What we are doing and what we can do in coming years?

Chengqing ZONG

National Laboratory of Pattern Recognition,

Institute of Automation, Chinese Academy of Sciences

cqzong@nlpr.ia.ac.cn

No.95, Zhongguancun East Road
Beijing 100190, China

http://www.ia.ac.cn
Tel. No.: +86-10-6255 4263

What we are doing?

- Spoken Language Translation (SLT)

  Many MT methods have been experienced for many years.

  + Rule-based

  + Example-based

  + Interlingua-based

  + Statistical MT

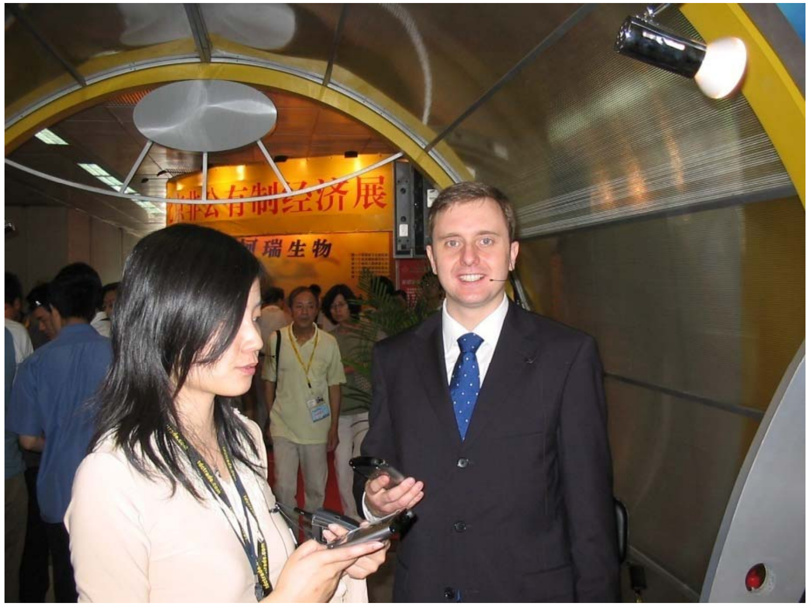
  ……

Spoken Language Translation

- We have been working with SLT for many years and have experienced many methods.

Spoken Language Translation

Spoken Language Translation



(1)

(2)

Spoken Language Translation

- Now almost all researchers use statistical MT based on very large scale corpus.

Spoken Language Translation

- **In IWSLT-2007**

Performance of the CE clean text translation:

BLEU Score ＝ 0.3648

It was the best one according to the results of human rankings.

CE Clean	
SYSTEM	% BETTER
NLPR	37.6
I2R	37.0
ICT	34.8
RWTH	32.4
FBK	30.6
CMU	30.6
UPC	28.3
XMU	28.1
HKUST	25.5
MIT	25.0
NTT	24.6
ATR	24.2
UMD	23.6
DCU	18.6
NUDT	16.1

Table 4 Human Rankings: CE Clean.

Spoken Language Translation

- **In IWSLT-2008**

CT_{CE}

MT	Ranking	NormRank
NLPR	0.5274	2.48
tch.ASR.1	0.4657	2.37
ict.ASR.1	0.3869	2.13
i2r.ASR.1	0.3863	2.11
mitll.ASR.SLF	0.3686	2.00
rwth.ASR.1	0.3423	1.96
dcu.ASR.1	0.3331	1.95
ntt.ASR.1	0.3327	1.87
nict.ASR.1	0.3127	1.89
fbk.ASR.1	0.2585	1.71
tottori.ASR.1	0.2074	1.53

BT_{CE}

MT	Ranking	NormRank
NLPR	0.5255	2.60
tch.ASR.1	0.4900	2.54
ict.ASR.1	0.4668	2.44
i2r.ASR.1	0.4393	2.38
rwth.ASR.1	0.4060	2.20
cmu.ASR.1	0.4051	2.24
dcu.ASR.1	0.3302	2.02
fbk.ASR.1	0.3291	2.01
nict.ASR.1	0.2965	1.83
tubitak.ASR.1	0.2813	1.88
tottori.ASR.1	0.2342	1.66
postech.ASR.1	0.2138	1.58
greyc.ASR.1	0.1468	1.26
qmul.ASR.1	0.1603	1.35

