Machine Translation in a Nutshell

Cristina España-Bonet

UdS & DFKI

Summer Semester 2018

18th April 2018

Again...

Thanks to Josef van Genabith
for some of the slides!

Outline

What is Translation?

A Definition

Translation is the **conversion of text** from one language (source language, SL) to another (target language, TL)

What is Translation?

A Definition

Translation is the **conversion of text** from one language (source language, SL) to another (target language, TL)

Translation is the **communication of the meaning** of a SL text by means of an equivalent TL text

How we Translate?

Translation Methods I

Word-for-word translation: the SL word order is preserved and the words translated singly by their most common meanings, out of context

Literal translation: the SL grammatical constructions are converted to their nearest TL equivalents, but the lexical words are again translated singly, out of context

How we Translate?

Translation Methods II

Faithful translation: producing the precise contextual meaning of the original within the constraints of the TL grammatical structures

Semantic translation: differs from 'faithful translation' only in as far as it must take more account of the aesthetic value of the SL text

How we Translate?

Translation Methods III

Adaptation: used mainly for plays and poetry; the themes, characters, plots are usually preserved, the SL culture is converted to the TL culture and the text is rewritten

Free translation: producing the TL text without the style, form, or content of the original

How we Translate?

Translation Methods IV

Idiomatic translation: reproducing the 'message' of the original but tends to distort nuances of meaning by preferring colloquialisms and idioms where these do not exist in the original

Communicative translation: rendering the exact contextual meaning of the original in such a way that both content and language are readily acceptable and comprehensible to the readership

How we Translate?

Is it easy/direct for a human?

Several problems, several choices...

How we Translate?

Remember? –Cultures & Language–

https://reportsfromtherock.wordpress.com/page/2/

"Of all the different types of snow, **slabb** is the worst. The word really says it all: sticky, dirty, treacherous and wet. Slabb is created after a sudden rise in temperature and more so when rain falls into a layer of snow. In a second, that bright and cheerful winter wonderland transforms into a slushy pool filled with ice-cold water."

How we Translate?

Translating Culture-Specific Concepts

1. Making up a new word
2. Explaining the meaning of the SL expression in lieu of translating it
3. Preserving the SL term intact
4. Opting for a word in the TL which seems similar to or has the same "relevance" as the SL term

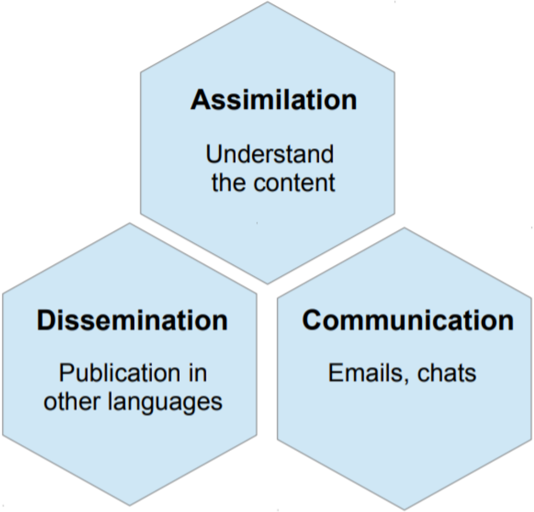
How we Translate?

Should we expect a machine to do it?

Use cases of machine translation

How we Translate?

Machine Translation Uses



How we Translate?

Naïve Machine Translation Taxonomy

Machine Translation systems

$$\overbrace{\text{Human Translation with Machine Support} \qquad \text{Fully Automated Translation} \qquad \text{Machine Translation with Human Support}}$$

How we Translate?

Naïve Machine Translation Taxonomy

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
  - Rule-based systems
- Machine Translation with Human Support

How we Translate?

Naïve Machine Translation Taxonomy

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
  - Rule-based systems
    - Direct MT
    - Transfer MT
    - Interlingua
- Machine Translation with Human Support

How we Translate?

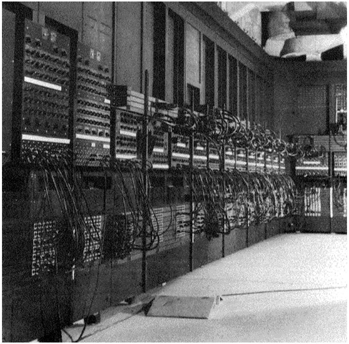
Naïve Machine Translation Taxonomy

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
    - **Neural MT**
    - **Statistical MT**
    - Example-based
  - Rule-based systems
    - Direct MT
    - Transfer MT
    - Interlingua
- Machine Translation with Human Support

Brief History of MT

Early Days



Brief History of MT

Early Days

- **1946** ENIAC, the first digital computer
- **1949** Weaver memorandum, computers could be used to translate natural languages
- **1955** "Translation" (1955), in W.N. Locke and A.D. Booth (eds.), Machine Translation of Languages (MIT Press).

  One naturally wonders if the problem of translation could conceivably be treated as a problem in cryptography. When I look at an article in Russian, I say: 'This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode.'

Brief History of MT

The Rule-based Era

SYSTRAN
beyond language

Brief History of MT

The Rule-based Era

- **1950s** Predominance of rule-based approaches
- **1960s** Predominance of rule-based approaches
- **1966** ALPAC report: general discouragement for MT (in the US)
- **1970s** Predominance of rule-based approaches, firsts commercial systems
- **1980s** Example-based MT proposed in Japan (Nagao), interlingual systems, statistical approaches to speech recognition (Jelinek et al. at IBM)

Brief History of MT

The Statistical Era I

IBM

CANDIDE system

Brief History of MT

The Statistical Era I

Late 80s	Statistical POS taggers, SMT models at IBM, work on translation alignment at Xerox (M. Kay)
Early 90s	Many statistical approaches to NLP in general, IBM's Candide claimed to be as good as Systran
Late 90s	Statistical MT successful as a fallback approach within Verbmobil System (Ney, Och). Wide distribution of translation memory technology (Trados) indicates big commercial potential of SMT

Brief History of MT

The Statistical Era II

Brief History of MT

The Statistical Era II

- 2001 BLEU score for automatic evaluation (Papineni) speeds up evaluation
- 2003 Koehn, Och & Marcu propose Statistical Phrase-Based MT
- 2006 Johns Hopkins workshop on OS factored SMT decoder Moses
- 2007 Google Translate based on SMT

Brief History of MT

The Neural Era

Brief History of MT

The Neural Era

- 2007 GPUs & CUDA
- 2010s Deep Learning leads NLP tasks
- 2013 First paper on NMT (Kalchbrenner & Blunsom)
- 2016 Neural systems state-of-the art in MT evaluation campaigns
- 2017 Google Translate moves to NMT in almost all languages

