

# PA153

## Strojový překlad

Vít Baisa

## Překlad

Překlad je převod textu ze zdrojového jazyka do jazyka cílového.

- ▶ odborný překlad × literární překlad
- ▶ přesná reprodukce × volná převodová parafráze

## Maimonidés, 12. stol.

Pro překlad slova je rozhodující kontext.

## Werner Winter

Každé slovo je element vytržený z celkového jazykového systému a jeho vztahy k jiným segmentům systému jsou v jednotlivých jazycích rozdílné.

Každý význam je element z celého systému segmentů, v něž mluvčí rozčleňuje skutečnost. V jazyce Mohave: otec ženy ≠ otec muže

## Otázky překladu

- ▶ Je vůbec přesný překlad mezi jazyky možný?
- ▶ Jak se pozná, že  $w_1$  je překladový ekvivalent slova  $w_2$ ?
- ▶ anglické typy větru, eskymácké typy sněhu, ...
- ▶ jak přeložit slova jako *alkáč*, *večerníček*, *telka*, *čoklbuřt*,  
*knížečka*, *ČSSD* ... ?

# Strojový překlad I – definice

## Strojový překlad

Obor počítačové lingvistiky zabývající se návrhem, implementací a aplikací automatických systémů (programů) pro překlad textů s minimálním zásahem člověka.

Např. používání elektronických slovníků při překladu nepatří do strojového překladu.

## Strojový překlad II – předmět zájmu

Zejména:

- ▶ webové stránky
- ▶ technické manuály
- ▶ vědecké dokumenty
- ▶ prospekty, katalogy
- ▶ právnické texty
- ▶ obecně texty z omezených domén

Nuance na různých jazykových vrstvách v umělecké literatuře jsou mimo schopnosti současných nástrojů NLP.

## Strojový překlad III

Ve skutečnosti je výstup z MT vždy revidován. Mluví se o *před-překladu* resp. o *post-editaci*.

Ta je někdy nutná i u člověka, ovšem systémy MT dělají zcela rozdílné chyby.

Pro člověka jsou typické chyby:

- ▶ špatné předložky (*I am in school*)
- ▶ chybějící členy (*I saw man*)
- ▶ špatné tvary slov: *mouses, breaked, ...*

Pro počítač jsou typické zejména chyby významové: *Kiss me, honey.*

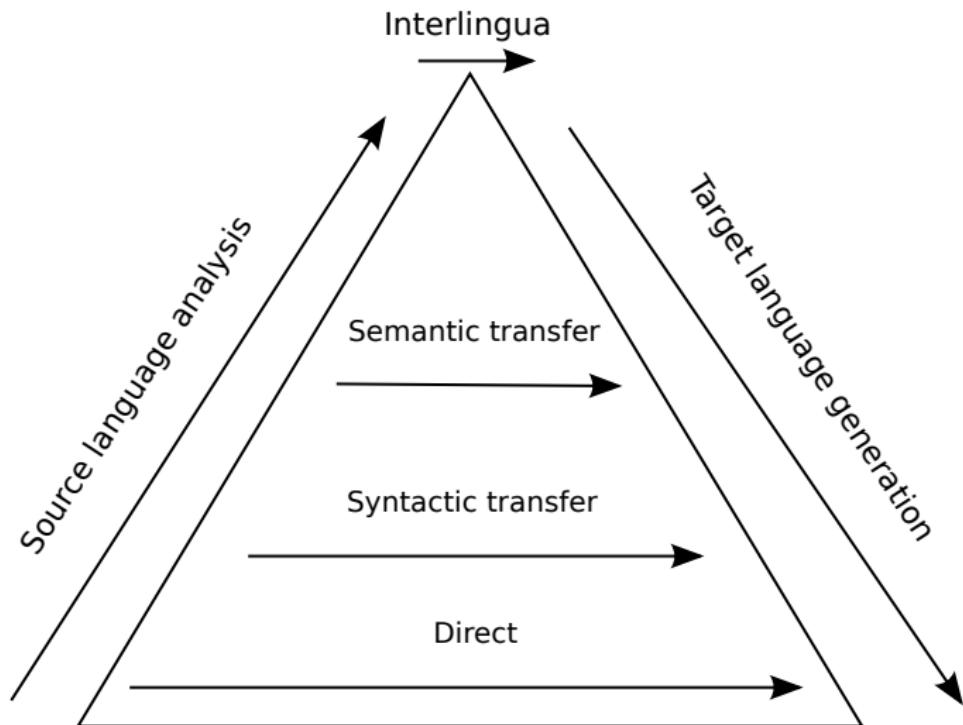
# Metody přímého zlepšení kvality strojového překladu

