Modeling Multilingual Grounded Language

BILL DOLAN
NATURAL LANGUAGE PROCESSING GROUP, MICROSOFT RESEARCH

XLITE: CROSS-LINGUAL TECHNOLOGIES
DECEMBER 7, 2012
Same Intent, Different Words

Make sure the cut sides are facing up.

The beveled edge should be on top.

Check that the diagonal-cut sides are on the top of the seat.

斜边应在顶部

A borda chanfrada deve enfrentar

ыман ряхे कि काष्ट के टुकड़े ऊपर की तरफ हो।

Скошенные края должны быть сверху.
Grounded Multilingual Semantics

Goal: a pragmatic, application-driven model of natural language semantics that is:

- Able to handle both monolingual and multilingual alternations in a robust way
- Data-driven
- Grounded, actionable: captures a mapping between language and objects/actions in the world

What clusters of utterances seem to “mean the same thing”, as far as some desired action in the world is concerned?
Agenda

Translation
- Statistical machine translation as a semantic modeling framework

Paraphrase as Translation
- A statistical machine translation model of paraphrase generation

Grounding as Translation
- Learning to map language to objects, states, events, and actions
Statistical Machine Translation
If you have a hammer...

The Pivot

Translate from: Russian (Auto-Detected)
Translate to: English

Enter text or webpage URL

During the times of Dickens and Dostoenvsky
Vivre près d’un garage ou une station d’essence pourrait quadrupler le risque de leucémie infantile, suggère une étude française.

Vivir cerca de una gasolinera puede llegar a cuadriplicar el riesgo de leucemia en niños.

A child who lives near a petrol (gas) station is four times more likely to develop leukemia than a child who lives far away from one, according to a new study.
Bilingual Alignment

<table>
<thead>
<tr>
<th>quadrupler</th>
<th>four</th>
<th>times</th>
<th>more</th>
<th>likely</th>
<th>to</th>
<th>develop</th>
<th>leukemia</th>
</tr>
</thead>
<tbody>
<tr>
<td>le</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>risque</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leucémie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>infantile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Pivot

Translators can (and do) translate the same phrases in one language into different phrases in another.

- Translators differ
- Contexts differ

Statistical machine translation models utilize parallel corpora created by human translators.

- Two or more sentences, pivoted on a single semantic “intent”
- Large datasets are potentially available
Paraphrase as Translation
Defining Paraphrase

A paraphrase is a pair of formally different expressions that can be interpreted as semantically equivalent in some context

- A good approximation of semantic equivalence is ordinarily sufficient
- Exact synonymy or semantic equivalence (as in entailment relations) is usually of less importance.

Note: paraphrase is mostly accidental

- Paraphrase usually arises because humans independently approach the same information using different forms of expression.
A child who lives near a petrol (gas) station is four times more likely to develop leukemia than a child who lives far away from one, according to a new study.

Living near to a petrol station or garage may increase the risk of acute childhood leukaemia by 400%.

Children who live in close proximity to gas stations and auto body shops have a dramatically higher rate of leukemia, according to a new study.

Living near a petrol station may quadruple the risk for children of developing leukaemia, new research says.

Children who live near petrol stations may be four times more susceptible to leukaemia.
## Monolingual Alignment

<table>
<thead>
<tr>
<th>quadruple</th>
<th>four</th>
<th>times</th>
<th>more</th>
<th>likely</th>
<th>to</th>
<th>develop</th>
<th>leukemia</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>for</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>developing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leukemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
An Uncommon Event

A four-metre python exploded when it tried to swallow a two-metre alligator whole in Florida. Animal experts in the US are baffled after a huge snake exploded when trying to snack on a giant alligator.

Scientists in Florida are puzzling over a Burmese python that scarfed down a six-foot alligator before its stomach ruptured.

A python in Florida's Everglades clashed with a 6-foot-long alligator. And won. But then tried to swallow him whole. And exploded.

Scientists in Florida still aren't quite sure how a 13-foot Burmese python managed to devour a six-foot alligator in the Everglades.

