Machine Learning Shouldn't Be a Black Box

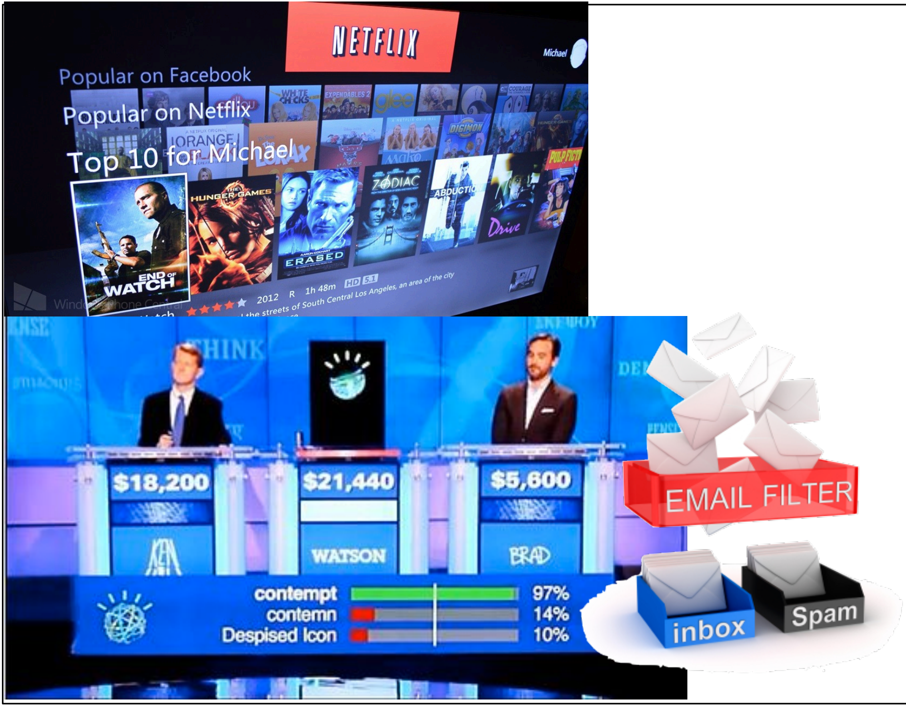
Jordan Boyd-Graber
University of Colorado Boulder

2016

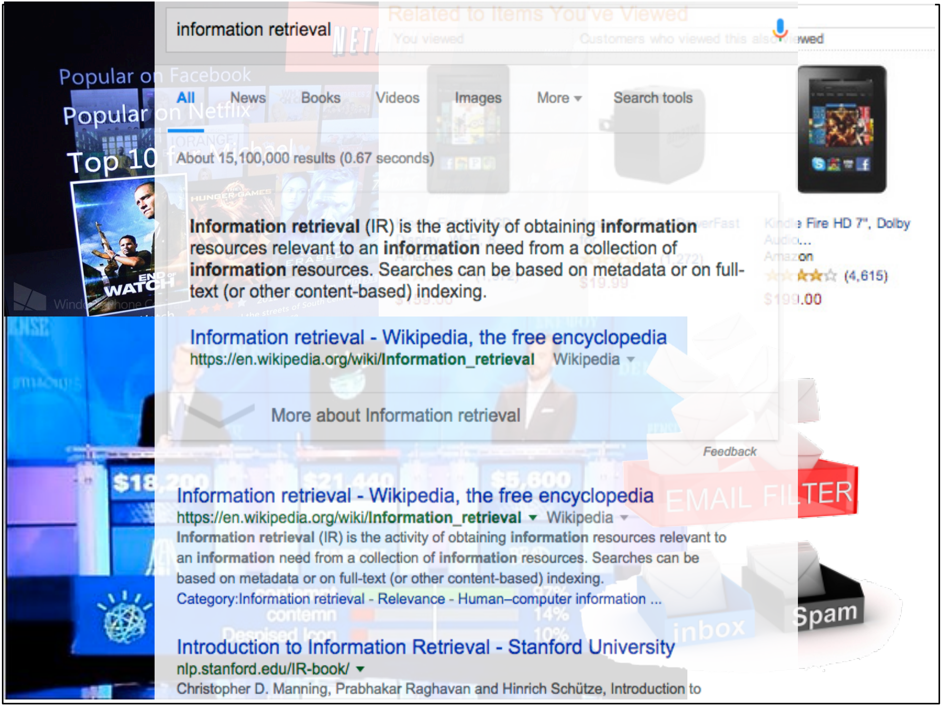

NETFLIX
Michael
Popular on Facebook
Popular on Netflix
Top 10 for Michael
THE HUNGER GAMES
ERASED
ZODIAC
END OF WATCH
Drive
2012 R 1h 48m HD 5.1
the streets of South Central Los Angeles, an area of the city