Spoken Language Translation

- **In IWSLT-2009**

IWSLT 2009 Evaluation Campaign

Preliminary Automatic Evaluation Results

last update: 2009/09/11

official evaluation specifications (case+punc) | additional evaluation specifications (no_case+no_punc)

CHALLENGE Chinese-English (CT_CE)

bleu	meteor	f1	prec	recl	wer	per	ter	gtm	nist		bleu	meteor	f1	prec	recl	wer	per	ter	gtm	nist
0.3552	0.6559	0.6956	0.7042	0.6873	0.5344	0.4186	47.5080	0.7101	6.6590	NLPR	0.3717	0.6408	0.6792	0.6912	0.6677	0.5462	0.4237	49.8440	0.6981	6.9972
0.3161	0.5783	0.6403	0.6924	0.5955	0.5651	0.4725	50.9050	0.6469	5.6137	dcu_ASR.1	0.3064	0.5495	0.6110	0.6673	0.5635	0.5894	0.4882	54.2110	0.6280	5.6163
0.3013	0.5907	0.6342	0.6348	0.6336	0.6218	0.4895	57.3050	0.6654	5.7632	fbk.ASR.1	0.2866	0.5608	0.6055	0.6093	0.6018	0.6476	0.5019	60.9660	0.6410	5.9682
0.2859	0.5921	0.6430	0.6549	0.6315	0.6230	0.4859	56.0280	0.6389	5.8057	ict.ASR.20	0.2853	0.5762	0.6215	0.6269	0.6161	0.6438	0.4921	59.7400	0.6366	6.0329
0.2667	0.5834	0.6101	0.5764	0.6480	0.7290	0.5545	66.8850	0.6656	5.2850	nict.ASR.1	0.2570	0.5569	0.5835	0.5519	0.6190	0.7518	0.5636	70.5390	0.6398	5.4834
0.2482	0.5489	0.5910	0.5773	0.6053	0.6943	0.5456	64.8360	0.6136	5.0705	tottori.ASR.1	0.2323	0.5270	0.5624	0.5438	0.5822	0.7239	0.5707	68.9040	0.5947	5.1394
(primary run not (yet) submitted)										ASR.1	(primary run not (yet) submitted)									
0.3644	0.6799	0.7181	0.7253	0.7110	0.5086	0.3919	45.1000	0.7354	6.9772	NLPR	0.3808	0.6692	0.7082	0.7203	0.6965	0.5227	0.3917	47.1550	0.7250	7.3495
0.3691	0.6415	0.6918	0.7235	0.6627	0.5302	0.4235	47.1490	0.7074	6.6291	dcu	0.3675	0.6234	0.6733	0.7077	0.6421	0.5447	0.4292	49.8010	0.7016	6.8250
0.3192	0.6323	0.6715	0.6653	0.6777	0.6015	0.4575	53.5840	0.6976	6.1671	fbk	0.3127	0.6093	0.6492	0.6468	0.6517	0.6210	0.4627	56.6630	0.6816	6.4422
0.3078	0.6310	0.6805	0.6868	0.6743	0.6042	0.4506	52.8110	0.6825	6.2825	ict	0.3185	0.6259	0.6685	0.6673	0.6697	0.6162	0.4413	55.8890	0.6959	6.6640
0.2970	0.6309	0.6479	0.6051	0.6973	0.7109	0.5201	63.9190	0.7008	5.7255	nict	0.2867	0.6082	0.6249	0.5844	0.6715	0.7390	0.5273	67.7210	0.6796	5.8891
0.2797	0.5971	0.6306	0.6092	0.6536	0.6590	0.5099	61.3850	0.6592	5.5309	tottori	0.2716	0.5791	0.6092	0.5846	0.6359	0.6807	0.5218	65.2680	0.6506	5.6891
(primary run not (yet) submitted)										nus	(primary run not (yet) submitted)									