A 13-foot Burmese python recently burst after it apparently tried to swallow a live, six-foot alligator whole, authorities said.
SMT for Paraphrase

Use news text to build statistical translation models of paraphrase:

- *Quirk et al.*, 2004
- *Dolan et al.*, 2004

Microsoft Research Paraphrase Corpus (2005):

- 5800 paraphrase pairs extracted from news articles
- *Dolan & Brockett*, 2005
- *Brockett & Dolan*, 2005
A Stumbling Block

Naturally-occurring paraphrase data is hard to obtain

- There is no such profession as a “paraphraser”
- Full-sentence paraphrases are relatively rare
  - Most paraphrases, even in news data, are fragmentary
  - Usually a single source, edited
- People find it difficult to paraphrase

A major challenge for ML approaches to the problem of modeling monolingual alternations.
The Pivot

Translators can (and do) translate the same phrases in one language into different phrases in another.

- Translators differ
- Contexts differ

Statistical machine translation models utilize parallel corpora created by human translators.

- Two or more sentences, pivoted on a single semantic “intent”
- Large datasets are potentially available.

Aligned data intended for use in SMT models can be a rich source of paraphrase data, if we pivot on the foreign language phrases.

- Bannard & Callison-Burch, 2005
- Callison-Burch et al. 2006
- Callison-Burch, 2008
- Zhao, et al., 2008
# Pivots as Bipartite Graphs

## Japanese - English

<table>
<thead>
<tr>
<th>セルホン</th>
<th>cellphone</th>
</tr>
</thead>
<tbody>
<tr>
<td>セルホン</td>
<td>cellular</td>
</tr>
<tr>
<td>セルホン</td>
<td>cellular phone</td>
</tr>
<tr>
<td>携帯電話</td>
<td>cell</td>
</tr>
<tr>
<td>携帯電話</td>
<td>cell phone</td>
</tr>
<tr>
<td>携帯電話</td>
<td>cellphone</td>
</tr>
<tr>
<td>携帯電話</td>
<td>cellular phone</td>
</tr>
<tr>
<td>携帯電話</td>
<td>mobile</td>
</tr>
<tr>
<td>携帯電話</td>
<td>mobile phone</td>
</tr>
<tr>
<td>携帯電話</td>
<td>phone</td>
</tr>
</tbody>
</table>

## Chinese - English

<table>
<thead>
<tr>
<th>手机</th>
<th>cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>手机</td>
<td>cell phone</td>
</tr>
<tr>
<td>手机</td>
<td>cellphone</td>
</tr>
<tr>
<td>手机</td>
<td>cellular phone</td>
</tr>
<tr>
<td>手机</td>
<td>mobile</td>
</tr>
<tr>
<td>手机</td>
<td>mobile phone</td>
</tr>
</tbody>
</table>
Under the Hood

携帯 ||| " , for mobile
携帯 ||| " applicant for registration "
携帯 ||| " applicant for registration
携帯 ||| " applicant for
携帯 ||| " applicant
携帯 ||| " campaign so that mobile
携帯 ||| " cell
携帯 ||| " cellphones
携帯 ||| " cellular phone
携帯 ||| " diamond " mobile
携帯 ||| " encouraging " mobile carriers to
携帯 ||| " encouraging " mobile carriers
携帯 ||| " encouraging " mobile
携帯 ||| " from mobile
携帯 ||| " how cheap mobile
携帯 ||| " in the mobile
携帯 ||| " keitai credit
携帯 ||| " keitai credit
携帯 ||| " keitai
携帯 ||| " mobile carriers to
携帯 ||| " mobile carriers
携帯 ||| " mobile
携帯 ||| " newsy
携帯 ||| " on your
携帯 ||| " portability
携帯 ||| " portable
携帯 ||| " the mobile
携帯 ||| " the service promises to provide mobile
携帯 ||| " when cell
携帯 ||| " which controls home appliance using cellular
携帯 ||| " wireless
携帯 ||| " you hear the phone
携帯 ||| # on your mobile
携帯 ||| & cell
携帯 ||| & mobile
携帯 ||| > cell
携帯 ||| > regional information - mobile
携帯 ||| ' a mobile
携帯 ||| ' awareness
携帯 ||| ' cell
携帯 ||| ' cellphones and
Bilingual Data, Multilingual Space
A Random Walk Model

Treat the pivot data as a multipartite graph
- *Kok & Brockett, NAACL/HLT 2010*
- Implement a random walk over the graph
- MT pivot meets graphical models for search (*Mei, et al., CIKM 2008*)
- Use translations among other languages as “bridge” languages

Phrase tables from Europarl parallel data:
- 6 languages, 15 language pairs, 1M sentences ea.