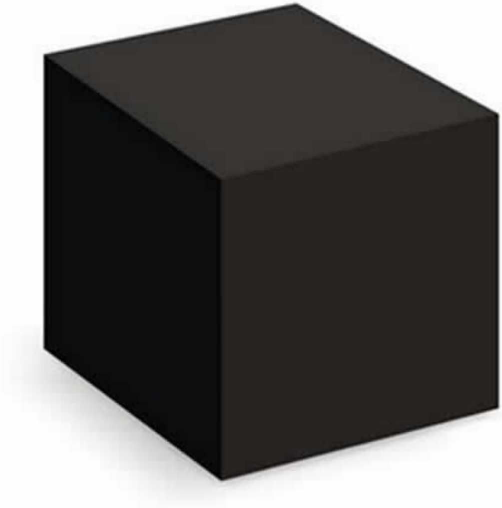

NETFLIX
Michael
Popular on Facebook
Popular on Netflix
Top 10 for Michael
HUNGER GAMES
ZODIAC
ERASED
ABDUCTION
Drive
END OF WATCH
2012 R 1h 48m HD 5.1
the streets of South Central Los Angeles, an area of the city


$18,200
$21,440
$5,600
KEN
WATSON
BRAD
contempt 97%
contemn 14%
Despised Icon 10%


NETFLIX
Michael
Popular on Facebook
Popular on Netflix
Top 10 for Michael
2012 R 1h 48m HD 5.1
$18,200
$21,440
$5,600
KEN
WATSON
BRAD
contempt 97%
contemn 14%
Despised Icon 10%
EMAIL FILTER
inbox
Spam

NETFLIX
Popular on Facebook
Popular on Netflix
Top 10 for Michael
HUNGER GAMES
END OF WATCH
ERASED
2012 R 1h 48m HD 5.1
Related to Items You've Viewed
You viewed
Customers who viewed this also viewed
Kindle Fire 7", LCD Display, Wi-Fi, 8...
Amazon
(1,972)
$159.00
Amazon Kindle PowerFast for...
(1,272)
$19.99
Kindle Fire HD 7", Dolby Audio...
Amazon
(4,615)
$199.00
$18,200
$21,440
$5,600
KEN
WATSON
BRAD
contempt 97%
contemn 14%
Despised Icon 10%
EMAIL FILTER
inbox
Spam


information retrieval
Related to Items You've Viewed
You viewed
Customers who viewed this also viewed
Popular on Facebook
Popular on Netflix
Top 10 for Michael
All
News
Books
Videos
Images
More
Search tools
About 15,100,000 results (0.67 seconds)
Information retrieval (IR) is the activity of obtaining information resources relevant to an information need from a collection of information resources. Searches can be based on metadata or on full-text (or other content-based) indexing.
Kindle Fire HD 7", Dolby Audio...
Amazon
(4,615)
$199.00
Information retrieval - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/Information_retrieval Wikipedia
More about Information retrieval
Feedback
Information retrieval - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/Information_retrieval Wikipedia
Information retrieval (IR) is the activity of obtaining information resources relevant to an information need from a collection of information resources. Searches can be based on metadata or on full-text (or other content-based) indexing.
Category:Information retrieval - Relevance - Human–computer information ...
Introduction to Information Retrieval - Stanford University
nlp.stanford.edu/IR-book/
Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, Introduction to
EMAIL FILTER
inbox
Spam
$18,200
$21,440
$5,600
END OF WATCH
HUNGER GAMES



