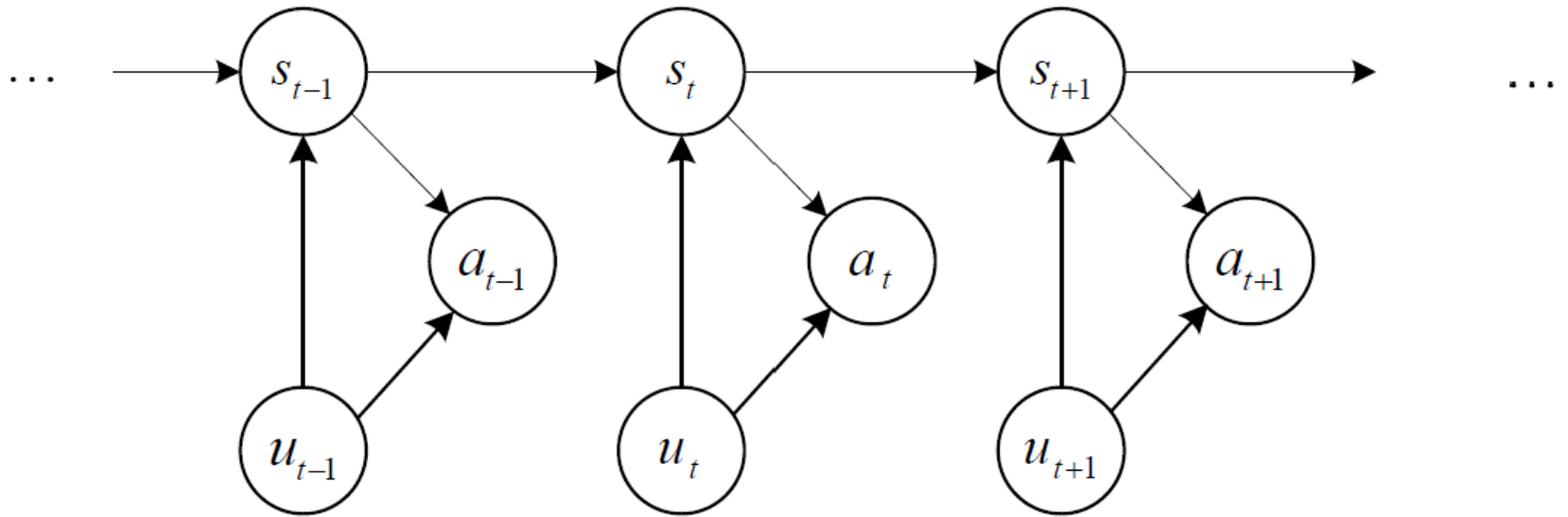
Dialogue management

- Object
 - ▶ The system response (system act)
- Factors should be considered
 - ▶ The dialog history (system state)
 - ▶ The user utterance (user act)

$$P(a_t | s_t, u_t)$$



The DBN Model



A labeled dialog

<1><usr><3><你们有货币兑换业务吗>
 <1><state><0,0,0,0,0,0,>
 <1><sys><7><这里有货币兑换业务>
 <2><usr><12><我想把美元兑换成人民币。当前的兑换率是多少>
 <2><state><1,1,0,0,0,0,>
 <2><sys><11><一百美元兑换六百零九元人民币>
 <3><usr><8><我想兑换四百美元现金>
 <3><state><1,1,0,1,0,0,>
 <3><sys><2><好的请稍等>
 <4><usr><15><NULL>
 <4><state><1,1,0,1,0,0,>
 <4><sys><3><请填一下表>
 <5><usr><9><好>
 <5><state><1,1,0,1,1,0,>
 <5><sys><1><给你钱请拿好>
 <6><usr><7><谢谢>
 <6><state><1,1,0,1,1,1,>
 <6><sys><5><不客气>

DBN model

<1><usr><3>
 <1><state><0,0,0,0,0,0,>
 <1><sys><7>
 <2><usr><12>
 <2><state><1,1,0,0,0,0,>
 <2><sys><11>
 <3><usr><8>
 <3><state><1,1,0,1,0,0,>
 <3><sys><2>
 <4><usr><15>
 <4><state><1,1,0,1,0,0,>
 <4><sys><3>
 <5><usr><9>
 <5><state><1,1,0,1,1,0,>
 <5><sys><1>
 <6><usr><7>
 <6><state><1,1,0,1,1,1,>
 <6><sys><5>