Results from random 100 phrase sample (1...4-grams)
- Found paraphrases for all (vs. 49 in CCB 2008)
- 0.61 good at top-1
- 0.42 good in top-10 (~3 x CCB 2008)
Bridge Languages

```
F2  F1  F3  G1
E3  E2  E1  E4  E5

“autoroute”  “grand route”  “route principale”  “Hauptstrasse”

“motorway”  “highway”  “major road”  “main road”  “main street”
```
SMT for Paraphrase Generation

Desiderata:
- Generate paraphrases, up to and including whole sentences
- Provide a robust delivery system
- Add in data from other sources
- Generate new contextual paraphrase candidates – not just memorized ones
SMT for Paraphrase Generation

Solution:
- Exploit our existing phrasal SMT pipeline and large language models
- Use bipartite phrasal mappings as training corpus
  - Currently: 350M paired phrases from 17 language pairs
  - Future: 560 M paired phrases from 24 language pairs
  - Require each phrase to have been seen in at least 2 language pairs
- Build a monolingual phrasal machine translation system
  - Train weights using Max BLEU on 2K sentence pair data set selected for diversity
- Add features derived from original translation scores
The Quality Bar

SMT meets an information need that cannot otherwise be met
- A user may be satisfied with low-grade readable output
- n-best lists are exposed only in special cases

With paraphrase, the grammatical quality bar may be higher
- The human user knows what the input means, is a fluent speaker
- We will be exposing the n-best list, not normal in SMT

Present initial prototype as a contextual thesaurus
Contextual Thesaurus

Translate from English to English to explore alternate ways of expressing the same idea. (Learn more...)
Contextual Thesaurus

Characteristics of the English-English phrasal MT system

- It tends to prefer more common words and sequences
- It has no grammatical knowledge
- Output tends to be monotone, with only local inversions
- Output is best on short sequences of words -- 3-8 words typically
- Better in certain domains: business, finance, government
Paraphrase Public API

Contextual Thesaurus web page:
http://labs.microsoftTranslator.com/thesaurus

Information about public Paraphrase API:

Information about subscribing to the API with an Azure subscription:

Contact to boost subscription limit for research projects:
◦ mtlic@Microsoft.com
Lexical Diversity

Decoding in SMT is NP-complete problem (Knight 1999)
  ◦ Use heuristic limits on search to limit the search space
  ◦ SMT traditionally seeks single best translation -- usually by optimizing BLEU

Paraphrase generation is premised on diversity
  ◦ Most naturally-occurring paraphrases are similar to each other
  ◦ In paraphrase, we often want to expose to the user a selection of alternatives
  ◦ The number of potential alternatives is huge -- a 10 word sentence with 10 synonyms per word yields $10^{10}$ combinations
  ◦ Risky to expose to user because lower ranked alternatives may be of poorer grammatical and semantic quality
Stress Testing

tend to eat under stress
tend to overeat under stress
tend to over-eat under stress
tends to overeat under stress
tended to overeat under stress
tend to overeat in stress
tend to overeat under stress
tendency to overeat under stress
tend to graze under stress
tend to eating under stress
inclined to eat under stress
tends to overeat in stress
tend to overeat on stress
tend to eat too much under stress
inclined to overeat under stress
tend to over-eat in stress
tend to overeat with stress
tend to overeat by stress
tend to binge under stress
tend to overeat under tension
tend to overeat under pressure
tendency to over-eat under stress
tends to overfeed under stress	
tended to overeat in stress
tend to eat too under stress
are inclined to eat under stress
tend to over-eat under tension
tend to over-eat under pressure
usually eating under stress
tends to overeat on stress
tend to overeat the stress
tend to overeat during stress
tendency to overeat in stress
tend to over-eating under stress
tends to graze under stress
tends to over-eat under stress
tend overeating under stress
tends to overeat with stress
generally overeat under stress
tends to overeat by stress
prone to overeat under stress
often overeat under stress
tend to overeat under strain
tend eating under stress
tends to overeat under tension
generally eating under stress
tends to overeat under pressure
tends to overeat under pressure
tend to eat in stress
inclined to eating under stress
The Office Nuisance