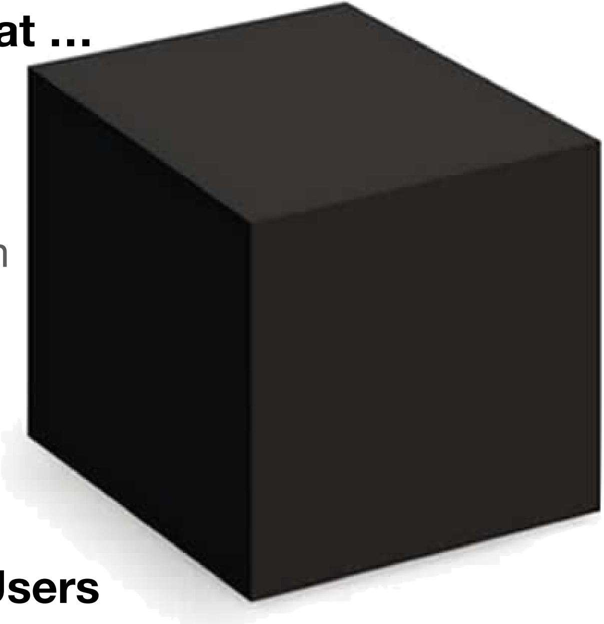
Algorithms that …

Inform

Collaborate with

Compete with

Understand



their Human Users

Algorithms that …

Inform

Collaborate with

Compete with

Understand

their Human Users

The Challenge of Big Data

Every second . . .

- 600 new blog posts appear
- 34,000 tweets are tweeted
- 30 GB of data uploaded to Facebook

The Challenge of Big Data

Every second . . .

- 600 new blog posts appear
- 34,000 tweets are tweeted
- 30 GB of data uploaded to Facebook

Unstructured

No XML, no semantic web, no annotation. Often just raw text.

The Challenge of Big Data

Every second . . .

- 600 new blog posts appear
- 34,000 tweets are tweeted
- 30 GB of data uploaded to Facebook

Unstructured

No XML, no semantic web, no annotation. Often just raw text.

Common task: what's going on in this dataset.

- Intelligence analysts
- Brand monitoring
- Journalists
- Humanists

The Challenge of Big Data

Every second . . .

- 600 new blog posts appear
- 34,000 tweets are tweeted
- 30 GB of data uploaded to Facebook

Unstructured

No XML, no semantic web, no annotation. Often just raw text.

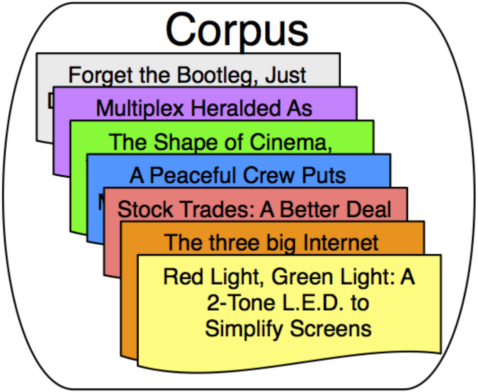
Common task: what's going on in this dataset.

- Intelligence analysts
- Brand monitoring
- Journalists
- Humanists

Common solution: topic models

What does a Topic Model do?

From an **input corpus** and number of topics $K \rightarrow$ words to topics



What does a Topic Model do?

From an input corpus and number of topics K $\rightarrow$ **words to topics**

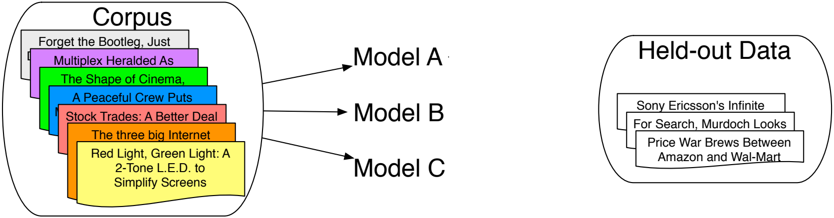
TOPIC 1	TOPIC 2	TOPIC 3
computer, technology, system, service, site, phone, internet, machine	sell, sale, store, product, business, advertising, market, consumer	play, film, movie, theater, production, star, director, stage

Evaluating Topic Models

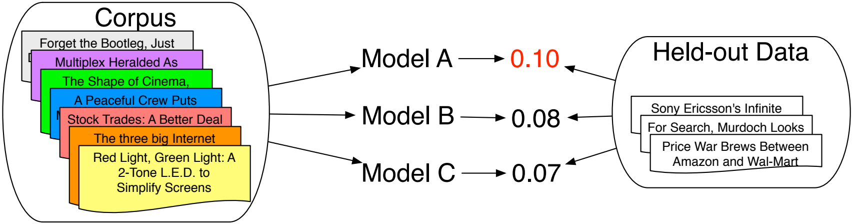
Reading Tea Leaves: How Humans Interpret Topic Models

Jonathan Chang, Jordan Boyd-Graber, Chong Wang, Sean Gerrish, and David M. Blei. Reading Tea Leaves: How Humans Interpret Topic Models. Neural Information Processing Systems, 2009.