- ▶ omezení vstupu na:
  - ▶ podjazyk (oznamovací věty)
  - ▶ doménu (informatika)
  - ▶ typ dokumentu (patentové dokumenty)
- ▶ pre-processing textu (např. ruční syntaktická analýza)

## Klasifikace podle přístupu (approach)

- ▶ pravidlový (znalostní) strojový překlad  
*rule-based, knowledge-based – RBMT, KBMT*
  - ▶ transferový
  - ▶ interlingua
- ▶ statistický strojový překlad  
*statistical machine translation – SMT*  
*Example-based machine translation – EBMT*
- ▶ hybridní strojový překlad  
*hybrid machine translation – HMT, HyTran*

# Vauquoisův trojúhelník



## Klasifikace podle interakce s uživatelem

- ▶ (ruční překlad)
- ▶ ruční překlad s pomocí počítače  
*machine-aided human translation – MAHT*
- ▶ automatický překlad s interagujícím překladatelem  
*human-aided machine translation – HAMT*
- ▶ plně automatický překlad  
*fully automated high-quality (M)T – FAHQT*

HAMT a MAHT někdy souhrnně označovány jako CAT – computer-aided translation.

# Rané názory na strojový překlad

- ▶ překlad je často opakovaná činnost – věřilo se, že bude tuto proceduru možné počítačem napodobit
- ▶ úspěchy použití počítačů v kryptografii: vhodné i pro strojový překlad?

## Warren Weaver

When I look at an article in Chinese, I say: This is really written in English, but it has been coded in some strange symbols.  
I will now proceed to decode.

## Georgetown experiment

První funkční prototyp strojového překladu.

- ▶ 50 vět (zřejmě pečlivě vybraných)
- ▶ spolupráce s IBM
- ▶ slovník obsahoval 250 slov
- ▶ překlad z ruštiny do angličtiny
- ▶ gramatika pro ruštinu obsahovala 6 pravidel

Demonstrace systému vyvolala nadšení. MT bylo očividně možné. Následně odstartovalo mnoho nových projektů, hlavně v USA a Rusku.

# Vývoj v 50. letech

- ▶ MT oblast podnítila rozvoj a výzkum na poli
  - ▶ teoretické lingvistiky (Chomsky)
  - ▶ počítačové lingvistiky
  - ▶ umělé inteligence (60. léta)
- ▶ s větším pokrytím kvalita strojového překladu klesala
- ▶ i nejlepší systémy (GAT, Georgetown, RU→EN) poskytovaly nepoužitelný výstup

## Zklamání ze slabých výsledků

- ▶ i přes nevalné výsledky přetrával optimismus
- ▶ Yehoshua Bar-Hillel píše v roce 1959 kritiku stavu strojového překladu
- ▶ tvrdí, že počítače nejsou schopné provádět lexikální desambiguaci
- ▶ fully automated high-quality translation (FAHQT) podle Bar-Hillela stěží dosažitelné

### Yehoshua Bar-Hillel – příklad pro desambiguaci

Little John was looking for his toy box. Finally, he found it. The box was in the pen. John was very happy.

Výdaje na projekty strojového překladu se začaly snižovat.

## ALPAC report

- ▶ Automatic Language Processing Advisory Committee
- ▶ organizace pod U.S. National Academy of Science
- ▶ analýzy a vyhodnocení kvality a použitelnosti systémů MT
- ▶ doporučila omezit výdaje na podporu strojového překladu
- ▶ vývoj strojového překladu v Evropě a Japonsku pokračoval nepřerušeně dál
- ▶ celých 15 let trvalo než MT v USA znova získal vážnost a původní postavení

# TAUM, METEO

## TAUM

- ▶ Traduction Automatique à l'Université de Montréal
- ▶ Université de Montréal in 1965
- ▶ prototypy MT systémů: TAUM-73, TAUM-METEO
- ▶ jedny z prvních systémů provádějící automatický překlad přes analýzu zdrojového jazyka a syntézu cílového jazyka
- ▶ překlad z angličtiny do francouzštiny

## METEO

- ▶ 1981–2001 používán pro překlad předpovědí počasí
- ▶ autor John Chandiou, Kanada

# Systran

- ▶ jedna z nejstarších MT firem (1968)
- ▶ velmi populární překladový systém
- ▶ základ Yahoo Babel fish, do r. 2007 využíván v Google
- ▶ RBMT, od r. 2010 hybridní překlad
- ▶ od r. 1976 oficiální MT systém používaný Evropským hospodářským společenstvím