Spoken Language Translation

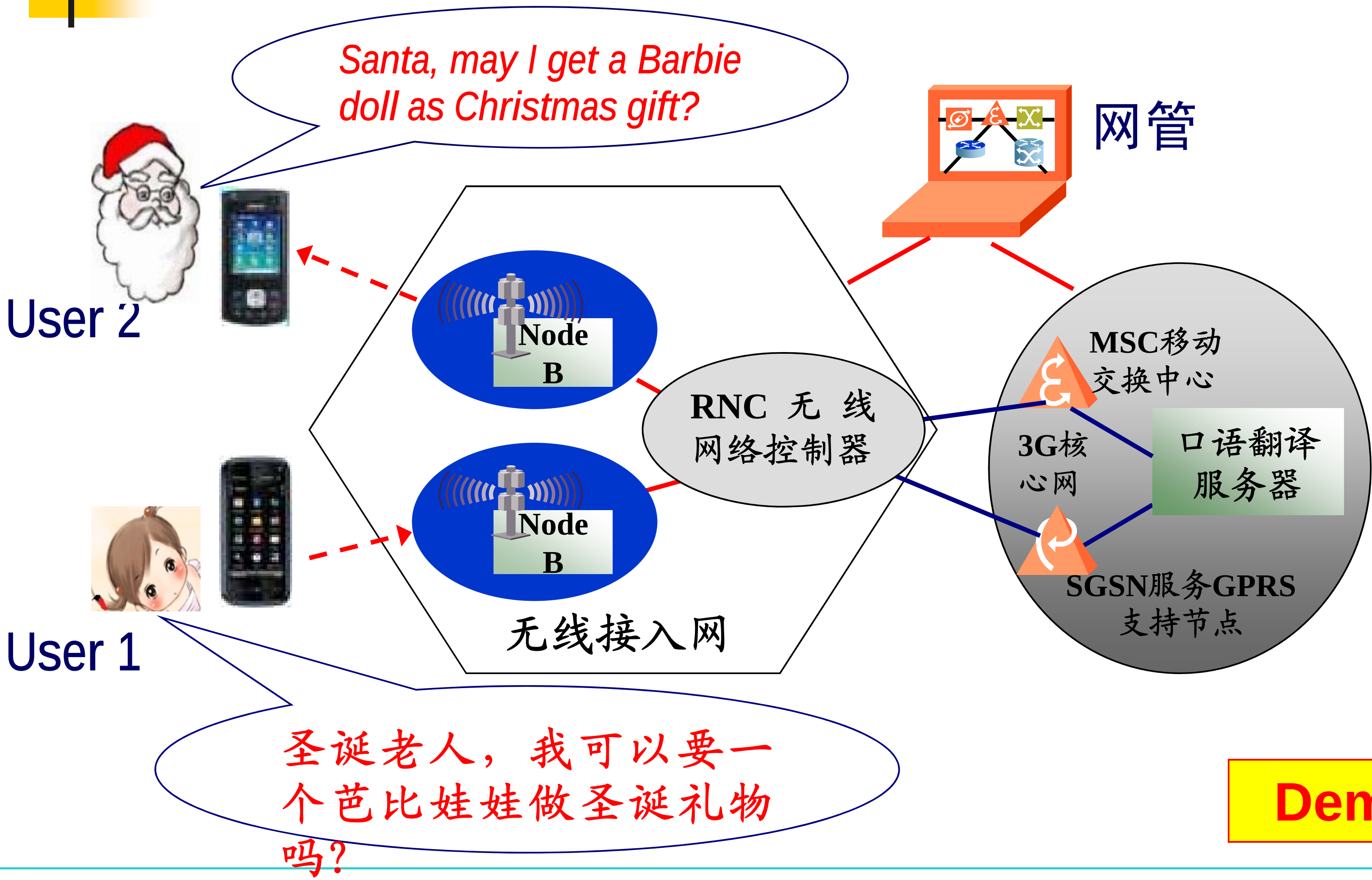
CHALLENGE English-Chinese (CT_EC)																				
bleu	meteor	f1	prec	recl	wer	per	ter	gtm	nist		bleu	meteor	f1	prec	recl	wer	per	ter	gtm	nist
0.3566	X	0.6479	0.6806	0.6183	0.5457	0.4094	48.6150	0.7007	6.3964	NLPR	0.3756	X	0.6456	0.6891	0.6074	0.5534	0.4151	48.3440	0.7022	6.5527
0.3583	X	0.6282	0.6693	0.5918	0.6065	0.4314	51.8730	0.6814	6.0257	nict.ASR.1	0.3544	X	0.6086	0.6573	0.5666	0.6344	0.4563	53.7920	0.6665	5.9179
0.3337	X	0.6127	0.6324	0.5943	0.6121	0.4484	54.3870	0.6857	6.1168	fbk.ASR.1	0.3333	X	0.5963	0.6228	0.5720	0.6321	0.4646	55.7060	0.6707	6.1445
0.3282	X	0.6010	0.6368	0.5690	0.5941	0.4586	52.8690	0.6697	5.8528	dcu.ASR.1	0.3315	X	0.5941	0.6337	0.5592	0.6105	0.4690	54.0680	0.6588	5.9079
0.2901	X	0.5805	0.6138	0.5506	0.6354	0.4772	56.1420	0.6472	5.6574	ict.ASR.20	0.2935	X	0.5762	0.6082	0.5474	0.6530	0.4841	57.3280	0.6469	5.8012
0.2214	X	0.4516	0.4100	0.5025	0.8513	0.6447	80.8210	0.6399	4.5091	tottori.ASR.1	0.2256	X	0.4635	0.4279	0.5055	0.8440	0.6353	79.3250	0.6095	4.6631