a sure way to annoy one’s boss
one sure way to irritate one 's boss
a sure way to irritate one 's boss
sure-fire way to irritate one 's boss
the sure way to irritate one 's boss
surefire way to irritate one 's boss
one sure way to annoy one 's boss
a sure-fire way to irritate one 's boss
one sure way to irritate one 's boss
sure-fire way to irritate one 's boss
the boss
a guaranteed way to irritate one 's boss
one sure way to irritate one 's head
a sure way to annoy one 's boss
a safe way to irritate one 's boss
a certain way to irritate one 's boss
one sure way to antagonize one 's boss
one sure way to upset one 's boss
surefire ways to irritate one 's boss
a sure way to irritate one 's the boss
is sure way to irritate one 's boss
sure-fire way to irritate one 's boss
is a way to irritate one 's boss
one sure way to disturb one 's boss
for sure way to irritate one 's boss
a sure way to irritate one 's head
one sure way to annoy your boss
a way to irritate one 's boss
assured way to irritate one 's boss
sure fire way to irritate one 's boss
a secure way to irritate one 's boss
a sure way to antagonize one 's boss
a sure way to upset one 's boss
surefire way to irritate one 's the boss
one sure way to annoy his boss
reliable way to irritate one 's boss
the sure way to annoy one 's boss
one sure way to tease one 's boss
certain way to irritate one 's boss
safe way to irritate one 's boss
a sure way to disturb one 's boss
the sure way to irritate one 's the boss
a surefire way to irritate one 's head
unsure way to irritate one 's boss
a sure way to annoy your boss
secure way to irritate one 's boss
a surefire way to annoy one 's boss
one sure way to irritate one 's superiors
a reliable way to irritate one 's boss
surefire way to irritate one 's the boss
a surefire way to antagonize one 's boss
one sure way to annoy one 's the boss
one sure way to harass one 's boss
Context is Important

**RETIREMENT ADVICE**
- retirement tips
- pension advice
- pensions advice
- retirement counseling
- retirement guidance
- retirement consulting
- retirement counselling
- pension tips
- pensions tips
- pension counseling
- resignation advice
- superannuation advice
- retirement consultation

**TAX ADVICE**
- tax tips
- tax counselling
- taxation advice
- tax consulting
- tax counseling
- tax advising
- tax consultancy
- tax advise
- tax consultation
- fiscal advice
- tax guidance
- tax advisory
- vat advice
Another kind of Pivot: Query/URL

how to assemble ikea chair
assemble ikea furniture
assemble and disassemble ikea furniture
how to put ikea furniture together
how to put together ikea’s furniture
how difficult is it to assemble ikea furniture
how to put together ikea
ikea directions on how to put ikea furniture together
Ikea do I need tools for self assembly
tools required for assembly of ikea furniture
best way to assemble ikea furniture
directions on how to put together ikea roger chairs
assembling ikea furniture
best way to assemble ikea furniture
is it difficult to build ikea furniture
put together ikea furniture
Grounding as Translation
Non-Linguistic Grounding

So far, we have pivoted on linguistic objects
- Contextual Thesaurus $\rightarrow$ phrases in other languages
- Queries $\rightarrow$ documents

To make linguistic inputs actionable, though, we need data grounded in the world. And we’d like it in many languages.

And it doesn’t exist.

One approach: crowdsourcing
Grounding in Video

Use short video clips to elicit descriptions from people around the world
Parallel descriptions of same activities, in many languages

Translation data

Parallel descriptions in the same language
Paraphrase data (train SMT models on cross-product)

Pivoted on the action in the video

Chen & Dolan, 2011
Crowdsourced Video Captions

A man eats spaghetti sauce.
A man is eating food.
A man is eating from a plate.
A man is eating something.
A man is eating spaghetti.
The man tried his pasta and sauce.
A man is eating spaghetti.
A man is eating.
The man ate some pasta from a bowl.
The man is eating.

A man is eating spaghetti from a large bowl while standing.
A man is eating spaghetti out of a large bowl.
A man tasting some food in the kitchen is expressing his satisfaction.
...in many languages

Slovene: Moški je špagete z vilico.
German: Ein Mann isst Spagetti
Romanian: Un barbat mananca paste.
          Un barbat mananca spaghetti.
          Un bucatar mananca ce a preparat.
French: Un homme mande des pates.
Spanish: Un gordo saborea un plato de pasta
Serbian: Čovek jede špagete.
Filipino: Linasahan ng kusinero ang kanyang pagkain.
Dutch: De luie kok neemt gulzig een hap van zijn bord spaghetti met worstjes.
Russian: Мужчина что-то ест из тарелки.
Tamil: ஒருவர் சாப்பிட்டுக்காண்டு இருக்கிறார்.
        ஒருவர் முள் ரண்டியால் உணவவசாப்பிட்டு இருக்கிறார்.
        மனிதன் சாப்பிட்டுக்காண்டு இருக்கிறார்.
        மனிதன் சாப்பிட்டு காண்து இருக்கிறார்.
## Language Distribution