## Strojový překlad v současnosti

- ▶ intenzivní sběr paralelních dat
- ▶ vývoj systémů vzhledem k hodnoticím metrikám
- ▶ USA: zájem o angličtinu jako TL
- ▶ EU: překlad mezi 23 úředními jazyky EU (EuroMatrix)
- ▶ korporace (Microsoft) zaměřeny na En jako SL
- ▶ SMT obohacována syntaxí
- ▶ velké páry (En↔Sp, En↔Fr): velmi dobrý překlad
- ▶ Google Translate jako gold standard
- ▶ morfologicky bohaté jazyky jsou opomíjeny
- ▶ En-\* a \*-En páry převažují

# Motivace pro strojový překlad ve 21. století

- ▶ překlad webových stránek pro pochopení obsahu (gisting)
- ▶ metody pro výrazné urychlení překladatelské práce (překladové paměti)
- ▶ extrakce a vyhledávání informací mezi jazyky (cross-lingual IR)
- ▶ instantní překlad instantní komunikace
- ▶ překlad na mobilních zařízeních

## EuroMatrix, 2006–2009

- ▶ překlad všech párů EU jazyků
- ▶ využití lingvistické znalosti ve statistickém MT
- ▶ vývoj a testování hybridních architektur
- ▶ každoroční vyhodnocování kvality MT
- ▶ open source: nástroje, software, data
- ▶ přehled stavu MT pro EU jazyky

## Závěr úvodu

- ▶ strojový překlad patří mezi UI-kompletní problémy
- ▶ máme k dispozici obrovskou výpočetní sílu
- ▶ tržní potenciál je větší než kdy dřív
- ▶ statistické metody se aktuálně zdají vhodnější
- ▶ nové nápady jsou vítány! (BP, DP)

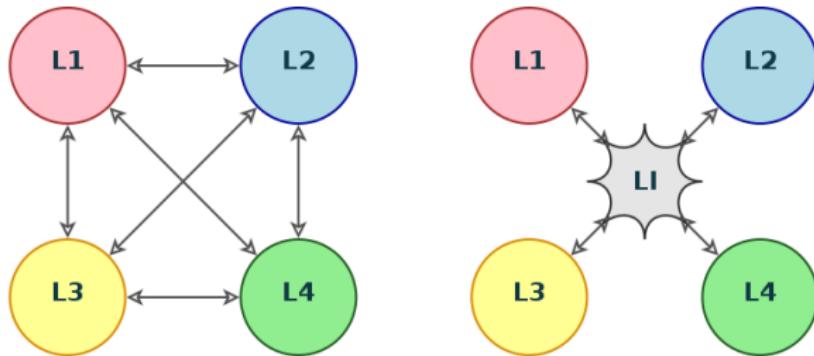
# Rule-based Machine Translation – RBMT

- ▶ lingvistické znalosti formou pravidel
- ▶ pravidla pro analýzu
- ▶ pravidla pro převod struktur mezi jazyky
- ▶ pravidla pro syntézu
- ▶ Knowledge-based Machine Translation

Rozdělení systémů RBMT:

- ▶ přímý překlad
- ▶ systémy používající interlinguu
- ▶ transferové systémy

# Interlingua vs. transferové KBMT



# Proces analýzy

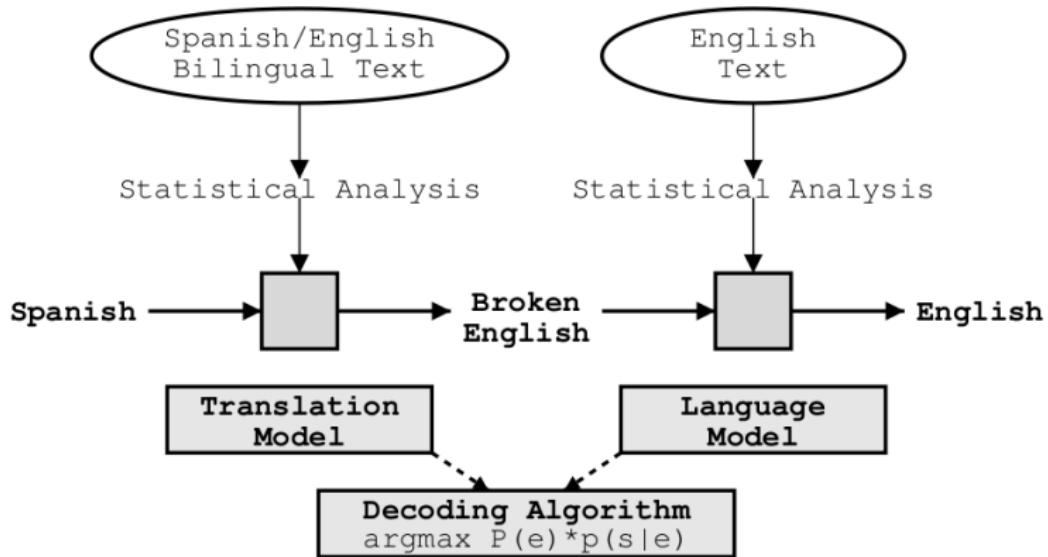
- ▶ tokenizace
- ▶ morfologická analýza (značkování)
- ▶ morfologická disambiguace (pravidlová, statistická)
- ▶ syntaktická analýza
- ▶ sémantická analýza: lexikální disambiguace