MT Systems

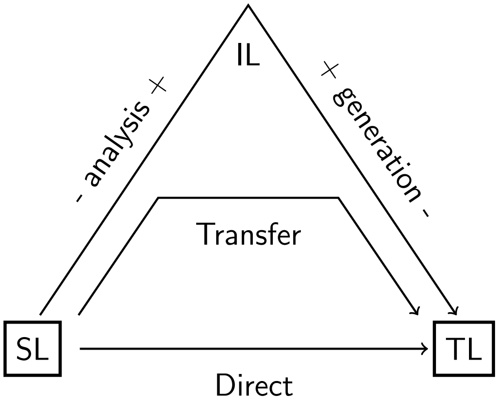
Rule-based MT

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
  - **Rule-based systems**
    - **Direct MT**
    - **Transfer MT**
    - **Interlingua**
- Machine Translation with Human Support

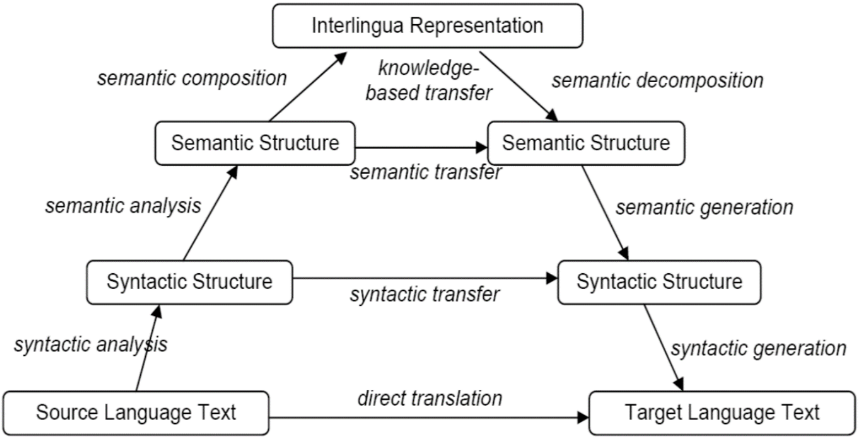
Rule-based MT

Schematic Vauquois triangle



Rule-based MT

Vauquois triangle



Rule-based MT

Direct Translation I

Word by word translation

- Do a little bit of analysis of local source context (but not global or non-local/long range —no parsing, no semantic analysis of sentence)
- May do a little bit of local re-ordering in target (e.g. French adjectives tend to follow noun)
- Requires very large bilingual dictionaries with rules of how to translate a word

Rule-based MT

Direct Translation II

function DIRECT_TRANSLATE_MUCH/MANY(word) **returns** Russian translation

if preceding word is *how* **return** *skol'ko*
else if preceding word is *as* **return** *stol'ko zhe*
else if word is *much*
  if preceding word is *very* **return** nil
  else if following word is a noun **return** *mnogo*
else /* word is many */
  if preceding word is a preposition and following word is a noun **return** *mnogii*
  else return *mnogo*

Figure 25.7 A procedure for translating *much* and *many* into Russian, adapted from Hutchins' (1986, pg. 133) discussion of Panov 1960. Note the similarity to decision list algorithms for word sense disambiguation.

From: Jurafsky & Martin II

Rule-based MT

Direct Translation III

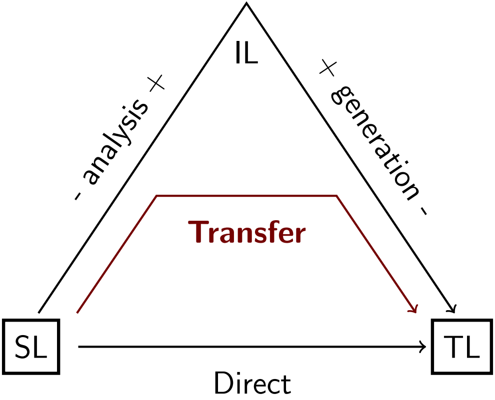
Characteristics (drawbacks!)

- 10s of thousands of manually constructed rules
- Time consuming
- Expensive
- Rule interaction hard to predict
- Any long range phenomena hard to capture

Need some global (syntactic/semantic) analysis

Rule-based MT

Schematic Vauquois triangle



Three phases:

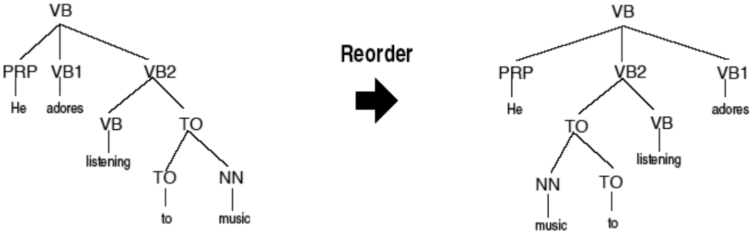
1. **Analysis**: analyse/parse source into syn/sem representation
2. **Transfer**: transform source syn/sem representation into corresponding target syn/sem representation
3. **Generation**: generate target string from target syn/sem representation

Rule-based MT

Transfer-based Translation II

EN: *He adores listening to music*
JA: *He music to listening adores*

From: Jurafsky & Martin II



Rule-based MT

Transfer-based Translation III

Characteristics (drawbacks!)

Need a lot of resources

- Analysis/generation lexica and grammars for SL and TL
- Transfer rule sets for any two languages you want to translate between

Example. To translate between 3 languages: A, B and C

- Need 3 grammars/lexica
- Need 6 transfer grammars: A→B, A→C, B→A, B→C, C→A, C→B

Rule-based MT

Transfer-based Translation III

Characteristics (drawbacks!)

Need a lot of resources

- Analysis/generation lexica and grammars for SL and TL
- Transfer rule sets for any two languages you want to translate between

Example. To translate between n languages: L_1, L_2 ... L_n

- Need n grammars/lexica
- Need $n(n-1)$ transfer grammars
  - 10 languages: 90 transfer systems
  - 100 languages: 9000 transfer systems ...

Characteristics (drawbacks!)

- Need a lot of resources
  - Analysis/generation lexica and grammars for SL and TL
  - Transfer rule sets for any two languages you want to translate between
- Time consuming and expensive to hand-craft
- Not easy to achieve good coverage
- Large rule sets
- Difficult to manage rule interactions

Solutions

- Need a lot of resources
- Time consuming and expensive to hand-craft
- Try to use machine learning?
- Still many resources needed: treebanks, parallel data
- Smart ways of inducing transfer rules from parsed/semantically analysed parallel data

Rule-based MT

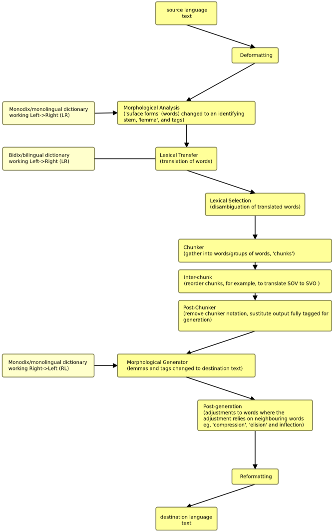
A transfer system: Apertium

Nothing better than their own description:

- Apertium in 5 slides:
  `http://slides.com/allysonallyson/deck#/`
- Workflow diagram:
  `http://wiki.apertium.org/wiki/Workflow_diagram`