(primary run not (yet) submitted)										nus.ASR.1	(primary run not (yet) submitted)									
0.4075	X	0.6897	0.7207	0.6613	0.4917	0.3614	43.5420	0.7480	7.0349	NLPR	0.4304	X	0.6907	0.7346	0.6518	0.4982	0.3671	43.5880	0.7507	7.2234
0.4005	X	0.6727	0.6846	0.6612	0.5488	0.3832	47.7820	0.7602	7.0276	fbk	0.4007	X	0.6597	0.6791	0.6413	0.5663	0.3991	49.1040	0.7362	7.0960
0.3886	X	0.6690	0.7199	0.6248	0.5145	0.3904	45.1190	0.7247	6.6597	ict	0.3998	X	0.6703	0.7235	0.6244	0.5244	0.3932	45.5670	0.7263	6.8031
0.3842	X	0.6705	0.7233	0.6249	0.5509	0.3944	46.5470	0.7296	6.3863	nict	0.3816	X	0.6537	0.7121	0.6041	0.5761	0.4164	48.3900	0.7085	6.3149
0.3734	X	0.6554	0.6740	0.6379	0.5652	0.4021	49.2990	0.7333	6.8211	dcu	0.3756	X	0.6516	0.6720	0.6323	0.5802	0.4107	50.3370	0.7283	6.9420
0.2759	X	0.5500	0.5150	0.5900	0.7421	0.5382	68.6970	0.6814	5.3888	tottori	0.2754	X	0.5473	0.5176	0.5805	0.7494	0.5415	68.3460	0.6681	5.5090
(primary run not (yet) submitted)										nus	(primary run not (yet) submitted)									