| slovo      | analýzy   | disambiguace |
|------------|---|--------------|
| Pravidelné | k2eAgMnPc4d1, k2eAgInPc1d1, k2eAgInPc4d1,<br>k2eAgInPc5d1, k2eAgFnSc2d1, k2eAgFnSc3d1,<br>k2eAgFnSc6d1, k2eAgFnPc1d1, k2eAgFnPc4d1,<br>k2eAgFnPc5d1, k2eAgNnSc1d1, k2eAgNnSc4d1,<br>k2eAgNnSc5d1, ... (+ 5) | k2eAgNnSc1d1 |
| krmení     | k2eAgMnPc1d1, k2eAgMnPc5d1, k1gNnSc1,<br>k1gNnSc4, k1gNnSc5, k1gNnSc6, k1gNnSc3,<br>k1gNnSc2, k1gNnPc2, k1gNnPc1, k1gNnPc4,<br>k1gNnPc5   | k1gNnSc1     |
| je         | k5eAalmP3nS, k3p3gMnPc4, k3p3gInPc4,<br>k3p3gNnSc4, k3p3gNnPc4, k3p3gFnPc4, k0  | k5eAalmP3nS  |
| pro        | k7c4  | k7c4         |
| správný    | k2eAgMnSc1d1, k2eAgMnSc5d1, k2eAgInSc1d1,<br>k2eAgInSc4d1, k2eAgInSc5d1, ... (+ 18)   | k2eAgInSc4d1 |
| růst       | k5eAalmF, k1glnSc1, k1glnSc4  | k1glnSc4     |
| důležité   | k2eAgMnPc4d1, k2eAgInPc1d1, k2eAgInPc4d1,<br>k2eAgInPc5d1, k2eAgFnSc2d1, k2eAgFnSc3d1,<br>k2eAgFnSc6d1, k2eAgFnPc1d1, k2eAgFnPc4d1,<br>k2eAgFnPc5d1, k2eAgNnSc1d1, k2eAgNnSc4d1,<br>k2eAgNnSc5d1, ... (+ 5) | k2eAgNnSc1d1 |

# Statistický MT

- ▶ pravidlové systémy motivovány lingvistikou
- ▶ SMT inspirován teorií informace a statistikou
- ▶ 50 miliónů stránek denně přeložených pomocí SMT
- ▶ **gisting**: stačí, má-li překlad nějaký užitek, nepotřebujeme přesný význam; nejčastější užití MT na internetu

# Schéma SMT



## Data pro SMT – (paralelní) korpusy

- ▶ Europarl: kolekce textů Evropského parlamentu
- ▶ OPUS: paralelní texty různého původu
- ▶ Acquis Communautaire: právní dokumenty Evropské únie (20 jazyků)
- ▶ volně dostupné jsou řádově 10 a 100 miliónů slov veliké
- ▶ vícejazyčné stránky (Wikipedie)
- ▶ srovnatelné korpusy (comparable corpora): texty ze stejné domény, ne přímé překlady: New York Times – Le Monde
- ▶ Kapradí – korpus překladů Shakespearových dramat (FI)
- ▶ InterCorp – ručně zarovnané beletr. texty (ČNK, FFUK)

## SMT – princip noisy channel

Vyvinut Shannonem (1948) pro potřeby samoopravujících se kódů, pro korekce kódovaných signálů přenášených po zašuměných kanálech na základě informace o původní zprávě a typu chyb vznikajících v kanálu.

Příklad s OCR. Rozpoznávání textu z obrázků je chybové, ale dokážeme odhadnout, co by mohlo být v textu (jazykový model) a jaké chyby často vznikají: záměna l-1-l, rn-m apod.

$$\begin{aligned} e^* &= \arg \max_e p(e|f) \\ &= \arg \max_e \frac{p(e)p(f|e)}{p(f)} \\ &= \arg \max_e p(e)p(f|e). \end{aligned}$$

## Jazykové modely

- ▶ LM pomáhají zajistit **plynulý výstup** (správný slovosled)
- ▶ LM pomáhají s **WSD v obecných případech**
- ▶ pokud má slovo více významů, můžeme vybrat nejčastější překlad (*pen* → *pero*)
- ▶ ve speciálních textech nelze použít, ale
- ▶ LM pomáhají s **WSD pomocí kontextu**
- ▶  $p_{LM}(i \text{ go home}) \geq p_{LM}(i \text{ go house})$

# N-gramové modely

- ▶ n-gram je nejdůležitější nástroj ve zpracování řeči a jazyka
- ▶ využití statistického pozorování dat

## Generování unigramy

To him swallowed confess hear both. Which. Of save on trail for  
are ay device and rote life have Every enter now severally so,  
let.