Evaluation


Corpus
Forget the Bootleg, Just
Multiplex Heralded As
The Shape of Cinema,
A Peaceful Crew Puts
Stock Trades: A Better Deal
The three big Internet
Red Light, Green Light: A
2-Tone L.E.D. to
Simplify Screens
Model A
Model B
Model C
Held-out Data
Sony Ericsson's Infinite
For Search, Murdoch Looks
Price War Brews Between
Amazon and Wal-Mart

Evaluation



Measures predictive power (likelihood / perplexity)

Qualitative Evaluation of the Latent Space

"segment 1"	"segment 2"	"matrix 1"	"matrix 2"	"line 1"	"line 2"	"power 1"	power 2"
imag	speaker	robust	manufactur	constraint	alpha	POWER	load
SEGMENT	speech	MATRIX	cell	LINE	redshift	spectrum	memori
texture	recogni	eigenvalu	part	match	LINE	omega	vlsi
color	signal	uncertainti	MATRIX	locat	galaxi	mpc	POWER
tissue	train	plane	cellular	imag	quasar	hsup	systolic
brain	hmm	linear	famili	geometr	absorp	larg	input
slice	source	condition	design	impos	high	redshift	complex
cluster	speakerind.	perturb	machinepart	segment	ssup	galaxi	arrai
mri	SEGMENT	root	format	fundament	densiti	standard	present
volume	sound	suffici	group	recogn	veloc	model	implement

Figure 3: Eight selected factors from a 128 factor decomposition. The displayed word stems are the 10 most probable words in the class-conditional distribution $P(w|z)$, from top to bottom in descending order.

[Hofmann 1999]

Qualitative Evaluation of the Latent Space

"Arts"	"Budgets"	"Children"	"Education"
NEW	MILLION	CHILDREN	SCHOOL
FILM	TAX	WOMEN	STUDENTS
SHOW	PROGRAM	PEOPLE	SCHOOLS
MUSIC	BUDGET	CHILD	EDUCATION
MOVIE	BILLION	YEARS	TEACHERS
PLAY	FEDERAL	FAMILIES	HIGH
MUSICAL	YEAR	WORK	PUBLIC
BEST	SPENDING	PARENTS	TEACHER
ACTOR	NEW	SAYS	BENNETT
FIRST	STATE	FAMILY	MANIGAT
YORK	PLAN	WELFARE	NAMPHY
OPERA	MONEY	MEN	STATE
THEATER	PROGRAMS	PERCENT	PRESIDENT
ACTRESS	GOVERNMENT	CARE	ELEMENTARY
LOVE	CONGRESS	LIFE	HAITI

[Blei et al. 2003]

Qualitative Evaluation of the Latent Space

DA	centralbank europæiske ecb s lån centralbanks
DE	zentralbank ezb bank europäischen investitionsbank darlehen
EL	τράπεζα τράπεζας κεντρική εκτ κεντρικής τράπεζες
EN	**bank central ecb banks european monetary**
ES	banco central europeo bce bancos centrales
FI	keskuspankin ekp n euroopan keskuspankki eip
FR	banque centrale bce européenne banques monétaire
IT	banca centrale bce europea banche prestiti
NL	bank centrale ecb europese banken leningen
PT	banco central europeu bce bancos empréstimos
SV	centralbanken europeiska ecb centralbankens s lån

[Mimno et al. 2009]

Qualitative Evaluation of the Latent Space

(a) Topic labeled as SSL

Keyword	Probability
ssl	0.373722
expr	0.042501
init	0.033207
engine	0.026447
var	0.022222
ctx	0.023067
ptemp	0.017153
mctx	0.013773
lookup	0.012083
modssl	0.011238
ca	0.009548

(b) Topic labeled as Logging

Keyword	Probability
log	0.141733
request	.036017
mod	0.0311
config	0.029871
name	0.023725
headers	0.021266
autoindex	0.020037
format	0.017578
cmd	0.01512
header	0.013891
add	0.012661

Table 2: Sample Topics extracted from Apache source code

[Maskeri et al. 2008]