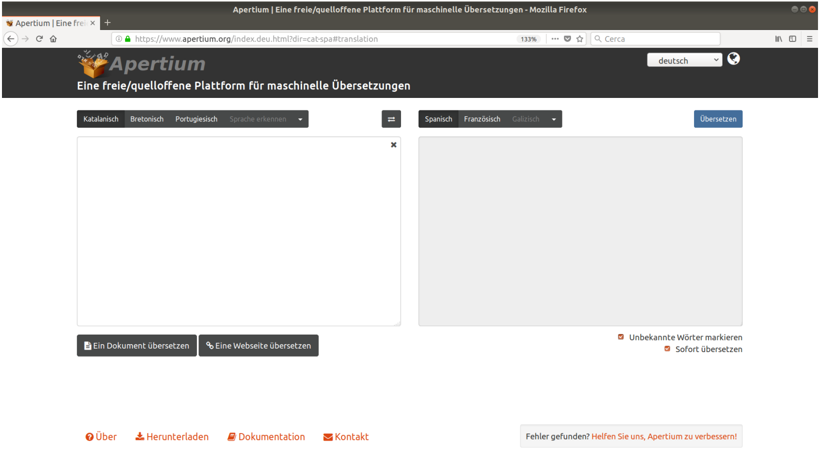
Rule-based MT

Apertium: Workflow diagram



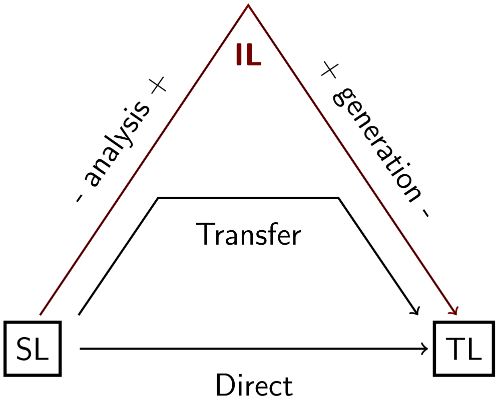
Rule-based MT

Apertium: Interface. Try it!



Rule-based MT

Schematic Vauquois triangle



Rule-based MT

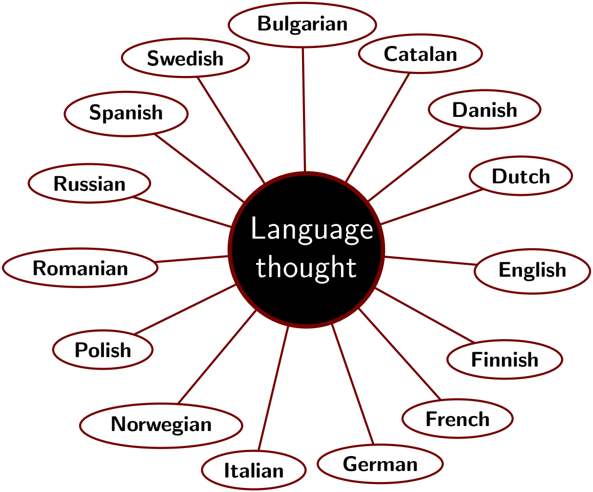
Interlingua Translation I

The $n(n-1)$ complexity in transfer-based MT comes from language specific syntactic and semantic representations

- Maybe we do not need these?
- Maybe we can have instead only one language independent abstract meaning representation?
- The language of thought... interlingua

Rule-based MT

Interlingua Translation II



Rule-based MT

Remember? –Cultures & Language–

- Problem: what is the language of thought?
- Are there universal concepts that work equally well for all languages?

- Problem: what is the language of thought?
- Are there universal concepts that work equally well for all languages?

Different cultures share basic **concepts** and **actions** and communicate them with words

Rule-based MT

Interlingua Translation IV

I'm not showing the *snow* picture again, but...

- Germans say *Wand* for inside (of) wall and *Mauer* for outside (of) wall
- Japanese differentiate between the *younger brother* and the *older brother*
- Spanish use the same word, *dedo* to say *toe* and *finger*

So far no success in coming up with single language independent interlingua

Rule-based MT

Interlingua Translation V

Example: GF. In-domain interlingual translation

- Limiting the number of languages may help
- Limiting the application domain may help
- Successful example: GF system
- (although hibridised with SMT for coverage)

What is GF?

- A **grammar formalism**: a notation for writing grammars.
- A **functional programming language**.

What is GF?

- A **grammar formalism**: a notation for writing grammars.
- A **functional programming language**.

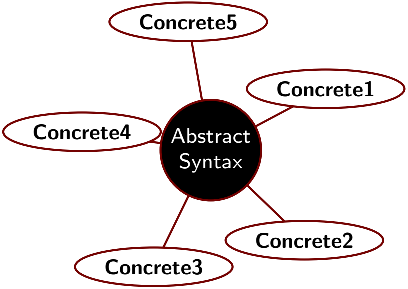
What is a multilingual grammar?

- A definition of a **parsing** and **generation** operations.
- **Concrete syntaxes** for many languages related by a common **abstract syntax**.

Rule-based MT

Abstract and Concrete syntaxes

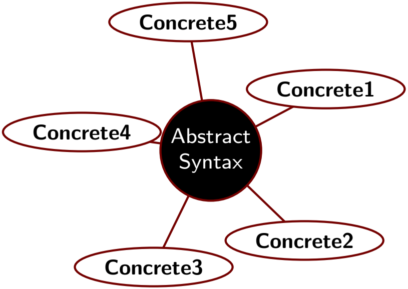
The abstract syntax acts as a **domain-specific interlingua**.



Rule-based MT

Abstract and Concrete syntaxes

The abstract syntax acts as a **domain-specific interlingua**.



Defines not only a linguistic structure but a semantic model for translation with:

- fixed word senses
- proper idioms

Rule-based MT

Translation with GF

Abstract Syntax

Nat : Set
Odd : Exp -> Prop
Gt : Exp -> Exp -> Prop
Sum : Exp -> Exp

Concrete Syntax (ENG)

Nat = "number"
Odd x = "x is odd"
Gt x y = "x is greater than y"
Sum x = "the sum of x"

Concrete Syntax (GER)

Nat = "Zahl"
Odd x = "x ist ungerade"
Gt x y = "x ist grösser als y"
Sum x = "die Summe von x"

Rule-based MT

Translation with GF

Abstract Syntax

Nat : Set
Odd : Exp -> Prop
Gt : Exp -> Exp -> Prop
Sum : Exp -> Exp

parsing

Concrete Syntax (ENG)

Nat = "number"
Odd x = "x is odd"
Gt x y = "x is greater than y"
Sum x = "the sum of x"

Concrete Syntax (GER)

Nat = "Zahl"
Odd x = "x ist ungerade"
Gt x y = "x ist grösser als y"
Sum x = "die Summe von x"