Spoken Language Translation

BTEC Chinese-English (BTEC_CE)

bleu	meteor	f1	prec	recl	wer	per	ter	gtm	nist		bleu	meteor	f1	prec	recl	wer	per	ter	gtm	nist
0.4969	0.7266	0.7604	0.7798	0.7420	0.4104	0.3555	33.6680	0.7252	7.6961	NLPR	0.4897	0.6917	0.7298	0.7573	0.7043	0.4540	0.3803	37.3750	0.7123	8.0287
0.4481	0.6808	0.7297	0.7741	0.6901	0.4404	0.3897	35.8560	0.6966	6.7795	nus	0.4402	0.6383	0.6894	0.7433	0.6428	0.4965	0.4267	40.4900	0.6750	6.9657
0.4595	0.6725	0.7274	0.7810	0.6808	0.4383	0.3938	35.7040	0.6956	6.3841	i2r	0.4526	0.6351	0.6897	0.7489	0.6392	0.4924	0.4341	40.0060	0.6679	6.3964
0.4053	0.6618	0.6974	0.7050	0.6900	0.5007	0.4241	42.1670	0.6936	7.0527	uw	0.3963	0.6216	0.6630	0.6838	0.6434	0.5491	0.4540	45.8720	0.6734	7.3015
0.4237	0.6447	0.7166	0.8017	0.6477	0.4568	0.4175	36.2630	0.6683	5.0626	dcu	0.4197	0.5978	0.6735	0.7774	0.5941	0.5102	0.4530	40.4900	0.6477	4.7111
0.3955	0.6419	0.6973	0.7407	0.6586	0.4846	0.4280	39.3680	0.6686	6.0958	bmrc	0.3943	0.5962	0.6541	0.7134	0.6040	0.5340	0.4615	43.6950	0.6437	6.0490
0.4014	0.6076	0.6653	0.7143	0.6226	0.4921	0.4378	41.4800	0.6768	6.1194	lium	0.3818	0.5571	0.6207	0.6887	0.5649	0.5554	0.4920	46.1450	0.6374	5.9042
0.3538	0.6269	0.6806	0.7141	0.6502	0.4997	0.4466	40.5900	0.6344	5.8624	tokyo	0.3544	0.5803	0.6343	0.6748	0.5984	0.5513	0.4782	45.7210	0.6188	6.0947
0.3529	0.6266	0.6838	0.7184	0.6523	0.5196	0.4486	41.8620	0.6593	6.0473	upv	0.3513	0.5799	0.6388	0.6846	0.5987	0.5713	0.4885	47.0820	0.6411	6.2323
0.3563	0.6226	0.6817	0.7230	0.6449	0.5080	0.4507	41.5820	0.6459	5.8408	ict	0.3479	0.5819	0.6398	0.6888	0.5973	0.5703	0.4890	46.6280	0.6251	5.9914
0.3151	0.6169	0.6569	0.6465	0.6676	0.5590	0.4760	48.0710	0.6478	6.3834	tottori	0.2935	0.5680	0.6092	0.6074	0.6110	0.6252	0.5209	54.0070	0.6193	6.6263
0.2795	0.5537	0.6125	0.6374	0.5896	0.5923	0.5324	51.6090	0.5964	5.6571	greyc	0.2773	0.5098	0.5635	0.5906	0.5388	0.6550	0.5788	57.2120	0.5653	5.9270

Spoken Language Translation



Demo

Spoken Language Translation

Unfortunately, it doesn't work in most real situation.

Spoken Language Translation

Unfortunately, it doesn't work in most real situation.

Why?

Spoken Language Translation

Unfortunately, it doesn't work in most real situation.

Why?

There are so many ill-formed expressions, out-of-vocabularies, unknown cultural custom, unknown world knowledge ...

Real Dialogue Corpus Collection

The first edition now consists of 792 dialogs belonging to tourism domain, which are selected from more than 14,000 spontaneous telephone recordings in real scenarios.