Qualitative Evaluation of the Latent Space

Probabilistic Models	model word probability set data number algorithm language corpus method
Prosody	prosodic speech pitch boundary prosody phrase boundaries accent repairs intonation
Semantic Roles*	semantic verb frame argument verbs role roles predicate arguments
Yale School Semantics	knowledge system semantic language concept representation information network concepts base
Sentiment	subjective opinion sentiment negative polarity positive wiebe reviews sentence opinions
Speech Recognition	speech recognition word system language data speaker error test spoken
Spell Correction	errors error correction spelling ocr correct corrections checker basque corrected detection
Statistical MT	english word alignment language source target sentence machine bilingual mt
Statistical Parsing	dependency parsing treebank parser tree parse head model al np
Summarization	sentence text evaluation document topic summary summarization human summaries score
Syntactic Structure	verb noun syntactic sentence phrase np subject structure case clause
TAG Grammars*	tree node trees nodes derivation tag root figure adjoining grammar
Unification	feature structure grammar lexical constraints unification constraint type structures rule
WSD*	word senses wordnet disambiguation lexical semantic context similarity dictionary
Word Segmentation	chinese word character segmentation corpus dictionary korean language table system
WordNet*	synset wordnet synsets hypernym ili wordnets hypernyms eurowordnet hyponym ewn wn

Table 2: Top 10 words for 43 of the topics. Starred topics are hand-seeded.

[Hall et al. 2008]

Word Intrusion

1. Take the highest probability words from a topic

Original Topic

dog, cat, horse, pig, cow

Word Intrusion

1. Take the highest probability words from a topic

Original Topic

dog, cat, horse, pig, cow

2. Take a high-probability word from another topic and add it

Topic with Intruder

dog, cat, apple, horse, pig, cow

Word Intrusion

1. Take the highest probability words from a topic

Original Topic

dog, cat, horse, pig, cow

2. Take a high-probability word from another topic and add it

Topic with Intruder

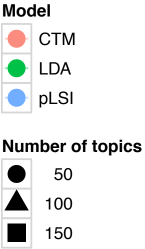
dog, cat, apple, horse, pig, cow

3. We ask users to find the word that doesn't belong

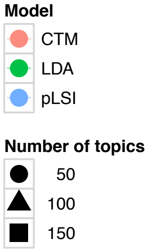
Hypothesis

If the topics are interpretable, users will consistently choose true intruder

Interpretability and Likelihood (NYT)

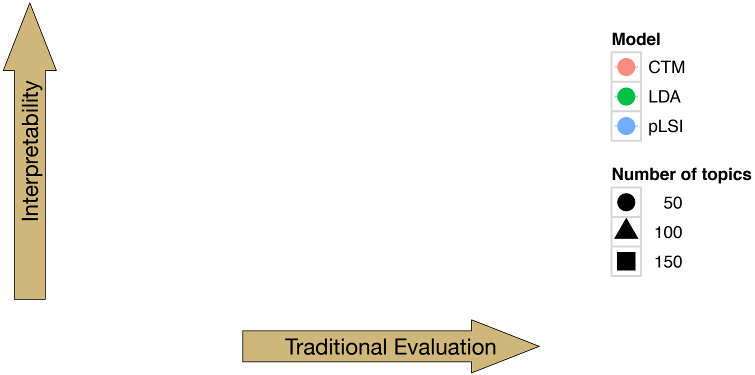


Interpretability and Likelihood (NYT)

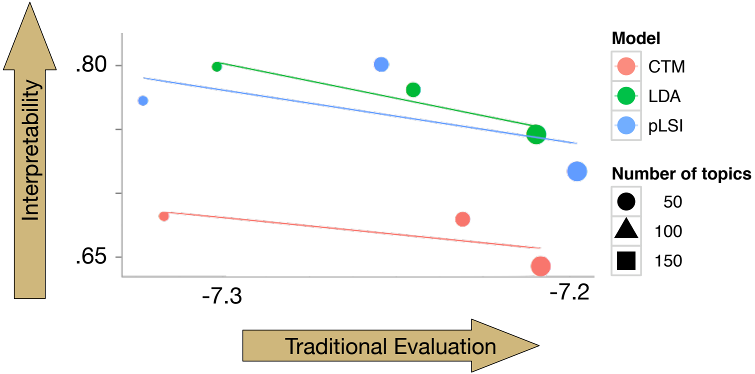


Traditional Evaluation

Interpretability and Likelihood (NYT)



Interpretability and Likelihood (NYT)



Within a model, higher likelihood $\neq$ higher interpretability

Since then . . .

