TrAva – a tool for evaluating Machine Translation – pedagogical and research possibilities
Belinda Maia, Diana Santos, Luís Sarmento & Anabela Barreiro - Linguateca

Why MT Matters

- Social reasons
- Political reasons
- Commercial importance
- Scientific and philosophical interest

Useful bibliography


Machine Translation (MT) – a few dates

- 1947 Warren Weaver - his ideas led to heavy investment in MT
- 1959 Bar-Hillel - considers FAHOMT - FULLY AUTOMATIC HIGH QUALITY MACHINE TRANSLATION philosophically impossible
- 1964 ALPAC Report - on limitations of MT > withdrawal of funds
- late 1970s the CEC purchase of SYSTRAN and beginning of EUROTRA project.
- Upward trend in the 1970s and 1980s
- Today: MT technology - high-end versus low-end systems
- MT and the Internet

Some Popular Misconceptions about MT

- False: MT is a waste of time because you will never make a machine that can translate Shakespeare.
- False: There aren't any MT systems which can translate. The spirit is willing, but the flesh is weak: and automatic translation is still in its infancy.
- False: Currently, the quality of translation you can get from an MT system is very low. This makes them useless in practice.
- False: MT threatens the jobs of translators.
- False: The Japanese have developed a system that you can talk to on the phone. It translates what you say into Japanese, and translates the other speaker's replies into English.
- False: There is an amazing South American Indian language with a structure of such logical perfection that it solves the problem of designing MT systems.
- False: MT systems are machines, and buying an MT system should be very much like buying a car.

Some Facts about MT

- True: MT is useful. The MERT systems have been in daily use for years, saving time by translating American diplomatic cables from English into Spanish. This word processing system can automatically translate a page of text into another language in a few minutes.
- True: MT can translate written English, but not spoken English. One application is the translation of legal texts.
- True: MT systems can vary in their output, from a literal translation to a more natural one, which may be preferred in some contexts.
- True: MT systems are being used in various industries, such as legal, medical, and financial.
- True: MT systems are being used in various industries, such as legal, medical, and financial.
- True: MT systems are being used in various industries, such as legal, medical, and financial.
- True: MT systems are being used in various industries, such as legal, medical, and financial.
- True: MT systems are being used in various industries, such as legal, medical, and financial.
- True: MT systems are being used in various industries, such as legal, medical, and financial.
MT architectures – Arnold et al

- Direct architecture - simple grammatical rules + a large lexical and phrasal database
- Transfer architecture - more complex grammar with an underlying approach of transformational-generative theory + considerable research into comparative linguistics in the two languages involved
- Interlingua architecture - L1 > a ‘neutral language’ (real, artificial, logical, mathematical…) > L2

Major Methods, Techniques and Approaches today

- Statistical vs. Linguistic MT
  - assimilation tasks: lower quality, broad domains – statistical techniques predominate
  - dissemination tasks: higher quality, limited domains – symbolic techniques predominate
- Rule-based vs. Example-based MT
- Transfer vs. Interlingual MT
- Multi-Engine MT
- Speech-to-Speech Translation

MT – present & future uses

- ‘Gist’ translation
- Ephemeral texts with tolerant users
- Human aided MT
  - Domain specific
  - Linear sentence structure
  - Pre-edited text or ‘controlled language’
  - Post-editing
- Improvement of MT – particularly for restricted domains and registers

MT and the Human translator

- MT is less of a threat to the professional human translator than English
- MT can encourage people’s curiosity for texts in languages they do not understand > and lead to human translation
- MT can be a tool for the human translator
- Professional translators can learn to work with and train MT

PoloCLUP’s experiment

- Background
  - Master’s seminar in Semantics and Syntax
  - Wish to raise students’ awareness of the strengths and weaknesses of MT
  - Wish to develop their interest in MT as a tool
  - Need to improve their knowledge of linguistics.
  - Availability of free MT online
  - Automation of process provided by computer engineer
Phase 1 - METRA

- Translation using 7 online MT programmes
- EN > PT
- PT > EN
- At present this tool is getting about 60 hits per day!