Hotel	Restaurant	Airport	Overall
206	263	323	792

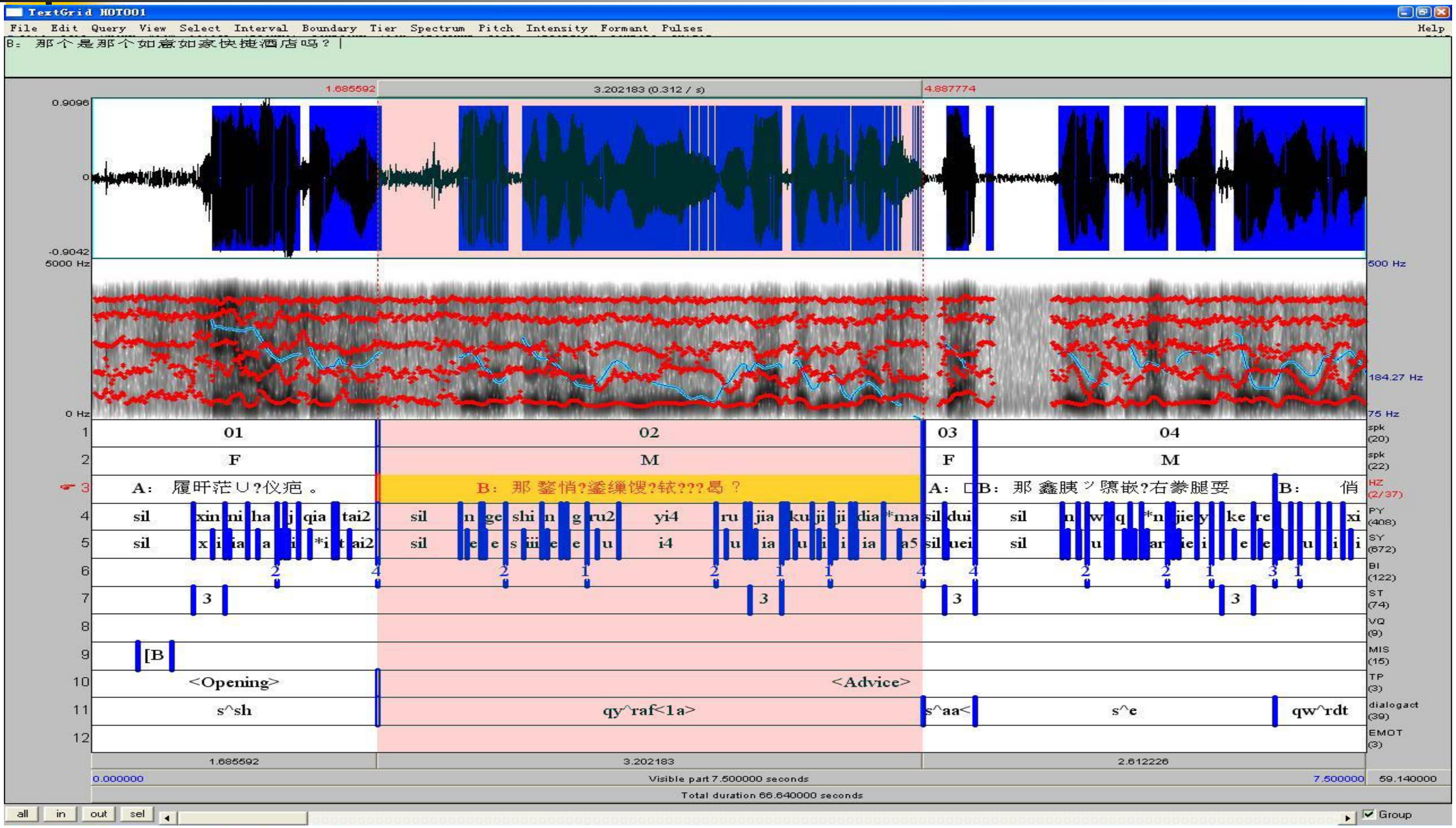
This version will be extended to 800-1,000 dialogs and released at the end of this year.

Real Dialogue Corpus Collection

Tags:

- Speaker Gender
- Orthographic Transcription
- Chinese Syllable
- Chinese phonetic transcription
- Prosodic boundary
- stress of the sentence
- Non-speech sounds
- Voice quality
- Topic
- Expressive emotion
- Dialectal accent of the speaker

Real Dialogue Corpus Collection



An Annotated Example using Praat

Questions

- How big of the corpus size is adequate, 10,000; 100,000; 1M ……?