- A way to get at an evaluation that matches **what we care about**
- A necessary step to improving topic models for navigating large datasets [Talley et al. 2011]
- Others have discovered automatic methods that uncover the same properties [Newman et al. 2010, Mimno et al. 2011]
- And extended the technique to structured topics and phrases [Lindsey et al. 2012, Weninger et al. 2012]

Algorithms that …

Inform

Collaborate with

Compete with

Understand

their Human Users

The Problem: User Perspective

bladder
spinal_cord
sci
spinal_cord_injury
spinal
urinary
urothelial
cervical
injury
recovery
urinary_tract
locomotor
lumbar

The Problem: User Perspective

bladder
spinal_cord
sci
spinal_cord_injury
spinal
urinary
urothelial
cervical
injury
recovery
urinary_tract
locomotor
lumbar

The Problem: User Perspective

bladder
spinal_cord
sci
spinal_cord_injury
spinal
urinary
urothelial
cervical
injury
recovery
urinary_tract
locomotor
lumbar

These words don't belong together!

Interactive Topic Modeling

Yuening Hu, Jordan Boyd-Graber, and Brianna Satinoff. Association for Computational Linguistics, 2011.

How to fix it?

bagel phone

constitution

tea

spinal_cord

god

nasa

president

space

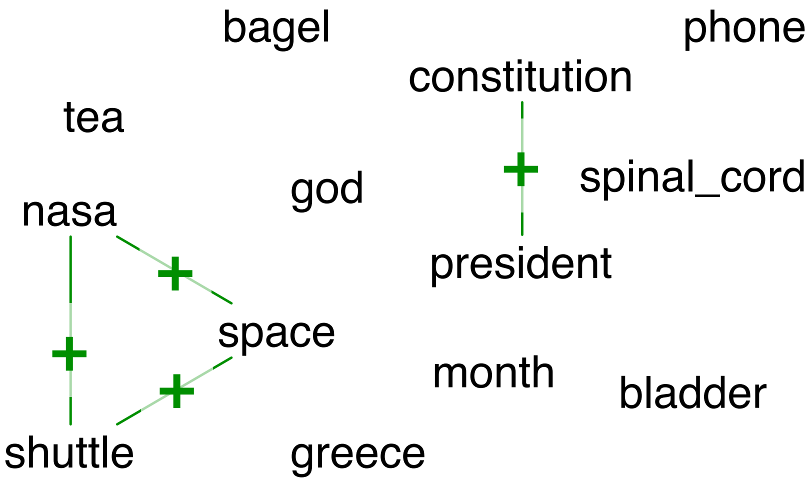
month

bladder

shuttle

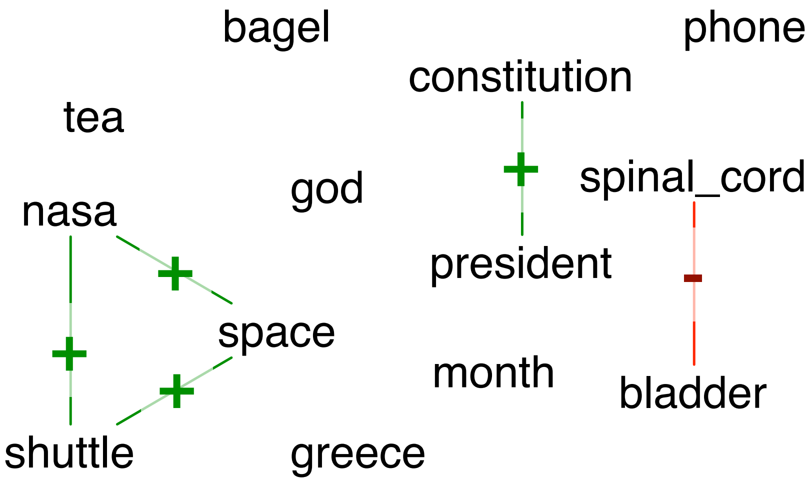
greece

How to fix it?



bagel
phone
constitution
tea
+
spinal_cord
god
nasa
president
+
space
+
month
bladder
shuttle
greece

How to fix it?


bagel
phone
constitution
tea
god
spinal_cord
nasa
president
space
month
bladder
shuttle
greece

