DEMOCRAT: Deciding between Multiple Outputs Created by Automatic Translation

Author：Menno van Zaanen & Harold Somers

Professor：陳嘉平

Reporter：陳逸昌

Outline

- Introduction
- Related work
- Method
- Evaluation
- Discussion
- Conclusions

Introduction(1/2)

- Many translation systems available free on the Internet.
- It attempt to combine the best of outputs to produce output than any individual systems.

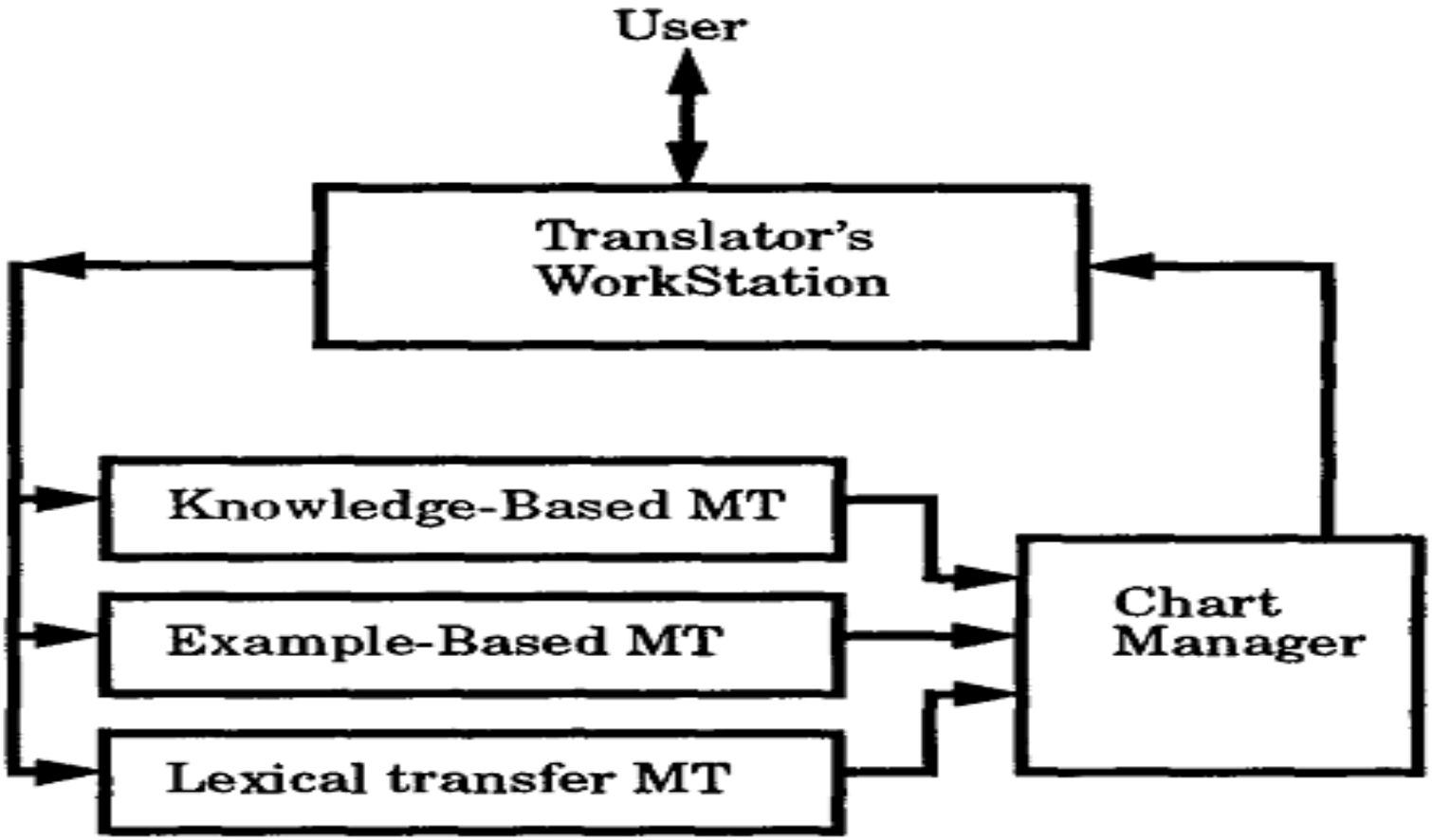
Introduction(2/2)

- The idea to combine outputs from multiple MT systems seems to have been first suggested by Bangalore(2001).
- They combining the outputs results better than any of the individual contributing outputs.