<table>
<thead>
<tr>
<th>Language</th>
<th>Count</th>
<th>Language</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>85550 (33855)</td>
<td>Spanish</td>
<td>1883</td>
</tr>
<tr>
<td>Hindi</td>
<td>6245</td>
<td>Gujarati</td>
<td>1437</td>
</tr>
<tr>
<td>Romanian</td>
<td>3998</td>
<td>Russian</td>
<td>1243</td>
</tr>
<tr>
<td>Slovene</td>
<td>3584</td>
<td>French</td>
<td>1226</td>
</tr>
<tr>
<td>Serbian</td>
<td>3420</td>
<td>Italian</td>
<td>953</td>
</tr>
<tr>
<td>Tamil</td>
<td>2789</td>
<td>Georgian</td>
<td>907</td>
</tr>
<tr>
<td>Dutch</td>
<td>2735</td>
<td>Polish</td>
<td>544</td>
</tr>
<tr>
<td>German</td>
<td>2326</td>
<td>Chinese</td>
<td>494</td>
</tr>
<tr>
<td>Macedonian</td>
<td>1915</td>
<td>Malayalam</td>
<td>394</td>
</tr>
</tbody>
</table>

Also: Tagalog, Portuguese, Norwegian, Filipino, Estonian, Turkish, Arabic, Urdu, Hungarian, Indonesian, Malay, Bulgarian, Danish, Bosnian, Marathi, Swedish, Albanian
Dataset Publicly Available


~2K videos
~120K captions
~$5K crowd costs
Grounding in an Inventory of Virtual Objects

green color flower design hat

green and yellow winter hat, daisy design on hem and pompom on top
green and brown striped beanie with white daisies on brim and pompom on top
light brown striped hat with flower pattern.
green color cute hat with white color flowers
dark and light green winter hat with white and yellow flower on border
green, flower print elastic clothing hat
round green striped beanie with flowers on around sides
green striped beanie cap with daisies around bottom and green puffball
brown hat with yellow bands and flower pattern
green and brown mixed color round hat, round ball on top
green striped knit beanie with a green ball on top and daisy pattern rim
drab and yellow beanie with flowers

a green striped beanie with a knob on top and sunflowers around the edge
poop green and yellow striped woolen hat with daisies
adorable brown and mustard striped skull cap with ball top and daisies on the rim
green bucket hat with white & yellow flower design upon hat
green and beige striped beanie with pompom on top and flowers on brim
a bright toboggan with daisies around the hat and a pompom.
a yellow knit cap with white stars around the bottom edge
green and dark yellow stripped snow hat with daisies going around the rim
green toboggan with flowers and stripes
Linguistic “Neighborhoods”
Grounded Mappings
Demanding Change

Remove measurement lines from background
remove graph lines
Remove grid lines
Remove the rows within the chart
remove all grid lines
Remove horizontal lines.
remove graph lines from chart
remove rulings
Remove lines from background
Remove grid lines from chart.
delete the background lines
no lines
Remove horizontal grid lines.
Take away the horizontal graph lines that show in the background.
Get rid of the lines in the background.
Remove horizontal lines in graph
remove the segment lines
Remove vertical background lines.
take out grid line
remove all grid lines from the chart
remove horizontal lines from behind the bars
Negotiating Change

Please hide the grid on the chart
Let's see how it looks without grid lines.
I think it would look better without the grid lines.
I think those lines in the background make it look cluttered
Please take away the vertical lines that represent the values on the left
We need to remove the background lines from the graph.
What about removing the grid lines?
Can we remove all background lines except for the bottom line?
What if we remove the lines?
Can we remove the background grid?
Let's see how it looks without the background lines
Can you remove result number delineations?
The grid lines aren't really needed, let's get rid of them.
Would no horizontal lines look better?
Would it be better to remove the lines?
Maybe you could take away the lines because they aren't very helpful anyway, and are distracting.
It will look cleaner without the increment lines
Try deleting the grid lines.
Can you remove the scale lines going across on the background?
Grounding in Parameterized Video
Linking Language to Actions and Objects

An old man jumps up in the air.
A man jumps up once.
Cartoon man jumps up in the air.
An old man jumps up off the ground.
An old guy jumps up in the air.
An old guy jumped up in the air in field.
A guy jumps up in a green field.