## Generování trigramy

Sweet prince, Falstaff shall die. Harry of Monmouth's grave.  
This shall forbid it should be branded, if renown made it empty.

## Kvalita a srovnání jazykových modelů

Chceme být schopni porovnávat kvalitu různých jazykových modelů (trénovány na různých datech, pomocí jakých n-gramů, jak vyhlazených apod.).

Je možné použít 2 přístupy: vnější (extrinsic) a vnitřní (intrinsic) vyhodnocení.

Dobrý model by měl přiřadit dobrému textu vyšší pravděpodobnost než špatnému textu.

Pokud máme nějaký testovací text, můžeme spočítat pravděpodobnost, jakou mu přiřazuje zkoumaný LM. Lepší LM by mu měl přiřadit vyšší pravděpodobnost.

## Vyhlažování jazykových modelů

Problém: pokud není v datech určitý n-gram, který se vyskytne v řetězci  $w$ , pro který hledáme pravděpodobnost, bude  $p(w) = 0$ .

Potřebujeme rozlišovat  $p$  i pro *neviděná data*. Musí platit

$$\forall w. p(w) > 0$$

Ještě větší je problém u modelů vyšších řadů.

Snaha o úpravení reálných počtů n-gramů na očekávané počty těchto n-gramů v libovolných datech (jiných korpusech).

Add-one, add-alpha, deleted estimation, Good-Turing vyhlazování

# Interpolace

Předchozí metody zacházely se všemi neviděnými n-gramy stejně. Předpokládejme 3-gramy:

*nádherná červená řepa*

*nádherná červená mrkev*

I když ani jeden nemáme v trénovacích datech, první 3-gram by měl být pravděpodobnější.

Budeme využívat pravděpodobnosti n-gramů nižších řádů, u kterých máme k dispozici více dat:

*červená řepa*

*červená mrkev*

# Velké jazykové modely – počet n-gramů

Kolik je různých n-gramů v korpusu?

| řad     | unikátní   | singletony          |
|---------|------------|---------------------|
| unigram | 86 700     | 33 447 (38,6 %)     |
| bigram  | 1 948 935  | 1 132 844 (58,1 %)  |
| trigram | 8 092 798  | 6 022 286 (74,4 %)  |
| 4-gram  | 15 303 847 | 13 081 621 (85,5 %) |
| 5-gram  | 19 882 175 | 18 324 577 (92,2 %) |

Europarl, 30 miliónů tokenů.

## Výpočet překladové pravděpodobnosti

Pro výpočet  $p(\mathbf{e}, \mathbf{a}|\mathbf{f})$  potřebujeme znát hodnotu funkce  $t$  pro všechna slova (věty).

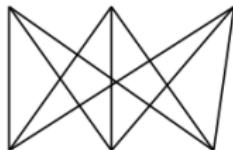
K tomu budeme využívat základní zdroj pro SMT: **paralelní korpus** se zarovnanými větami.

Bohužel nemáme zarovnání slov mezi sebou. To je úkol tzv. **word-alignment**.

Ke slovu přichází **expectation-maximization (EM)** algoritmus.

## Ilustre EM algoritmu

... la maison ... la maison blue ... la fleur ...



... the house ... the blue house ... the flower ...

... la maison ... la maison bleu ... la fleur ...



... the house ... the blue house ... the flower ...



$$p(\text{la}|\text{the}) = 0.453$$

$$p(\text{le}|\text{the}) = 0.334$$

$$p(\text{maison}|\text{house}) = 0.876$$

$$p(\text{bleu}|\text{blue}) = 0.563$$

...

## Matice zarovnání slov

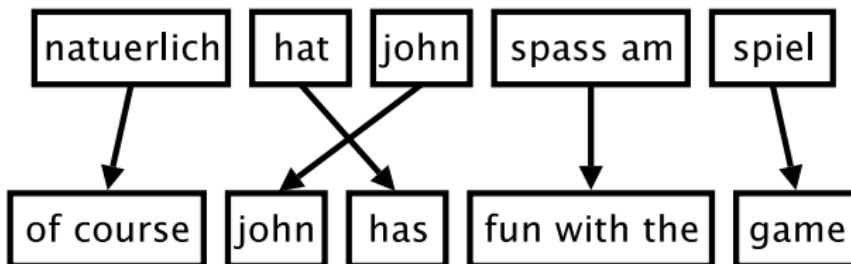
# Problémy se zarovnáním slov

|        |        |        |        |        |
|--------|--------|--------|--------|--------|
|        | john   | biss   | ins    | grass  |
| john   | ██████ | ██     | ██     | ██     |
| kicked | ██     | ██████ | ██████ | ██████ |
| the    | ██     | ██     | ██████ | ██████ |
| bucket | ██     | ██     | ██     | ██████ |

|      |        |        |        |        |
|------|--------|--------|--------|--------|
|      | john   | wohnt  | hier   | nicht  |
| john | ██████ | ██     | ██     | ██     |
| does | ██     | ████?█ | ██     | ████?  |
| not  | ██     | ██     | ██     | ██████ |
| live | ██     | ██     | ██     | ██     |
| here | ██     | ██     | ██████ | ██     |