Every even number that is greater than 0 is the sum of two odd numbers

Rule-based MT

Translation with GF

Abstract Syntax

Nat : Set
Odd : Exp -> Prop
Gt : Exp -> Exp -> Prop
Sum : Exp -> Exp

parsing

linearisation

Concrete Syntax (ENG)

Nat = "number"
Odd x = "x is odd"
Gt x y = "x is greater than y"
Sum x = "the sum of x"

Concrete Syntax (GER)

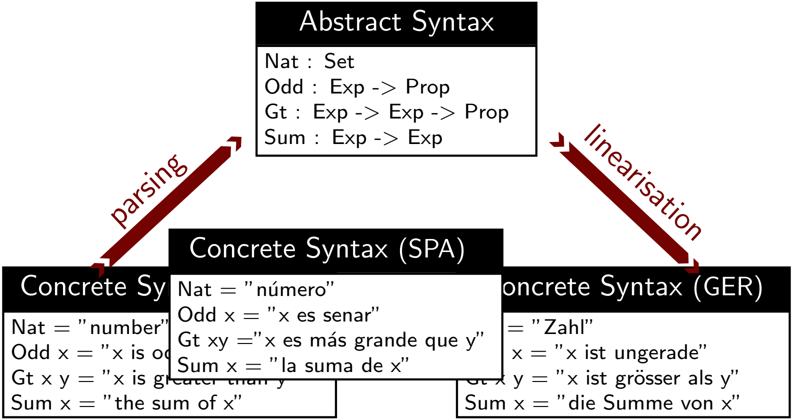
Nat = "Zahl"
Odd x = "x ist ungerade"
Gt x y = "x ist grösser als y"
Sum x = "die Summe von x"

Every even number that is greater than 0 is the sum of two odd numbers

Jede gerade Zahl, die gröβer als 0 ist, ist die Summe von zwei ungerader Zahlen

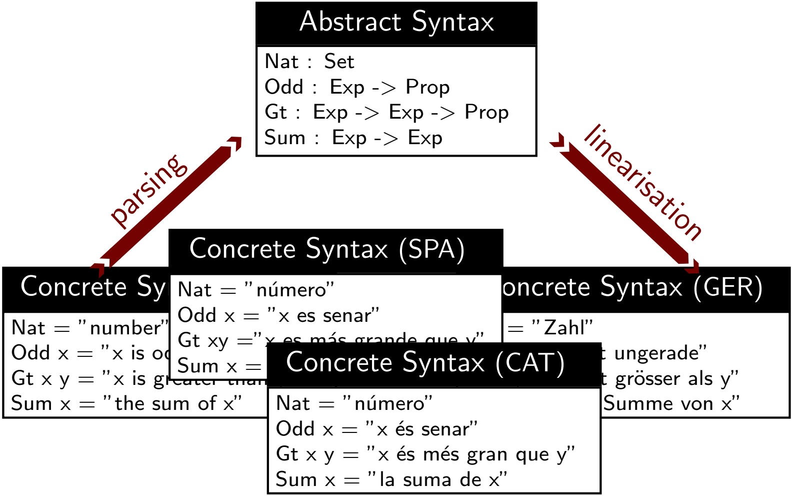
Rule-based MT

Translation with GF



Rule-based MT

Translation with GF



Rule-based MT

GF. Try it!

Grammatical Framework

A programming language for multilingual grammar applications

Use GF

- GF Cloud
- Android app
- Other Demos

- **Download GF**
- GF Eclipse Plugin
- GF Editor Modes

- User Group
- Bug Reports (old)
- Blog

Learn GF

- Google Tech Talk
- QuickStart
- QuickRefCard
- GF Shell Reference
- **GF Summer School**

- The GF Book
- GF Tutorial
- Reference Manual
- Best Practices [PDF]

- Library Synopsis
- Library Tutorial [PDF]
- Coverage Map

Develop GF

- GF Developers Guide
- GF on GitHub
- Contibutions GitHub
- Wiki
- Browse Source Code
- Authors

Develop Applications

- PGF library API (Old Runtime)
- PGF library API (New Runtime)
- GF on Android (new)
- GF on Android (old)

Related to GF

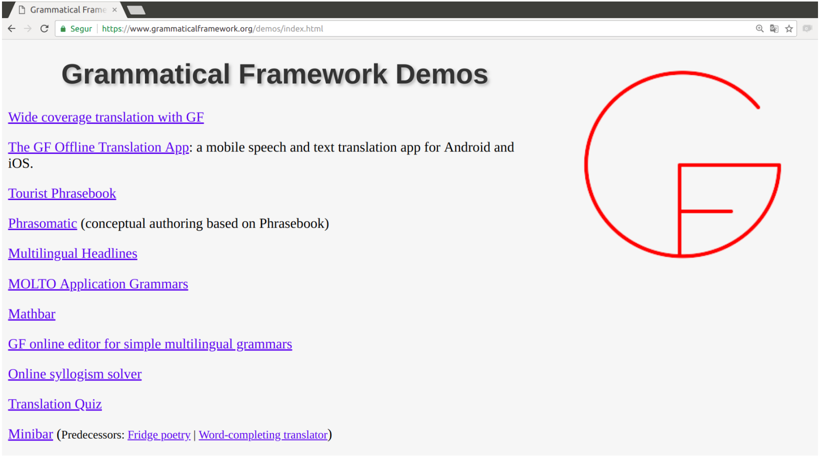
- Publications
- GF Summer Schools
- The REMU Project
- The MOLTO Project
- GF on Wikipedia

- Digital Grammars AB

`https://www.grammaticalframework.org/`

Rule-based MT

GF. Let's go for Minibar



`https://www.grammaticalframework.org/demos/index.html`

Empirical systems

Data-driven Machine Translation

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - **Empirical systems**
    - **Neural MT**
    - **Statistical MT**
    - **Example-based**
  - Rule-based systems
- Machine Translation with Human Support

Empirical systems

Parallel Corpora

- **Data** is the key aspect in empirical systems (by definition!)
- **Parallel corpora** are needed to learn translation models, sometimes **monolingual corpora** are needed to improve fluency
- Parallel corpora are especially **difficult to obtain**
- **Domain-specific corpora** are even more valuable and scarce

Where to find them?