Multi-engine MT (MEMT)

- The multi-engine approach to MT was pioneered by Frederking and Nirenburg (1994)
- They pass the output from three independent MT architectures- KBMT, EBMT and a lexical transfer system.

Multi-engine MT (MEMT)



Consensus translation

- The Bangalore system using a “progressive multiple alignment” technique found in the biological sciences.
- A least-cost traversal of the lattice corresponds to selecting the consensus by majority vote.
- Bangalore employ a simple n-gram language models to select not clear consensus words.

déme			direcciones	impulsoras por favor	a	área	de	middletown
déme			direcciones	por favor	a	área		
déme			direcciones	conductores por favor	al	área		middletown
déme		las	direcciones	que conducen satisfacen	al	área	de	middletown
déme	que	las	direcciones	tend en cia a gradan	al	área	de	middletown
*****			**********			****		**********

Figure 2: Result of aligning different translations for the English sentence *give me driving directions please to middletown area*

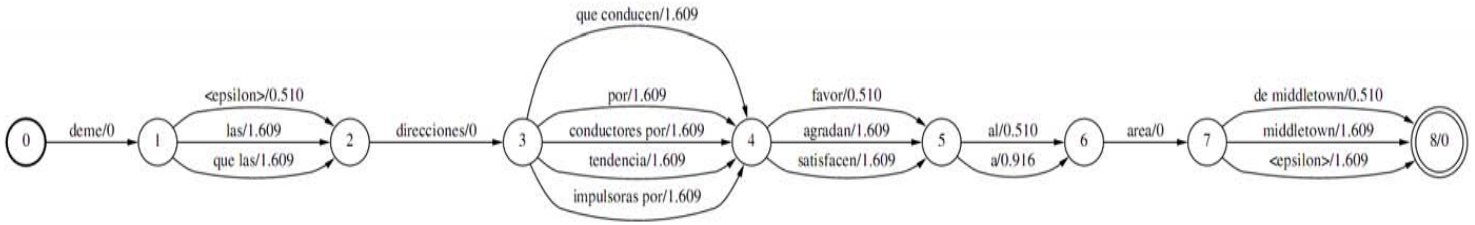


Figure 3: Lattice representation of the result of the multiple alignment. The weights on the arcs are negative logarithm of the probability that word.

DEMOCRAT

- The DEMOCRAT system wished to experiment with an algorithm requires nothing more than the raw outputs.
- If it works for a variety of language pairs, it would be instantly applicable to any situation.

Method

- All the other approaches require an element of training the system on language model.
- The system incorporates an attempt to align the multiple inputs, that doesn’t rely on any pre-processing.

Method

- DEMOCRAT select and combine the ”best” sequence of words from them to give one consensus translation.
- The system assume that if many MT systems use the same word or phrase, it’s likely to be good.

Method

- The first phase aligns all the outputs from MT in pairs. (using Levenshtein distance)
- The second phase builds a graph that contains the input sentences.
- The third phase walks through the graph and generates the output sentence.

Method

- They selection is based on the frequency .
- If no single edge with the highest frequency, then currently edge is chosen in a pseudo-random manner.

Example

- 3 MT systems' outputs
  1. Approval of the official report of the preceding meeting
  2. Approval of the verbal process of the preceding meeting
  3. Approbation of the minutes of the previous session