A woman throws a chair.
Lady throwing a tiny chair.
A woman tosses a chair.
A woman throws a chair at the camera.
Animated lady throws a small chair.
A woman throws a chair.
A woman throws a chair.
A woman in a shawl jumps up in front of a plane.
Lady standing in front of a plane jumps up.
A cartoon woman jumps up in the air.
Lady jumps up in front of a plane.
Woman jumping up with a plane in the background.
The lady jumped in the air next to a biplane.
Woman jumps once in front of plane.

A robot throws an orange ball.
Robot throwing basketball.
A robot threw a ball.
A cartoon robot throws something.
A robot throws a basketball.
Animated robot throws a basketball.
A robot throws an object with his right hand.

A man shrugs at a chair.
A man shrugging his shoulders at a chair.
An old man shrugs emphatically at a chair.
A man spreads his arms in disdain at a chair.
An old man shrugs at a chair.
A man is making a hopeless gesture at a chair.
An old guy puts his hands up in an "I don't know" fashion at a chair.

A clown is nodding his head yes to a mailbox.
A clown moves his head up and down in front of a mailbox.
A clown nods to a mailbox.
Animated clown shakes his fist at a chair.
A clown nodding several times at a chair while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A clown shaking his head while a snowman.
A man is shrugging his shoulders at a chair
An old man makes a hopeless gesture at a chair
An old man shrugs at a chair

男は椅子に向かって肩をすくめている。
老人が椅子の後ろで絶望的なジェスチャーをしている
男が、肩をすくめながら椅子に話しかけている。
Learn, over large volumes of data, how linguistic constructs map to coding constructs

A man is shrugging his shoulders at a chair
An old man makes a hopeless gesture at a chair
An old man shrugs at a chair

男は椅子に向かって肩をすくめている。
老人が椅子の後ろで絶望的なジェスチャーをしている
男が、肩をすくめながら椅子に話しかけている。
老人が雪だるまの後ろで絶望的なジェスチャーをしている。
男は雪だるまに向かって肩をすくめている。

A man is shrugging his shoulders at a snowman.
An old man is making a hopeless gesture behind a snowman.
Animation Captioning

*bares her teeth and looks confrontational.*
*does not wave, and steps forward.*
*doesn’t wave then stomps foot and looks angry*
*gets mad and steps forward.*
*had angry stance*
*has a growling face*
*has more dramatic facial expressions.*
*is acting aggressive.*
*is acting angry*
*is acting angry*
*is acting angry.*
*is acting very angry.*
*is angry and unfriendly*
*is angry looking.*
*is angry, not friendly*
*is demonstrating aggression*
*is flexing.*
*is gesturing angrily.*
*is growling and clinching*
*is making angry expressions*
*is making wrestling moves and an angry face.*
*is more aggressive*
*is pretending to be a sumo wrestler.*
*right one moved his right arm.Left moved both arms and legs*
*is ready to fight*
*is showing aggression.*
*is singing in an angry manner with hands down low*
*is standoffish and gestures angrily.*
*is telling about something that made her angry.*
*is threatening and aggressive*
*is upset which shows by her body language and facial expressions.*
*is very angry*
*is waving.*
*is wearing a ring*
*looks an angry and moves differently.*
*looks ready for a fight.*
*looks very angry.*
*makes angry gestures with mouth and fists.*
*makes an angry face*
*Quiet, then angry.*
*seems angry*
*seems angry.*
*she is mad.*
*shows her teeth looking angry*
*steps up and back and looks angry.*
*takes a step forward.*
*very angry*
Grounded Multilingual Models

Goal: a data-driven, ML framework that simultaneously captures:
• Multilingual + monolingual “neighborhoods” tied to shared semantic “intent”
• Language grounded in objects/actions/code state

• Statistical Machine Translation is a natural fit
  • Developed to capture meaning-preserving string transformations
  • Allows a unified approach to modeling bilingual + monolingual data
  • Can be readily adapted to add grounding features
  • Active research community feeds rapid development
Thanks
Query Expansion as Translation
Query Expansion as SMT

D: best home remedies for cold and flu
Q: how to deal with stuffy nose

Documents and queries as different languages

Cast QE as bridging the language gap via translation
  ◦ E.g. rewrite/translate a query Q to a title of a doc that is relevant to Q

Three lexicon models trained on billions of Q-D pairs mined from clickthrough logs