In the context of this course, go to the different evaluation campaings:

- **WMT**: `http://www.statmt.org/wmt18`
- **IWSLT**: `http://workshop2017.iwslt.org`
- **NIST**: `https://www.nist.gov/itl/iad/mig/lorehlt-evaluations`

Data-driven Machine Translation

Neural Machine Translation

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
    - **Neural MT**
    - Statistical MT
    - Example-based
  - Rule-based systems
- Machine Translation with Human Support

The Encoder–Decoder Model (with attention)

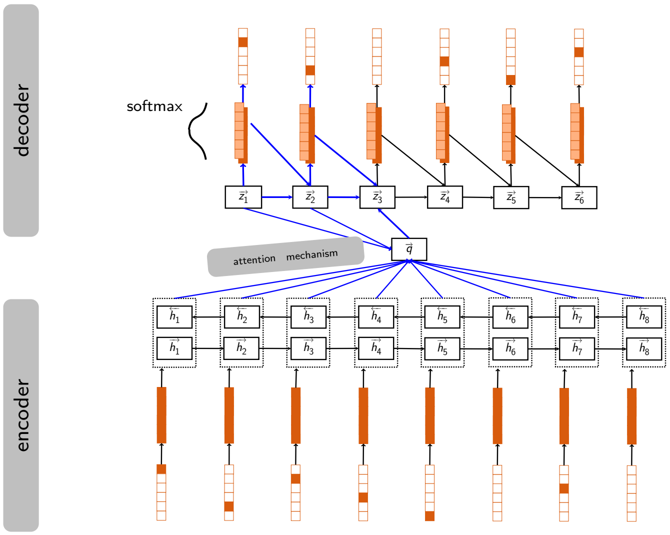
1. encodes a sequence of word vectors into a fixed-sized context vector
2. decodes the fixed-sized vector back into a variable-length sequence

The Encoder–Decoder Model (with attention)

1. encodes a sequence of word vectors into a fixed-sized context vector
2. decodes the fixed-sized vector back into a variable-length sequence

Several NLP tasks use nowadays seq2seq architectures:

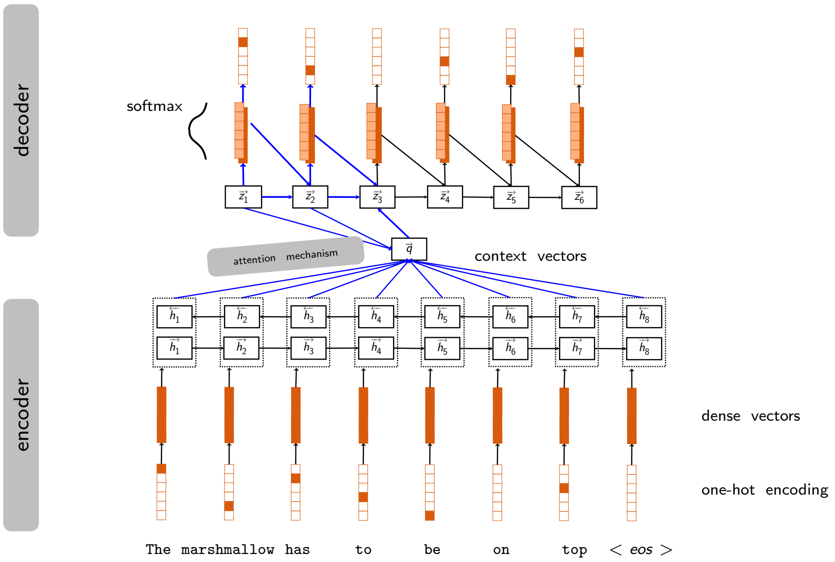
- Machine translation, but also...
- text summarisation, question answering, chatbots, speech recognition...



decoder
softmax
$\vec{z}_1$
$\vec{z}_2$
$\vec{z}_3$
$\vec{z}_4$
$\vec{z}_5$
$\vec{z}_6$
attention mechanism
$\vec{q}$
$\overleftarrow{h}_1$
$\overleftarrow{h}_2$
$\overleftarrow{h}_3$
$\overleftarrow{h}_4$
$\overleftarrow{h}_5$
$\overleftarrow{h}_6$
$\overleftarrow{h}_7$
$\overleftarrow{h}_8$
$\overrightarrow{h}_1$
$\overrightarrow{h}_2$
$\overrightarrow{h}_3$
$\overrightarrow{h}_4$
$\overrightarrow{h}_5$
$\overrightarrow{h}_6$
$\overrightarrow{h}_7$
$\overrightarrow{h}_8$
encoder