Example

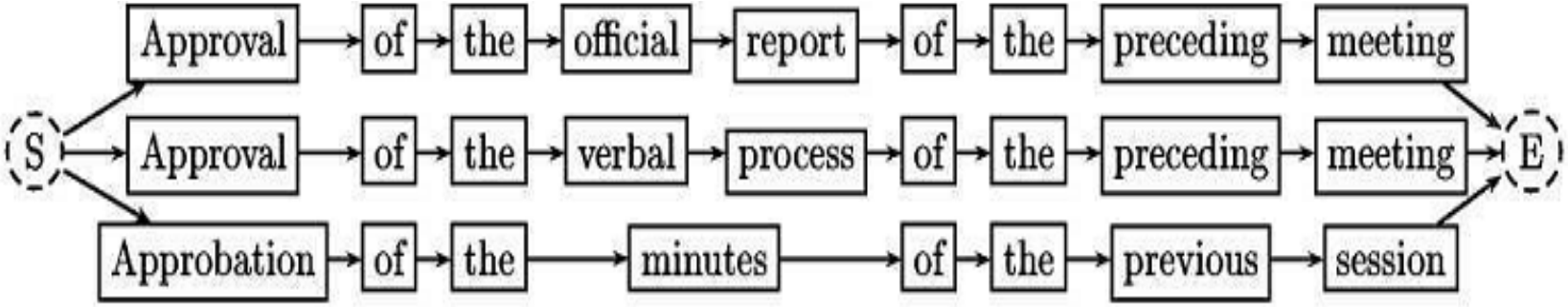


Figure 1: Initial graph of three outputs

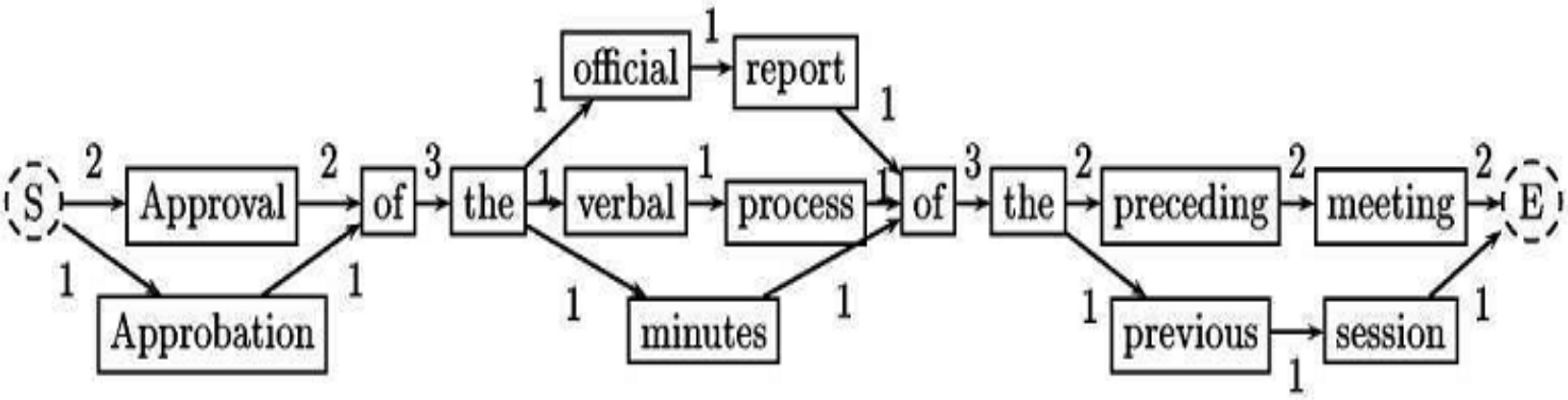


Figure 2: Result of compression of figure 1

Evaluation

- DEMOCRAT to be as nearly a “plug-and-play” system as possible.
- In this paper they focus on results of three French-English and one English-German results.

Corpus

- French-English
  1. the web pages of the Marseilles Tourist Office (mar)
  2. the Europarl corpus of European Parliament Proceedings (euFE)
  3. The Jules Verne's 20000 liecues sous les mers (jv)
- English-German
  1. The Europarl corpus(euEG).

Corpus statistical

Abbr.	Source	No. of sent.	Av. sent. length
mar	Web site	94	19.33
euFE	Europarl corpus	107	27.73
jv	Project Gutenberg	200	23.00
euEG	Europarl corpus	101	22.09