DBN model 1

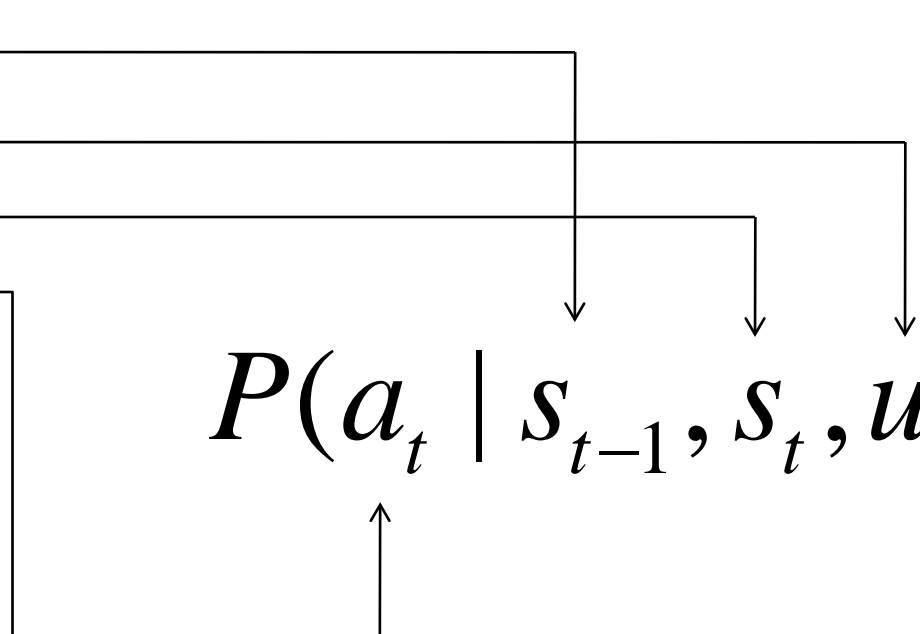
$$P(a_t | s_t, u_t)$$


DBN model

<1><usr><3>
 <1><state><0,0,0,0,0,0,>
 <1><sys><7>
 <2><usr><12>
 <2><state><1,1,0,0,0,0,>
 <2><sys><11>
 <3><usr><8>
 <3><state><1,1,0,1,0,0,>
 <3><sys><2>
 <4><usr><15>
 <4><state><1,1,0,1,0,0,>
 <4><sys><3>
 <5><usr><9>
 <5><state><1,1,0,1,1,0,>
 <5><sys><1>
 <6><usr><7>
 <6><state><1,1,0,1,1,1,>
 <6><sys><5>

DBN model 2

$$P(a_t \mid s_{t-1}, s_t, u_t)$$



DBN model

<1><usr><3>
 <1><state><0,0,0,0,0,0,>
 <1><sys><7>
 <2><usr><12>
 <2><state><1,1,0,0,0,0,>
 <2><sys><11>
 <3><usr><8>
 <3><state><1,1,0,1,0,0,>
 <3><sys><2>
 <4><usr><15>
 <4><state><1,1,0,1,0,0,>
 <4><sys><3>
 <5><usr><9>
 <5><state><1,1,0,1,1,0,>
 <5><sys><1>
 <6><usr><7>
 <6><state><1,1,0,1,1,1,>
 <6><sys><5>

DBN model 3



$$P(a_t | u_t)$$

Experiments on DM

	DBN			MDP
	$P(a_t s_t, u_t)$	$P(a_t s_{t-1}, s_t, u_t)$	$P(a_t u_t)$	$P(a_t s_t, u_t)$
precision	90.58%	89.13%	88.41%	81.88%