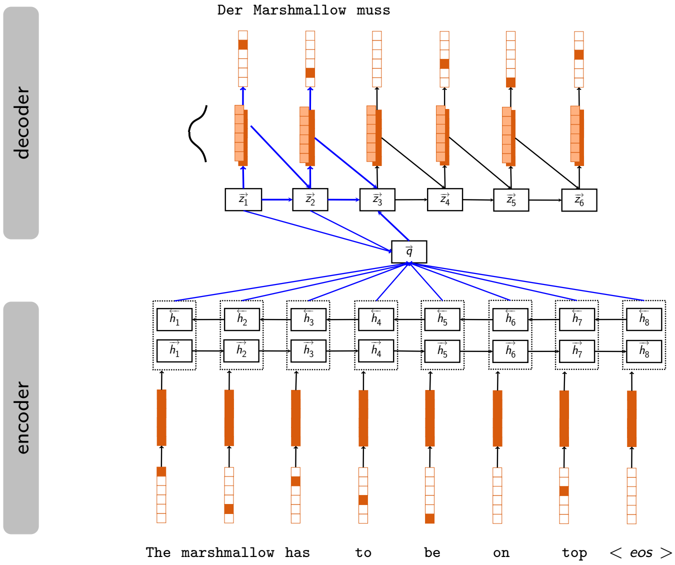
Data-driven Machine Translation

Neural Machine Translation II



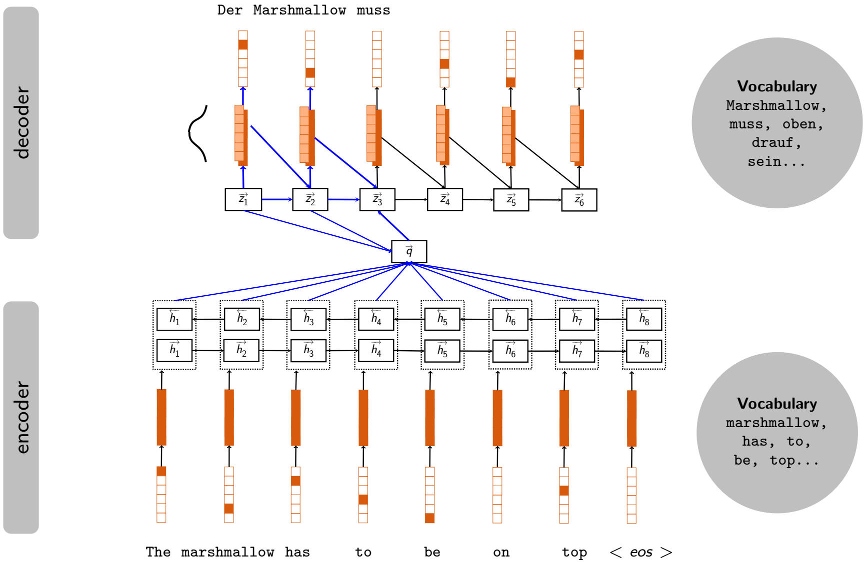
Data-driven Machine Translation

Neural Machine Translation II



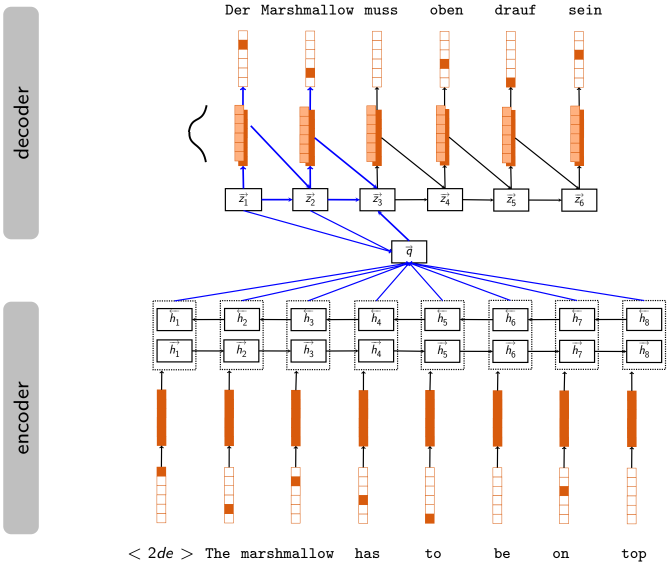
Data-driven Machine Translation

Neural Machine Translation II



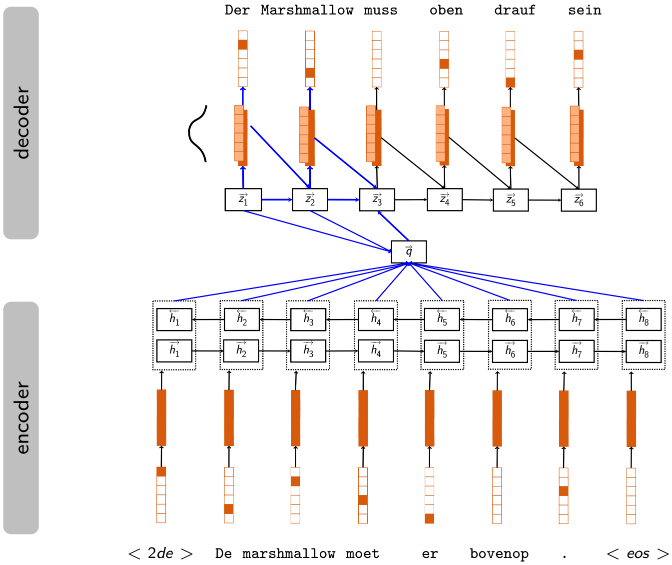
Data-driven Machine Translation

Neural Machine Translation II



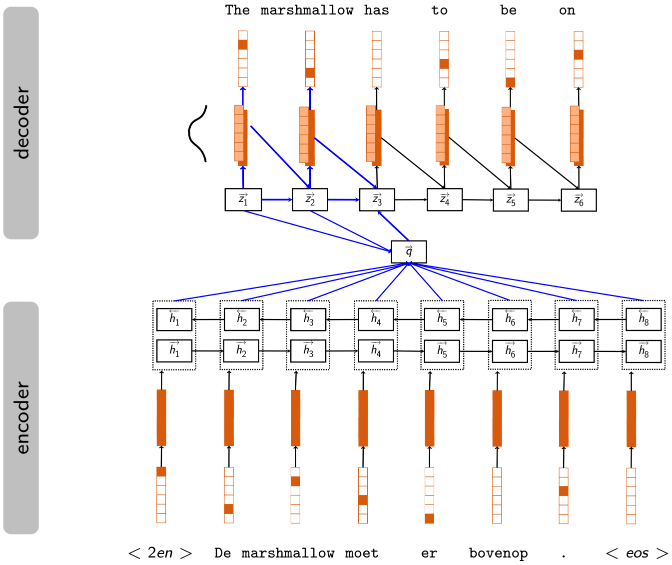
Data-driven Machine Translation

Neural Machine Translation II



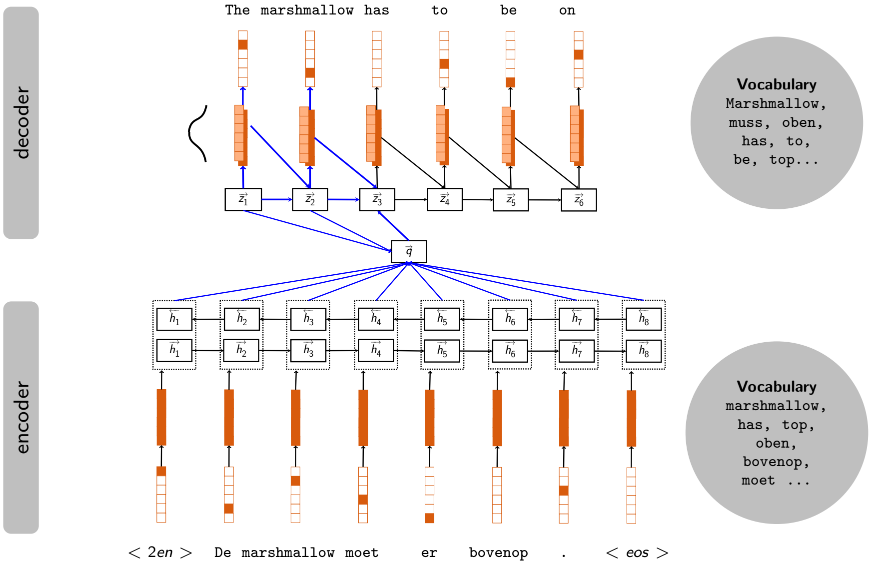
Data-driven Machine Translation

Neural Machine Translation II



Data-driven Machine Translation

Neural Machine Translation II



The Encoder–Decoder Model (with attention)

- and recurrent neural networks (LSTMs, GRUs)
  - OpenNMT, Google (today)
- and convolutional neural networks
  - Facebook

The Transformer Model (almost only attention!)

- OpenNMT, Google (tomorrow?)

Data-driven Machine Translation

Statistical Machine Translation

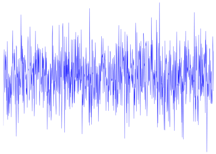
Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
    - Neural MT
    - **Statistical MT**
    - Example-based
  - Rule-based systems
- Machine Translation with Human Support

Data-driven Machine Translation

Statistical Machine Translation I

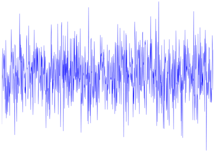
The Noisy Channel as a statistical approach to translation:

Good morning! $\rightarrow$ 

Data-driven Machine Translation

Statistical Machine Translation I

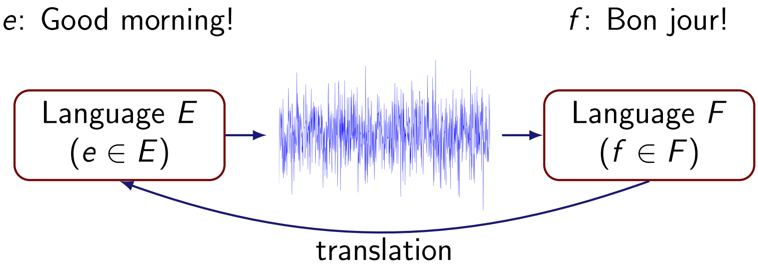
The Noisy Channel as a statistical approach to translation:

Good m  → Bon jour!