Topic	Before
1	election, yeltsin, russian, political, party, democratic, russia, president, democracy, boris, country, south, years, month, government, vote, since, leader, presidential, military
2	new, york, city, state, mayor, budget, giuliani, council, cuomo, gov, plan, year, rudolph, dinkins, lead, need, governor, legislature, pataki, david
3	nuclear, arms, weapon, defense, treaty, missile, world, unite, yet, soviet, lead, secretary, would, control, korea, intelligence, test, nation, country, testing
4	president, bush, administration, clinton, american, force, reagan, war, unite, lead, economic, iraq, congress, america, iraqi, policy, aid, international, military, see
	⋮
20	soviet, lead, gorbachev, union, west, mikhail, reform, change, europe, leaders, poland, communist, know, old, right, human, washington, western, bring, party

Topic	Before
1	election, yeltsin, russian, political, party, democratic, russia, president, democracy, boris, country, south, years, month, government, vote, since, leader, presidential, military
2	new, york, city, state, mayor, budget, giuliani, council, cuomo, gov, plan, year, rudolph, dinkins, lead, need, governor, legislature, pataki, david
3	nuclear, arms, weapon, defense, treaty, missile, world, unite, yet, soviet, lead, secretary, would, control, korea, intelligence, test, nation, country, testing
4	president, bush, administration, clinton, american, force, reagan, war, unite, lead, economic, iraq, congress, america, iraqi, policy, aid, international, military, see
	⋮
20	soviet, lead, gorbachev, union, west, mikhail, reform, change, europe, leaders, poland, communist, know, old, right, human, washington, western, bring, party

Topic	Before
1	election, yeltsin, russian, political, party, democratic, russia, president, democracy, boris, country, south, years, month, government, vote, since, leader, presidential, military
2	new, york, city, state, mayor, budget, giuliani, council, cuomo, gov, plan, year, rudolph, dinkins, lead, need, governor, legislature, pataki, david
3	nuclear, arms, weapon, defense, treaty, missile, world, unite, yet, soviet, lead, secretary, would, control, korea, intelligence, test, nation, country, testing
4	president, bush, administration, clinton, american, force, reagan, war, unite, lead, economic, iraq, congress, america, iraqi, policy, aid, international, military, see
	⋮
20	soviet, lead, gorbachev, union, west, mikhail, reform, change, europe, leaders, poland, communist, know, old, right, human, washington, western, bring, party

Suggestion

boris, communist, gorbachev, mikhail, russia, russian, soviet, union, yeltsin

Topic	Before
1	election, yeltsin, russian, political, party, democratic, russia, president, democracy, boris, country, south, years, month, government, vote, since, leader, presidential, military
2	new, york, city, state, mayor, budget, giuliani, council, cuomo, gov, plan, year, rudolph, dinkins, lead, need, governor, legislature, pataki, david
3	nuclear, arms, weapon, defense, treaty, missile, world, unite, yet, soviet, lead, secretary, would, control, korea, intelligence, test, nation, country, testing
4	president, bush, administration, clinton, american, force, reagan, war, unite, lead, economic, iraq, congress, america, iraqi, policy, aid, international, military, see
	⋮
20	soviet, lead, gorbachev, union, west, mikhail, reform, change, europe, leaders, poland, communist, know, old, right, human, washington, western, bring, party

Topic	After
1	election, democratic, south, country, president, party, africa, lead, even, democracy, leader, presidential, week, politics, minister, percent, voter, last, month, years
2	new, york, city, state, mayor, budget, council, giuliani, gov, cuomo, year, rudolph, dinkins, legislature, plan, david, governor, pataki, need, cut
3	nuclear, arms, weapon, treaty, defense, war, missile, may, come, test, american, world, would, need, lead, get, join, yet, clinton, nation
4	president, administration, bush, clinton, war, unite, force, reagan, american, america, make, nation, military, iraq, iraqi, troops, international, country, yesterday, plan
	⋮
20	soviet, union, economic, reform, yeltsin, russian, lead, russia, gorbachev, leaders, west, president, boris, moscow, europe, poland, mikhail, communist, power, relations