Questions

- **How big of the corpus size is adequate, 10,000; 100,000; 1M ……?**


Questions

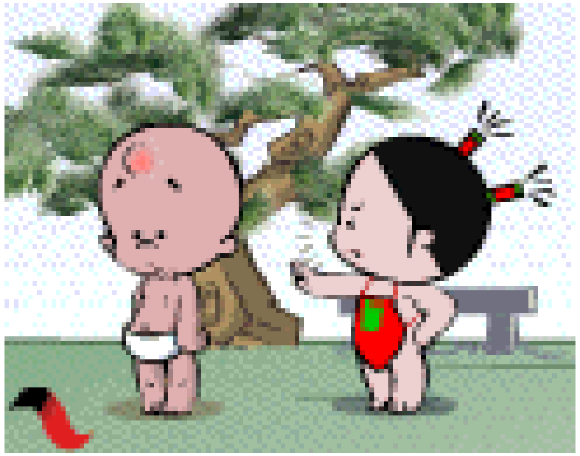
- **How big of the corpus size is adequate, 10,000; 100,000; 1M ……?**

The system needs the ability to learn from a small size corpus and generalize the learnt knowledge.

Questions

- As a human being, how we talk with a foreigner and understand each other?
  - Speech (tone, stress, rhythm …)
  - Gesture
  - Face expression

  … …



Questions

- As a human being, how we talk with a foreigner and understand each other?

  - Speech (tone, stress, rhythm …)
  - Gesture
  - Face expression

  … …

Besides text, the speech and body language information should be used in parsing and translation model. Interaction is necessitated.

Questions

- Can all knowledge for MT be learnt from corpus?

Questions

- Can all knowledge for MT be learnt from corpus?

No

Questions

- Can all knowledge for MT be learnt from corpus?

No

Interaction between speaker and system is necessitated. The system needs to learn from the procedure of human-machine interaction.

Questions

- Even if a system can correctly translate a sentence, can a listener always correctly understand the speaker's intention and meanings?

Questions

- Even if a system can correctly translate a sentence, can a listener always correctly understand the speaker's intention and meanings?

Sometimes not!

Questions

e.g, food name/ menu translation:

(1) 馒头

Questions

e.g, food name / menu translation:

(1) 馒头

steamed bread

Questions

e.g, food name/ menu translation:

(1) 馒头

steamed bread

(2) 夫妻肺片

Questions

e.g, food name/ menu translation:

(1) 馒头

steamed bread

(2) 夫妻肺片

Piece of wife and husband's lung

Questions

e.g, food name / menu translation:

(1) 馒头

steamed bread

(2) 夫妻肺片

Piece of wife and husband's lung

(3) 童子鸡

Questions

e.g, food name/ menu translation:

(1) 馒头

steamed bread

(2) 夫妻肺片

Piece of wife and husband’s lung

(3) 童子鸡

Children chicken without sex life

Questions

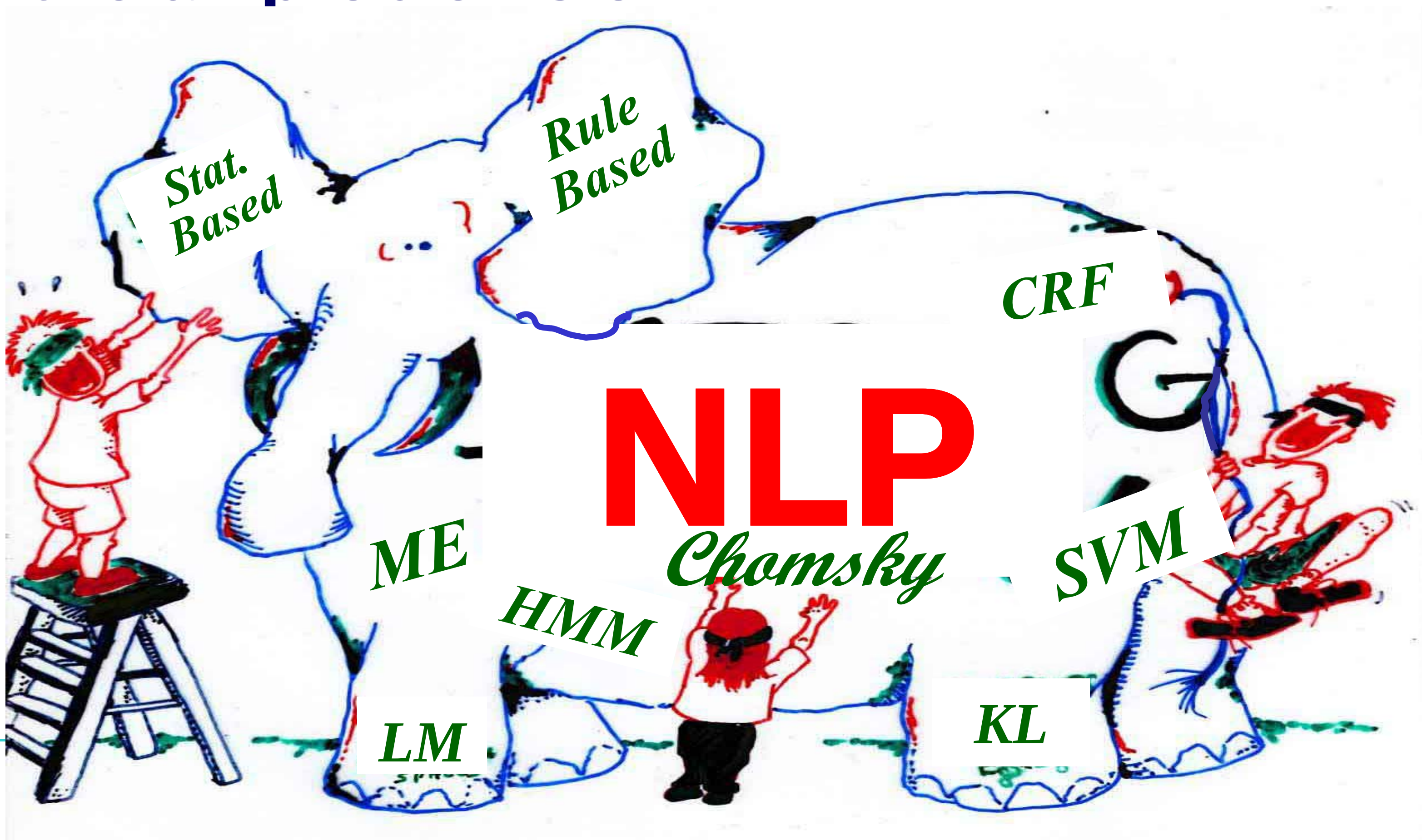
The culture translation or explanation sometimes becomes necessitated.

Questions

- Do the current methods can finally solve the all problems of NLP ?

Questions

- Do the current methods can finally solve the all problems of NLP ?



What we can do in coming years?

In summary, theoretically, study the new methods and approaches to spoken language understanding and translation:

What we can do in coming years?

In summary, theoretically, study the new methods and approaches to spoken language understanding and translation:

- **Rich information joint methods to understand the meaning of an utterance or a dialog**

What we can do in coming years?

In summary, theoretically, study the new methods and approaches to spoken language understanding and translation:

- Rich information joint methods to understand the meaning of an utterance or a dialog
- Meaning based machine translation

What we can do in coming years?

In summary, theoretically, study the new methods and approaches to spoken language understanding and translation:

- **Rich information joint methods to understand the meaning of an utterance or a dialog**
- **Meaning based machine translation**
- **Culture translation and explanation**

What we can do in coming years?

In summary, theoretically, study the new methods and approaches to spoken language understanding and translation:

- Rich information joint methods to understand the meaning of an utterance or a dialog
- Meaning based machine translation
- Culture translation and explanation
- Incremental knowledge learning

What we can do in coming years?

For application:

What we can do in coming years?

For application:

- **Development of practical SLT systems**

What we can do in coming years?

For application:

- **Development of practical SLT systems**
- **Dialogue information extraction**

What we can do in coming years?

For application:

- Development of practical SLT systems
- Dialogue information extraction
- Dialog summarization

What we can do in coming years?

For application:

- Development of practical SLT systems
- Dialogue information extraction
- Dialog summarization
- Multimodal human-machine interaction

Thanks

谢谢!