Supporting Communications in Global Networks

Kevin Duh & 歐陽靖民
Supporting Communications in Global Networks

Machine Translation

Kevin Duh
6000
Number of Languages in the World
世界中の言語の数
There are 6000 languages in the world.

世界には6000の言語があります。
When I look at an article in Russian, I say: “This is really written in English, but has been coded in some strange symbols. I will now proceed to decode.” [1947]

Warren Weaver,
American scientist (1894-1978)
Rule-Based Machine Translation (1960—now)

Manual encoding of linguistic knowledge into rules

Data-Driven Machine Translation (2000—now)

Exploiting Statistics from Large Data
Architecture of Data-Driven Machine Translation

There are 6000 languages in the world

Bilingual Text Data 対訳データ → Statistics 統計的學習 → Probabilistic Models 確率モデル

世界には6000の言語があります

Statistics 統計的學習
There are 6000 languages in the world.

- あります 6000 言語 には 世界
- 世界 には 6000 の言語 が あります

Architecture of Data-Driven Machine Translation

Word Translation Model

Fluency Model
How to build these models?
Exercise:
Let’s build one now!
1a) evas **dlrow-eht**

2a) **dlrow-eht** si detcennoc

3a) h**craeser** si **tnatropmi**

4a) ew **eb-ot-mia** tseb ni **dlrow-eht**

**Notes:**
- The text appears to be scrambled or encoded, with some symbols and characters rearranged.
- The formatting suggests a visual representation of characters or symbols, possibly for a puzzle or encoded message.
- The text and symbols might be part of a larger challenge or exercise in cryptography or decryption.
1a) evas dlrow-eht

1b) 🟩 🟠

2a) dlrow-eht si detcennoc

2b) 🟩 🟦 🟢

3a) hcraeser si tnatropmi

3b) 🟧 🟦 🟢

4a) ew eb-ot-mia tseb ni dlrow-eht

4b) 🟩 🔵 🔹 🔹
1a) save the-world

1b) □ △

2a) the-world is connected

2b) □  △  β

3a) research is important

3b)  △  △  β

4a) we aim-to-be best in the-world

4b) □  δ  ∆  φ  γ
1a) save the-world
1b) 世界を救え

2a) the-world is connected
2b) 世界は繋がってる

3a) research is important
3b) 研究は大事

4a) we aim-to-be best in the-world
4b) 世界一を狙う
in practice, it’s not that easy
Recent Research Efforts

Duh+(2010), “Analysis of Translation Model Adaptation”

What if data is insufficient?

Bilingual Text Data 対訳データ

Statistics 統計的学習

Probabilistic Models 確率モデル
Recent Research Efforts

**English:** Subject → Verb → Object

**Japanese:** 主語 → 目的語 → 動詞

How to use Japanese-specific knowledge?

Visit us @ Posters 5-7 for more info

Bilingual Text Data
対訳データ

Statistics
統計的学習

Probabilistic Models
確率モデル
Recent Research Efforts

Can we translate faster?

Bilingual Text Data 対訳データ

Statistics 統計的学習

Probabilistic Models 確率モデル

Using technology from Speech Recognition:
Hori+(2007), “Efficient WFST one-pass decoding w/ on-the-fly rescoring”
Killer Application for Translation?
Three General Scenarios

- **translate in**

- **translate out**

- **translate 2-way**
The Future Knowledge Worker

Translation

Technical Articles

Breaking News

Social Media

© Robert Brook
The entire world is your market.

Export Things:
- Cars
- Electronics
- Machinery
- Steel
- Chemicals

Export Culture:
- Movies
- Manga
- Apps
- Music
- ...

translate out
Real-time Conversation—now

6000 languages in the world
How will you use translation technology?

あなたなら、翻訳をどのようなコミュニケーションに使う？
Supporting Communication in Global Networks
Social Network Mining

歐陽靖民
Japanese Programmers on GitHub

Created by franckcuny@Flickr
Social Network Analysis
Methods and Applications
Stanley Wasserman and Katherine Faust
Six Degree of Separation

Stanley Milgram
American Social Psychologist
1933-1984

BBC recreating the experiment in 2009
(From Nairobi to Boston)
Obesity and Social Network
(Christakis and Fowler, 2007)

Contagious Happiness
A new study finds that a person’s happiness is influenced by the happiness of close friends and family, and to a lesser extent by the friends of friends, up to three degrees of separation.

Happiness in clusters
The diagram above shows the relationships among 1,020 people in 2000. Groups of happy and unhappy people are visible as light and dark clusters.

KEY
Most happy
Least happy
Men
Women

Happiness and Social Network
(Fowler and Christakis, 2008)
500,000,000 Active Users
30 billion content shared

200,000,000 Users
65 million tweets per day

The Age of Online Social Networking
Social Network Mining
Why social network mining?

Identifying **influential users** in a network
Why social network mining?

Understanding how information spreads

I like this!

I heard that it is good!
Why social network mining?

Making Social Recommendations

30 of your best friends have already bought this!
Understanding Social Networks
Measuring a Social Network

Example: Studying Twitter

Four Degree of Separation in Twitter

Kwak et al. What is Twitter, a Social Network or a News Media? WWW, 2010.
Modeling a Social Network

‘Exploring complex networks’ (Strogatz, 2001)

Small World Networks

‘Emergence of scaling in random networks’ (Barabási & Albert, 1999)

Scale-free Networks
Supporting Users in Social Networks
Social Information Overloading
Supporting Users in Social Networks

Learning the relations among people

Many interactions / Similar Opinions

Few interactions / Different opinions
Users participating in the social network

New algorithms for supporting the social network

Measuring and Modelling the social network
Our Research
Research Question

How trust relations shape user opinions?

How trust can be used to predict user opinions?
Find Reviews

- **Cameras & Photo**
  - Digital Cameras, Film Cameras, Camera Lenses...
- **Clothing & Apparel**
  - Women, Men, Shirts and Tops, Sweaters, Pants, Outerwear, Accessories, Shoes...
- **Electronics**
  - MP3 Players, GPS, Cell Phones, Camcorders, Audio, Home Theater...
- **Home & Garden**
  - Home Furnishings, Furniture, Tools, Cooking, Heaters...
- **Video Games**
  - Game Consoles, Computer Games, XBox 360, PS3, Wii

Epinions Most Helpful Reviews

**Good To The Last Drop**
Cuisinart Brew Central DCC-1200 12-Cup Coffee Maker

Reviewed by [newtownbeans](#) - Top 1000

Author's rating: ★★★☆☆

The Cuisinart Brew Central Coffeemaker, Model DCC-1200 is a solidly built and efficient electric coffee maker that brews really fresh and aromatic coffee. It is a bit larger than most coffee makers, measuring approx. 14x8x8 inches and it is available in ... [Read the full review](#)

Compare prices for Cuisinart Brew Central DCC-1200 12-Cup Coffee Maker
Trust Relations

Product Ratings
Similarity between users increases before and after trust.

Difference in ratings decreases before and after trust.
Estimating strength of relations + Predicting user ratings

Does he/she like this type of product?

Influence from different trusted friends
Supporting Communication in Global Networks