# Frázový překladový model

State-of-the-art statistického strojového překladu. Nepřekládají se pouze samostatná slova. Když to jde, tak i celé sekvence slov.



Fráze nejsou lingvisticky motivované, pouze statisticky.  
Německé *am* se zřídka překládá jedním slovem *with*. Statisticky významný kontext *spass am* pomáhá správnému překladu.  
Klasické fráze by se dělily jinak: (*fun (with (the game))*).

## Výhody PBTM

- ▶ často překládáme  $n : m$  slov, slovo je tedy nevhodný atomický prvek
- ▶ překlad skupin slov pomáhá řešit překladové víceznačnosti
- ▶ můžeme se učit překládat delší a delší fráze
- ▶ jednodušší model: neuvažujeme fertilitu, NULL token atd.

## Extrahování frází

michael geht davon aus , dass er im haus bleibt

michael  
assumes  
that  
he  
will  
stay  
in  
the  
house

## Automatické hodnocení překladu

- ▶ výhody: rychlosť, cena; nevýhody: měříme opravdu kvalitu?
- ▶ gold standard: ručně připravené referenční překlady
- ▶ kandidát  $c$  se srovnává s  $n$  referenčními překlady  $r_i$
- ▶ různé přístupy: n-gramová shoda mezi  $c$  a  $r_i$ , editační vzdálenost, ...

# Pokrytí a přesnost na slovech

Nejjednodušší způsob automatického hodnocení



- ▶ přesnost

$$\frac{\text{correct}}{\text{output-length}} = \frac{3}{6} = 50\%$$

- ▶ pokrytí

$$\frac{\text{correct}}{\text{reference-length}} = \frac{3}{7} = 43\%$$

- ▶ f-score

$$\frac{\text{precision} \times \text{recall}}{(\text{precision} + \text{recall})/2} = \frac{.5 \times .43}{(.5 + .43)/2} = 46\%$$

# Pokrytí a přesnost – nedostatky



| metrika  | systém A | systém B |
|----------|----------|----------|
| přesnost | 50%      | 100%     |
| pokrytí  | 43%      | 100%     |
| f-score  | 46%      | 100%     |

Nepostihuje se nesprávný slovosled.

## BLEU

- ▶ nejznámější (standard), nejpoužívanější, nejstarší (2001)
- ▶ IBM, Papineni
- ▶ n-gramová shoda mezi referencí a kandidáty
- ▶ počítá se přesnost pro 1 až 4-gramy
- ▶ extra postih za krátkost (**brevity penalty**)

$$\text{BLEU} = \min \left( 1, \frac{\text{output-length}}{\text{reference-length}} \right) \left( \prod_{i=1}^4 precision_i \right)^{\frac{1}{4}}$$

## BLEU – příklad

SYSTEM A: **Israeli officials** responsibility of **airport** safety  
2-GRAM MATCH 1-GRAM MATCH

REFERENCE: Israeli officials are responsible for airport security

SYSTEM B: **airport security** **Israeli officials are responsible**  
2-GRAM MATCH 4-GRAM MATCH

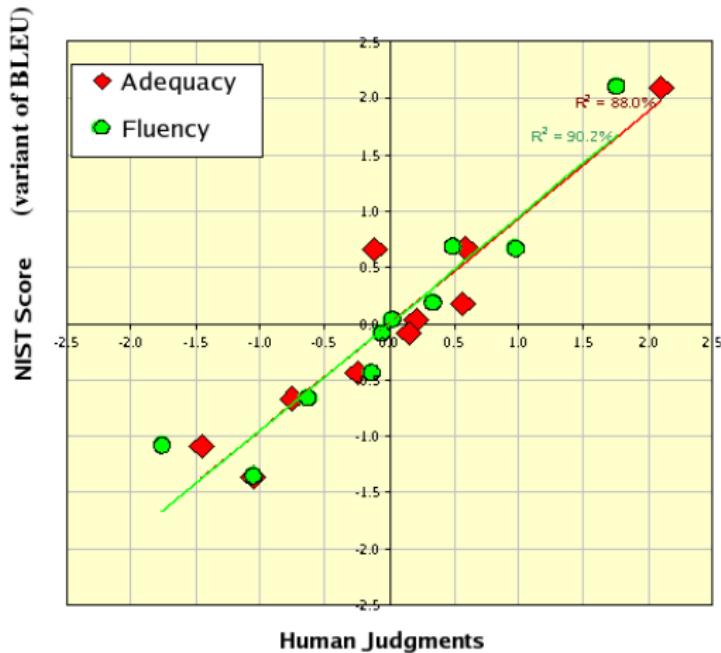
| metrika          | systém A | systém B |
|------------------|----------|----------|
| přesnost (1gram) | 3/6      | 6/6      |
| přesnost (2gram) | 1/5      | 4/5      |
| přesnost (3gram) | 0/4      | 2/4      |
| přesnost (4gram) | 0/3      | 1/3      |
| brevity penalty  | 6/7      | 6/7      |
| BLEU             | 0 %      | 52 %     |