Online translation system

- Five online translation systems
  1. Babelfish

     http://world.altavista.com/babelfish
  2. Freetranslation

     http://www.freetranslation.com
  3. Systran

     http://systranbox.com/systran
  4. TranslateRU

     http://www.translate.ru/eng/
  5. Worldlingo

     http://www.worldlingo.com/en/products_services/worldlingo_translator.html

Results

System	mar		euFE		jv		euEG	
	BLEU	F-score	BLEU	F-score	BLEU	F-score	BLEU	F-score
ba	0.23033	0.25396	0.17827	0.21214	0.19091	0.22997	0.10611	0.17575
fr	0.15087	0.21803	0.14106	0.18859	0.15175	0.20908	0.10233	0.17121
sy	0.22111	0.25466	0.18277	0.21428	0.18866	0.23053	0.10385	0.17491
tr	0.18683	0.22940	0.18458	0.21511	0.15342	0.21159	0.10789	0.18274
wo	0.22732	0.25223	0.17889	0.21266	0.16674	0.21711	0.10334	0.17141
de.0 ba fr	0.18927	0.23790	0.16157	0.20141	0.17298	0.22086	0.10543	0.17321
de.1 ba sy	0.22317	0.25386	0.18047	0.21331	0.19292	0.23043	0.10692	0.17591
de.2 ba tr	0.19340	0.23756	0.18502	0.21566	0.16554	0.21768	0.10770	0.18293
de.3 ba wo	0.23077	0.25456	0.17900	0.21281	0.17859	0.22398	0.10618	0.17446
de.4 fr sy	0.18416	0.23982	0.15939	0.20015	0.17439	0.22165	0.10462	0.17308
de.5 fr tr	0.16635	0.22592	0.16753	0.20498	0.15407	0.21144	0.10912	0.17710
de.6 fr wo	0.19167	0.23835	0.15468	0.19782	0.16025	0.21409	0.09590	0.16868
de.7 sy tr	0.18917	0.23616	0.18548	0.21662	0.16383	0.21984	0.11104	0.18454
de.8 sy wo	0.22784	0.25515	0.18093	0.21361	0.17805	0.22553	0.10606	0.17416
de.9 tr wo	0.19057	0.23832	0.17476	0.21047	0.14702	0.21134	0.11022	0.17615
de.10 ba fr sy	0.20976	0.25220	0.17922	0.21242	0.18745	0.22897	0.10711	0.17679
de.11 ba fr tr	0.19772	0.24646	0.17294	0.21063	0.17029	0.22325	0.11145	0.17975
de.12 ba fr wo	0.21441	0.25016	0.17801	0.21191	0.18054	0.22629	0.10477	0.17478
de.13 ba sy tr	0.21208	0.25171	0.18307	0.21457	0.17910	0.22596	0.10865	0.17789
de.14 ba sy wo	0.22935	0.25464	0.17859	0.21255	0.18238	0.22730	0.10764	0.17465
de.15 ba tr wo	0.21503	0.25104	0.18132	0.21340	0.16400	0.22010	0.11196	0.17924
de.16 fr sy tr	0.20043	0.25059	0.17396	0.21131	0.16894	0.22189	0.12255	0.18616
de.17 fr sy wo	0.20571	0.25168	0.17771	0.21222	0.17599	0.22682	0.10628	0.17559
de.18 fr tr wo	0.19162	0.24397	0.16883	0.20987	0.16436	0.22013	0.11727	0.18268
de.19 sy tr wo	0.21011	0.24921	0.18264	0.21435	0.15892	0.22007	0.11017	0.17828
de.20 ba fr sy tr	0.20793	0.25511	0.17255	0.21217	0.17419	0.22481	0.11042	0.18056
de.21 ba fr sy wo	0.21576	0.25234	0.17574	0.21119	0.18048	0.22811	0.10402	0.17486
de.22 ba fr tr wo	0.20877	0.25369	0.16949	0.21100	0.16947	0.22451	0.10469	0.17757
de.23 ba sy tr wo	0.20977	0.25025	0.17824	0.21230	0.16732	0.22312	0.11172	0.17874
de.24 fr sy tr wo	0.20994	0.25536	0.17665	0.21360	0.16803	0.22504	0.11389	0.18366
de.25 ba fr sy tr wo	0.20337	0.25014	0.17704	0.21328	0.16952	0.22554	0.10549	0.17749

Discussion

- DEMOCRAT is quite sensitive to the quality of its input.
- DEMOCRAT does best when the contributing inputs are of good quality.

Discussion

- Many on-line MT systems are actually derived from the same underlying engine like Systran.
- When ran DEMOCRAT with several systems derived from the same underlying engine, it was rarely able to come first in the evaluation were often very close.

Future work

- If any of the translations are particularly poor, this can adversely affect out result.
- Keeping track of which MT systems' inputs get used more often.

conclusions

- They have implemented a simple system to compile consensus translation from the output of free on-line MT systems.
- There is not a single individual best MT system.
- DEMOCRAT doesn’t require any training material, and can work in any language.