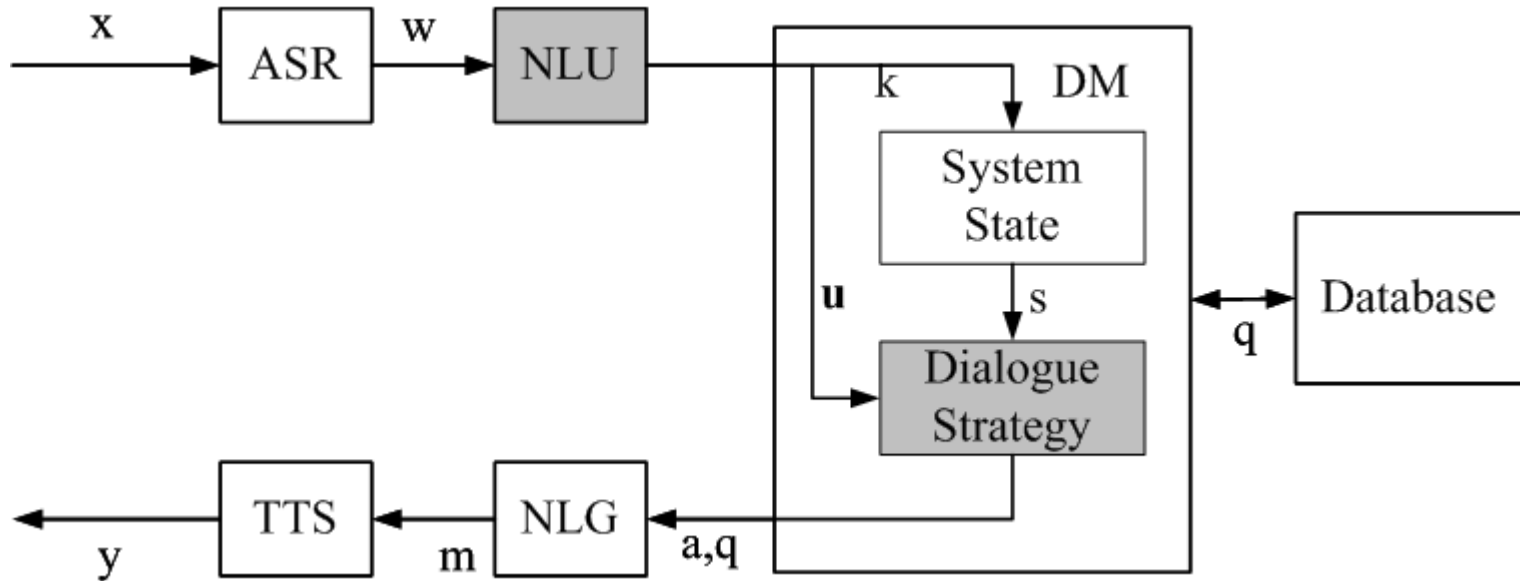
Experiments on DM

	DBN		MDP
	$P(a_t s_t, u_t)$	$P(a_t s_{t-1}, s_t, u_t)$	$P(a_t u_t)$
precision	90.58%	89.13%	88.41%

Outline

- Task
- Problems in NLU
- FST in NLU
- Problems in DM
- DBN in DM
- **The whole system**
- Conclusion

The structure of our spoken dialogue system



对话管理演示系统

EXCHANGE 货币兑换	
CANADA CAD	095.12 08883
CHINA CNY	73.169 609.10
EURO EUR	65.44 06.100
JAPAN JPY	10.900 10.200
SINGAPORE SGD	13.712 12.630
HONG KONG HKD	70.043 64.072
NEW ZEALAND NZD	11.646 10.675
ANGKIA MYR	3.2536 2.18.18

选择场景

退出

USER: 我想换钱
SYS: 您想换成什么货币
USER: 换美元
SYS: 请问您想兑换多少?
USER: 今天的汇率是多少
SYS: 现在的比率是6.4800
USER: 给我换五百美元吧
SYS: 请您填写一下兑换单, 填好了请告诉我。
USER: 填好了
SYS: 好的, 算上手续费, 一共是三千二百五十六
USER: 好的, 给你
SYS: 好的
USER: 谢谢
SYS: 不用客气
USER: 再见
SYS: 再见, 您慢走

字体设置

录音开始

录音结束

确认输入

Change dialog scenes

Input text

Input speech

A user survey of the spoken dialogue system

user types	total dialogues	dialogues end with satisfaction
researcher	30	27
ordinary user	30	22

Outline

- Task
- Problems in NLU
- FST in NLU
- Problems in DM
- DBN in DM
- The whole system
- **Conclusion**

Conclusion

- FST model can be competent for NLU
 - ▶ The model can be modified easily when data increase
 - ▶ The labeled data can be used in other dialogue scene

- DBN model performs great in DM
 - ▶ It needs a little mount of data
 - ▶ The training is fast



THANKS!