Data-driven Machine Translation

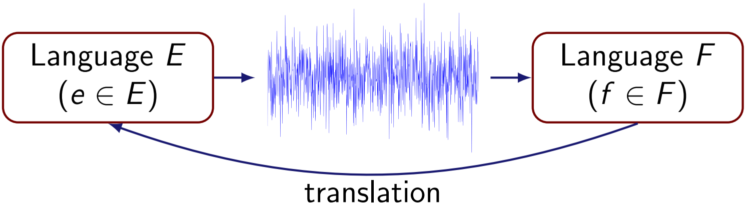
Statistical Machine Translation I

The Noisy Channel as a statistical approach to translation:



Data-driven Machine Translation

Statistical Machine Translation I

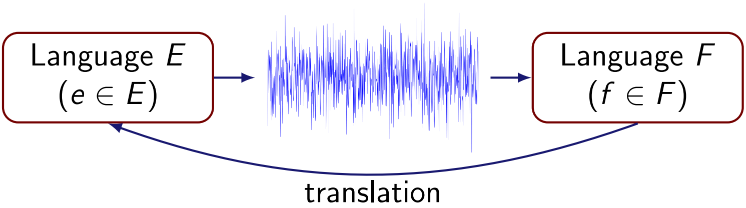


Mathematically:

$$P(e|f)$$

Data-driven Machine Translation

Statistical Machine Translation I



Mathematically:

$$P(e|f) = \frac{P(e)\,P(f|e)}{P(f)}$$

$$T(f) = \hat{e} = \mathrm{argmax}_{\mathrm{e}}\; P(e|f) = \mathrm{argmax}_{\mathrm{e}}\; P(e)\,P(f|e)$$

$$T(f) = \hat{e} = \mathrm{argmaxe}\ \mathbf{P(e)}\ P(f|e)$$

Language Model

- Takes care of fluency in the target language
- Data: corpora in the target language

Data-driven Machine Translation

Statistical Machine Translation II

$$T(f) = \hat{e} = \mathrm{argmaxe}\ P(e)\,\mathbf{P(f|e)}$$

Language Model

- Takes care of fluency in the target language
- Data: corpora in the target language

Translation Model

- Lexical correspondence between languages
- Data: aligned corpora in source and target languages

$$T(f) = \hat{e} = \textbf{argmax}_e \; P(e)\, P(f|e)$$

Language Model

- Takes care of fluency in the target language
- Data: corpora in the target language

Translation Model

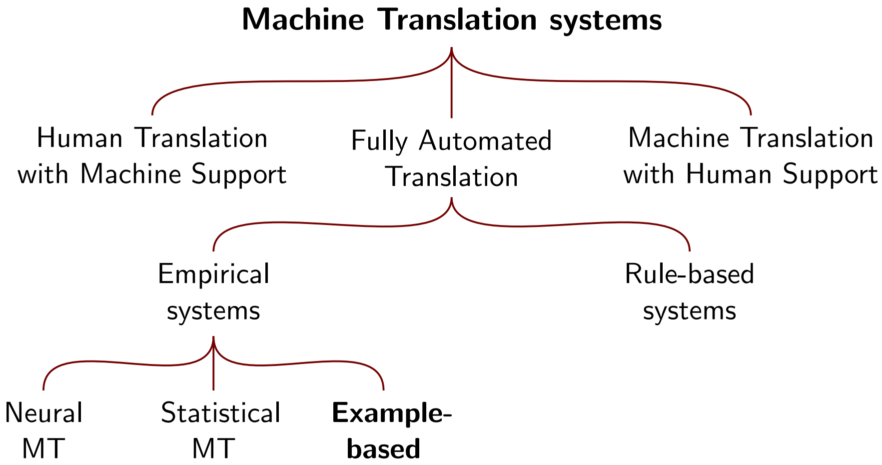
- Lexical correspondence between languages
- Data: aligned corpora in source and target languages

argmax

- Search done by the *decoder*

Data-driven Machine Translation

Example-based Machine Translation



Data-driven Machine Translation

Example-based Machine Translation

1. **Compile** and align a database of examples
2. **Match** input to a database of translation examples with **similarity** measures
3. **Identify** corresponding translation fragments
4. **Recombine** fragments into target text

It may also makes use of rules to find matches and to recombine aligned parts and build the final translation

Data-driven Machine Translation

SMT vs. EBMT

SMT

- Probabilities to access the merit of candidates
- Probabilities to rank candidates (decoding)
- Join together translated fragments

EBMT

- Similarity score between input fragments to fragments in database
- Syntactic and/or semantic similarity to rank candidates
- Join together translated fragments

Machine Translation systems

- Human Translation with Machine Support
- Fully Automated Translation
  - Empirical systems
    - **Neural MT**
    - **Statistical MT**
    - **Example-based**
  - Rule-based systems
    - **Direct MT**
    - **Transfer MT**
    - **Interlingua**
- Machine Translation with Human Support

Comparison

RBMT vs. SMT vs. NMT for High-Quality Systems

	RBMT	**SMT**	**NMT**
Data Amount	small	large	large
Training Time	–	days	weeks
CPU/GPU	CPU	CPU	GPU
Cost	expensive (in people)	cheap	expensive (in hardware)
Maintainability	weak	strong	superstrong
Grammaticality	strong	medium	strong
Reordering	strong	weak	strong
Consistency	strong	medium	weak
Coverage	weak	strong	weak
Multilinguality	medium	none	strong

Comparison

RBMT vs. SMT vs. NMT for High-Quality Systems II

I want a good translator, what MT should I use?

1. Think of your problem: language, domain and application
2. Think of your resources: time, hardware and money
3. Decide

RBMT vs. SMT vs. NMT for High-Quality Systems II

I want a good translator, what MT should I use?

1. Think of your problem: language, domain and application
2. Think of your resources: time, hardware and money
3. Decide

- The previuos slide shows general trends
- But quality depends on the language pair and domain
- All systems have pros and cons, why not **hybridisation**?

Hybrid Machine Translation

Current Systems

There are very few pure single MT systems, **hybridisation** is a must to take advantage of the strengths of the different methods

Still... almost no hybridisation yet in the Neural Era besides system combination!