BOOMERANG

- This tool submits a text for translation – and back-translation – and back-translation…. Until it reaches a fixed point
- This shows that the rules programmed for one language direction do not always correspond to the other language direction
EVAL > TrAva

- Informal class experiment led to a useful research tool
- Several versions of EVAL
  - Different types of classification of input
  - Different explanations of errors of output
- Production, correction and re-correction of procedure interesting

TrAva - procedure

- Online EN > PT MT using 4 MT systems:
  - Free Translation
  - Systran
  - ETS Server
  - Amikai
- Researcher chooses area for analysis – e.g.
  - ambiguity
  - lexical and structural mismatches
  - Homographs and polysemous lexical items
  - syntactic complexity
  - multiword units: idioms and collocations
  - anaphora resolution

TrAva - procedure

- Selection of 'genuine' examples from BNC, Reuter’s corpus, newspapers etc.
- Possible ‘pruning’ of unnecessary text (some systems accept limited text)
- No deliberate attempt to confuse the system
- BUT: avoidance of repetitive ‘test suites’

TrAva - procedure

- Sentence submitted to TrAva
- MT results
- Researcher:
  - Classifies part of sentence being examined in terms of the English lexicon or POS (BNC codes)
  - Examines results
  - Explains errors in terms of Portuguese grammar
Access to work done

- Researcher may access work done and review it
- Teacher / administrator can access student work and give advice
- FAQ
Present situation

• METRA and BOOMERANG are all free to use online at:
  • http://poloclup.linguateca.pt/ferramentas
• TraVa is free to use online at:
  • http://www.linguateca.pt/trava/
• The corpus CorTA – over 1000 sentences + 4 MT versions available for consultation at: http://www.linguateca.pt/

Conclusions

• It has been a successful experiment
• It is useful pedagogically
  – As linguistic analysis
  – As appreciation of MT
• It has interesting theoretical implications > emphasis on ‘real’ sentences and recognition of interconnection of lexicon + syntax + context

Conclusions

• Further work needs to be done on the classifications
  – E.g. the analysis of ‘error’ as ‘lexical choice’ needs to be able to combine with other possible reasons for error
• A lone researcher can use it to examine a restricted area
• BUT – a large team is needed to overhaul a system properly

Homographs and Polysemy

• Homographs = words with same spelling but different syntactic use
• Polysemy = words with same spelling, but different meaning according to use or context
• BUT – the difference is not as clear-cut as all that
• However – major problem for MT
Complex Noun Phrases

- DETerminante + ADJectivo + Nome
- DETerminante + ADJectivo Composto + ADJectivo + Nome
- DETerminante + ADVérbio (em –ly) + ADJectivo + ADJectivo + Nome
Lexical Bundles

EXAMPLES

- Now let us look at some examples:
  - Homographs
  - Polysemy
  - Complex noun phrases
  - Lexical bundles
### Table 1: Characteristics of the Study Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Range from 20 to 60 years</td>
</tr>
<tr>
<td>Gender</td>
<td>Male and female</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>Standard deviation ± 15 cm</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>Median 63, range 40-90</td>
</tr>
<tr>
<td>BMI</td>
<td>Range from 18.5 to 24.9</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Yes: 30%, No: 70%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Yes: 15%, No: 85%</td>
</tr>
<tr>
<td>Smoking</td>
<td>Yes: 20%, No: 80%</td>
</tr>
<tr>
<td>Alcohol Consumption</td>
<td>No: 80%, Yes: 20%</td>
</tr>
</tbody>
</table>

### Figure 1: Flowchart of the Study Design

1. Participant Selection
2. Baseline Assessment
3. Intervention Phase
4. Follow-up Assessment
5. Data Analysis

### Figure 2: Schematic Diagram of the Experimental Setup

- LED Light Source
- Photocell Sensor
- Data Acquisition System
- Control Unit

---

**Note:** This represents a simplified version of the actual data and figures presented in the document.