# Další metriky

- ▶ NIST, NEVA
- ▶ WAFT
  - ▶ Word Accuracy for Translation
  - ▶ editační vzdálenost mezi  $c$  a  $r$
  - ▶  $WAFT = 1 - \frac{d+s+i}{\max(l_r, l_c)}$
- ▶ TER
  - ▶ Translation Edit Rate
  - ▶ nejmenší počet kroků (smazání, přidání, prohození, změna)
  - ▶  $TER = \frac{\text{počet editací}}{\text{prům. počet ref. slov}}$
- ▶ METEOR
  - ▶ uvažuje synonyma (WordNet) a
  - ▶ morfologické varianty slov

# Hodnocení hodnoticích metrik

Korelace automatického hodnocení s manuálním.



# Hodnocení překladu podle jazykových párů

|    | Target language |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
|----|-----------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|    | EN              | BG   | DE   | CS   | DA   | EL   | ES   | ET   | R    | FR   | HU   | IT   | LT   | LV   | MT   | NL   | PL   | PT   | RO   | SK   | SL   | SV   |
| EN | 40.3            | 46.8 | 52.6 | 50.0 | 41.0 | 55.2 | 34.8 | 38.6 | 50.1 | 37.2 | 50.4 | 39.6 | 43.4 | 39.8 | 52.3 | 49.2 | 53.0 | 49.0 | 44.7 | 50.7 | 52.0 |      |
| BG | 61.8            | 38.7 | 39.4 | 39.6 | 34.5 | 46.9 | 25.5 | 26.7 | 42.4 | 22.0 | 48.5 | 29.3 | 29.1 | 25.9 | 44.9 | 35.1 | 45.9 | 36.8 | 34.1 | 34.1 | 39.9 |      |
| DE | 53.6            | 26.3 | 35.4 | 43.1 | 32.8 | 47.1 | 25.7 | 29.5 | 39.4 | 27.6 | 42.7 | 27.6 | 30.3 | 19.8 | 50.1 | 30.2 | 44.1 | 30.7 | 29.4 | 31.4 | 41.2 |      |
| CS | 58.4            | 32.0 | 42.6 | 35.7 | 43.6 | 34.6 | 48.9 | 30.7 | 30.5 | 41.6 | 27.4 | 44.3 | 34.5 | 33.8 | 26.1 | 46.5 | 39.2 | 43.7 | 36.5 | 43.6 | 41.3 | 42.9 |
| DA | 57.5            | 28.7 | 44.1 | 35.7 | 34.3 | 47.5 | 27.8 | 31.6 | 41.3 | 24.2 | 48.8 | 29.7 | 32.9 | 21.1 | 48.5 | 34.3 | 43.4 | 33.9 | 33.0 | 36.2 | 47.2 |      |
| EL | 59.5            | 32.4 | 43.1 | 37.7 | 44.5 | 35.7 | 34.0 | 26.5 | 29.0 | 48.3 | 23.7 | 49.6 | 29.0 | 32.6 | 23.8 | 48.9 | 34.2 | 52.5 | 37.2 | 33.1 | 36.3 | 43.3 |
| ES | 60.0            | 31.1 | 42.7 | 37.5 | 44.4 | 39.4 | 35.7 | 25.4 | 28.5 | 51.3 | 24.0 | 51.7 | 26.8 | 30.5 | 24.6 | 48.8 | 33.9 | 57.3 | 38.1 | 31.7 | 33.9 | 43.7 |
| ET | 52.0            | 24.6 | 37.3 | 35.2 | 37.8 | 28.2 | 40.4 | 35.7 | 33.4 | 30.9 | 37.0 | 35.0 | 36.9 | 20.5 | 41.3 | 32.0 | 37.8 | 28.0 | 30.6 | 32.9 | 37.3 |      |