Hybrid Machine Translation

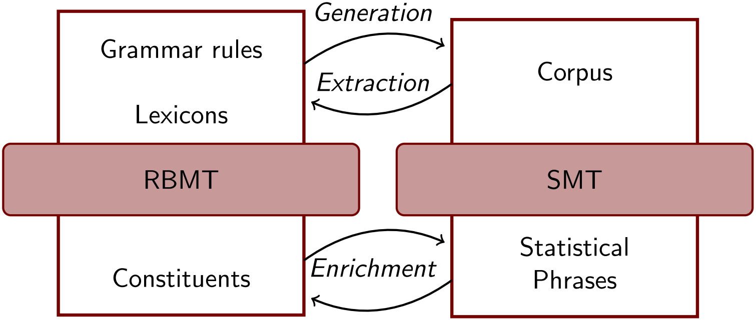
"The best of all worlds"

There are very sophisticated approaches, but most of them can be classified under these categories:

- System Combination
- Hybridisation lead by an SMT system
- Hybridisation lead by an RBMT system

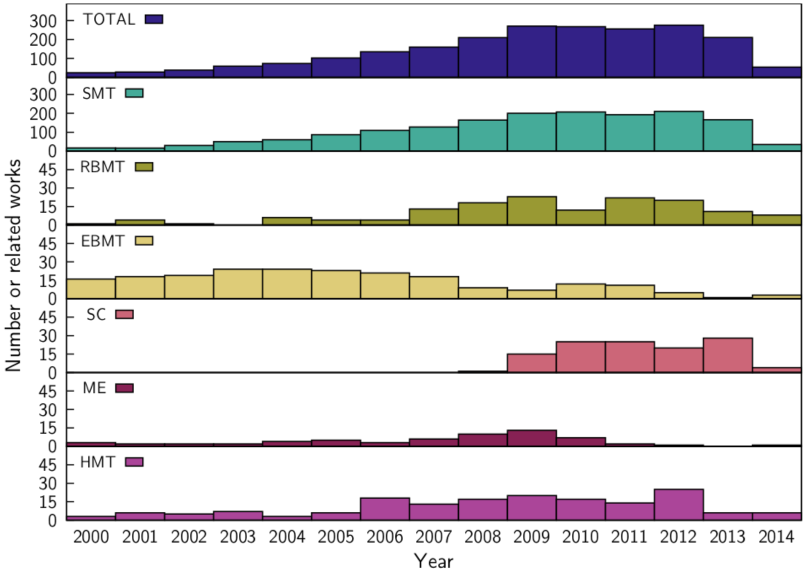
Hybrid Machine Translation

Simple Hybridisation examples between RBMT and SMT



Hybrid Machine Translation

Works per MT paradigm



Machine Translation Systems

Not Fully Automated Translation

Machine Translation systems

- **Human Translation with Machine Support**
- Fully Automated Translation
  - Empirical systems
  - Rule-based systems
- Machine Translation with Human Support

Not Fully Automated Translation

Human Translation with Machine Support

Machine-aided Human Translation, MAHT

Uses computer assisted translation tools (**CAT tools**) to:

- access a bilingual terminology
- access a translation memory
- submit parts ot the text to an MT server

Not Fully Automated Translation

Human Translation with Machine Support II

Translation Memories are aids for human translators

- Store and index entire existing translations
- Before translation, check the index to see if it's already been translated and reuse

- Strict matches: very reliable translation
- Fuzzy matches: more flexible, greater cover, but less reliable (similar to EBMT!)

Not Fully Automated Translation

CATs: SDL Trados

SDL Trados, a commercial tool

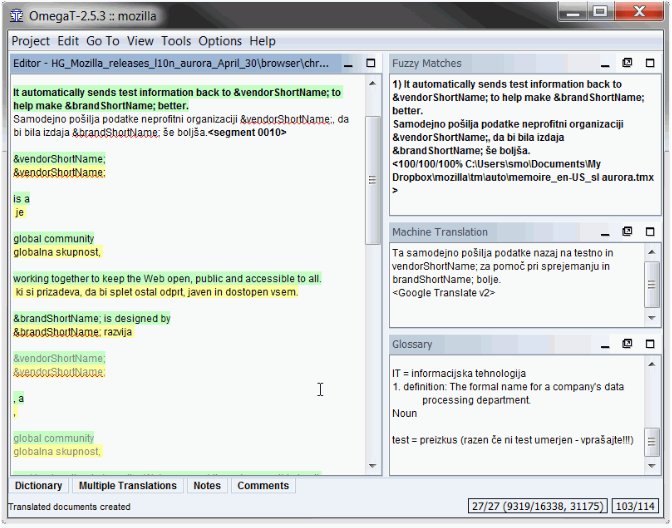
https://www.youtube.com/watch?v=FgBAyxFq3Ok

1:30-1:50; 3:25-5:06

Not Fully Automated Translation

CATs: OmegaT

OmegaT, a free tool



Not Fully Automated Translation

Not Fully Automated Translation

Machine Translation systems

Human Translation with Machine Support

Fully Automated Translation

Empirical systems

Rule-based systems

Machine Translation with Human Support

Not Fully Automated Translation

Machine Translation with Human Support

Human-aided Machine Translation, HAMT

Implies the automation of the translating function, with some human intervention in pre-editing, post-editing, or interaction

Not Fully Automated Translation

Machine Translation with Human Support

Human-aided Machine Translation, HAMT

Implies the automation of the translating function, with some human intervention in pre-editing, **post-editing**, or interaction

Come on Monday to Mihaela Vela's class!
Please, 14:00 sharp and bring your laptop to post-edit

References

Translation

- A Textbook of Translation. Peter Newmark. Prentice Hall.
- Mahmoud Ordudari. *Translation procedures, strategies and methods.* Translation Journal, volume 11, No. 3, July 2007.

References

History

- Translation. Warren Weaver. In William N. Locke and A. Donald Boothe, editors, Machine Translation of Languages, pages 15–23. MIT Press, 1949/1955.
- W. J. Hutchins. *Machine translation and machine-aided translation.* Journal of Documentation, 34(2):119–159, 1978.
- Jonathan Slocum. *A survey of machine translation: its history, current status, and future prospects.* Computational Linguistics, 11(1):1–17, 1985.
- Automatic Language Processing Advisory Committee (ALPAC). Language and Machines. Computers in Translation and Linguistics. Technical Report Publication 1416, National Research Council, Washington, 1966.

References

Rule-based MT

- Vauquois, B. *A survey of formal grammars and algorithms for recognition and transformation in mechanical translation.* IFIP Congress (2), 1114–1122, 1968.
- A. M. Corbí-Bellot, M. Forcada, S. Ortiz-Rojas, J. A. Pérez-Ortiz, G. Sánchez-Ramírez, F. Sánchez-Martínez, I. Alegria, A. Mayor, K. Sarasola. *An open-source shallow-transfer machine translation engine for the romance languages of Spain*. 10th Annual Conference of the European Assocation for Machine Translation, Budapest, Hungary, pages 79–86. (2005)
- Grammatical Framework: Programming with Multilingual Grammars. Aarne Ranta. CSLI Publications, Stanford (2011).

References

Others

- H. Somers. *An overview of EBMT.* In Michael Carl and Andy Way, editors, Recent advances in example-based machine translation, Kluwer Academic Publishers, pages 3–57. (2003)
- C. España-Bonet, Marta Ruiz Costa-jussà. *Hybrid Machine Translation Overview.* Chapter in Hybrid Approaches to Machine Translation, part of the series Theory and Applications of Natural Language Processing, pages 1–24. (2016)

References

SMT & NMT

See references at the corresponding sessions