| FI | 49.3            | 23.2 | 36.0 | 32.0 | 37.9 | 27.2 | 39.7 | 34.9 | 35.7 | 29.5 | 27.2 | 36.6 | 30.5 | 32.5 | 19.4 | 40.6 | 28.8 | 37.5 | 26.5 | 27.3 | 28.2 | 37.6 |
| FR | 64.0            | 34.5 | 45.1 | 39.5 | 47.4 | 42.8 | 60.9 | 26.7 | 30.0 | 35.7 | 25.5 | 56.1 | 28.3 | 31.9 | 25.3 | 51.6 | 33.7 | 61.0 | 48.8 | 33.1 | 35.6 | 45.8 |
| HU | 48.0            | 24.7 | 34.3 | 30.0 | 38.0 | 25.5 | 34.1 | 29.6 | 29.4 | 30.7 | 35.7 | 29.6 | 31.9 | 18.1 | 36.1 | 29.8 | 34.2 | 25.7 | 25.6 | 28.2 | 30.3 |      |
| IT | 61.0            | 32.1 | 44.3 | 38.9 | 45.8 | 40.6 | 26.9 | 25.0 | 29.7 | 32.7 | 24.2 | 35.7 | 29.4 | 32.6 | 24.6 | 50.5 | 35.2 | 56.5 | 39.3 | 32.5 | 34.7 | 44.3 |
| LT | 51.8            | 27.6 | 33.9 | 37.0 | 36.8 | 26.5 | 21.1 | 34.2 | 32.0 | 34.4 | 28.5 | 36.8 | 35.7 | 40.1 | 22.2 | 38.1 | 31.6 | 31.6 | 29.3 | 31.8 | 35.3 | 35.3 |
| LV | 54.0            | 29.1 | 35.0 | 37.8 | 38.5 | 29.7 | 23.3 | 34.2 | 32.4 | 35.6 | 29.3 | 38.9 | 38.4 | 35.7 | 23.3 | 41.5 | 34.4 | 39.6 | 31.0 | 33.3 | 37.1 | 38.0 |
| MT | 72.1            | 32.2 | 37.1 | 37.9 | 38.9 | 33.7 | 48.7 | 26.9 | 23.8 | 42.4 | 22.4 | 48.7 | 30.2 | 33.2 | 35.7 | 44.0 | 37.1 | 45.9 | 38.9 | 35.8 | 40.0 | 41.6 |
| NL | 56.9            | 29.3 | 46.9 | 37.0 | 43.4 | 33.3 | 49.7 | 27.5 | 29.8 | 43.4 | 25.3 | 44.5 | 28.6 | 31.7 | 22.0 | 32.0 | 47.7 | 33.0 | 30.1 | 34.6 | 43.6 |      |
| PL | 60.8            | 31.5 | 40.2 | 44.2 | 42.1 | 34.2 | 46.2 | 29.2 | 29.0 | 40.0 | 24.5 | 43.2 | 33.2 | 35.6 | 27.9 | 44.8 | 35.7 | 44.1 | 38.2 | 38.2 | 39.8 | 42.1 |
| PT | 60.7            | 31.4 | 42.9 | 38.4 | 42.8 | 40.2 | 60.7 | 26.4 | 29.2 | 53.2 | 23.8 | 52.8 | 28.0 | 31.5 | 24.8 | 49.3 | 34.5 | 35.7 | 39.4 | 32.1 | 34.4 | 43.9 |
| RO | 60.8            | 33.1 | 38.5 | 37.8 | 40.3 | 35.6 | 30.4 | 24.6 | 26.2 | 46.5 | 25.0 | 44.8 | 28.4 | 29.9 | 28.7 | 43.0 | 35.8 | 48.5 | 31.5 | 35.1 | 39.4 |      |
| SK | 60.8            | 32.6 | 39.4 | 48.1 | 41.0 | 33.3 | 46.2 | 29.8 | 28.4 | 39.4 | 27.4 | 41.8 | 33.8 | 36.7 | 28.5 | 44.4 | 39.0 | 43.3 | 33.3 | 42.6 | 41.8 |      |
| SL | 61.0            | 33.1 | 37.9 | 43.5 | 42.6 | 34.0 | 47.0 | 31.1 | 28.8 | 38.2 | 25.7 | 42.3 | 34.6 | 37.3 | 30.0 | 43.9 | 38.2 | 44.1 | 33.8 | 38.9 | 42.7 |      |
| SV | 58.3            | 26.9 | 41.0 | 35.6 | 46.6 | 33.3 | 46.6 | 27.4 | 30.9 | 38.9 | 22.7 | 42.0 | 28.2 | 31.0 | 29.7 | 43.6 | 32.2 | 44.2 | 32.7 | 31.3 | 33.5 |      |