Topic	Before
1	election, yeltsin, russian, political, party, democratic, russia, president, democracy, boris, country, south, years, month, government, vote, since, leader, presidential, military
2	new, york, city, state, mayor, budget, giuliani, council, cuomo, gov, plan, year, rudolph, dinkins, lead, need, governor, legislature, pataki, david
3	nuclear, arms, weapon, defense, treaty, missile, world, unite, yet, soviet, lead, secretary, would, control, korea, intelligence, test, nation, country, testing
4	president, bush, administration, clinton, american, force, reagan, war, unite, lead, economic, iraq, congress, america, iraqi, policy, aid, international, military, see
	⋮
20	soviet, lead, gorbachev, union, west, mikhail, reform, change, europe, leaders, poland, communist, know, old, right, human, washington, western, bring, party

Topic	After
1	election, democratic, south, country, president, party, africa, lead, even, democracy, leader, presidential, week, politics, minister, percent, voter, last, month, years
2	new, york, city, state, mayor, budget, council, giuliani, gov, cuomo, year, rudolph, dinkins, legislature, plan, david, governor, pataki, need, cut
3	nuclear, arms, weapon, treaty, defense, war, missile, may, come, test, american, world, would, need, lead, get, join, yet, clinton, nation
4	president, administration, bush, clinton, war, unite, force, reagan, american, america, make, nation, military, iraq, iraqi, troops, international, country, yesterday, plan
	⋮
20	soviet, union, economic, reform, yeltsin, russian, lead, russia, gorbachev, leaders, west, president, boris, moscow, europe, poland, mikhail, communist, power, relations

Topic	Before
1	election, yeltsin, russian, political, party, democratic, russia, president, democracy, boris, country, south, years, month, government, vote, since, leader, presidential, military
2	new, york, city, state, mayor, budget, giuliani, council, cuomo, gov, plan, year, rudolph, dinkins, lead, need, governor, legislature, pataki, david
3	nuclear, arms, weapon, defense, treaty, missile, world, unite, yet, soviet, lead, secretary, would, control, korea, intelligence, test, nation, country, testing
4	president, bush, administration, clinton, american, force, reagan, war, unite, lead, economic, iraq, congress, america, iraqi, policy, aid, international, military, see
	⋮
20	soviet, lead, gorbachev, union, west, mikhail, reform, change, europe, leaders, poland, communist, know, old, right, human, washington, western, bring, party

Topic	After
1	election, democratic, south, country, president, party, africa, lead, even, democracy, leader, presidential, week, politics, minister, percent, voter, last, month, years
2	new, york, city, state, mayor, budget, council, giuliani, gov, cuomo, year, rudolph, dinkins, legislature, plan, david, governor, pataki, need, cut
3	nuclear, arms, weapon, treaty, defense, war, missile, may, come, test, american, world, would, need, lead, get, join, yet, clinton, nation
4	president, administration, bush, clinton, war, unite, force, reagan, american, america, make, nation, military, iraq, iraqi, troops, international, country, yesterday, plan
	⋮
20	soviet, union, economic, reform, yeltsin, russian, lead, russia, gorbachev, leaders, west, president, boris, moscow, europe, poland, mikhail, communist, power, relations

Example: Negative Constraint

Topic	Words
318	bladder, sci, spinal_cord, spinal_cord_injury, spinal, urinary, urinary_tract, urothelial,injury, motor, recovery, reflex, cervical, urothelium, functional_recovery

Example: Negative Constraint

Topic	Words
318	bladder, sci, spinal_cord, spinal_cord_injury, spinal, urinary, urinary_tract, urothelial, injury, motor, recovery, reflex, cervical, urothelium, functional_recovery

Negative Constraint

spinal_cord, bladder

Example: Negative Constraint

Topic	Words
318	bladder, sci, spinal_cord, spinal_cord_injury, spinal, urinary, urinary_tract, urothelial,injury, motor, recovery, reflex, cervical, urothelium, functional_recovery

Topic	Words
318	sci, spinal_cord, spinal_cord_injury, spinal, injury, recovery, motor, reflex, urothelial, injured, functional_recovery, plasticity, locomotor, cervical, locomotion

Negative Constraint

spinal_cord, bladder

ALTO: Active Learning with Topic Overviews for Speeding Label Induction and Document Labeling

Forough Poursabzi-Sangdeh, Jordan Boyd-Graber, Leah Findlater, and Kevin Seppi. Association for Computational Linguistics, 2016.

Real-World Use Cases

Algorithms that …

Inform

Collaborate with

Compete with

Understand

their Human Users

When you at the dark side **look**, careful you must **be**.

ich bin mit dem Zug nach Ulm **gefahren**
I am with the train to Ulm **traveled**

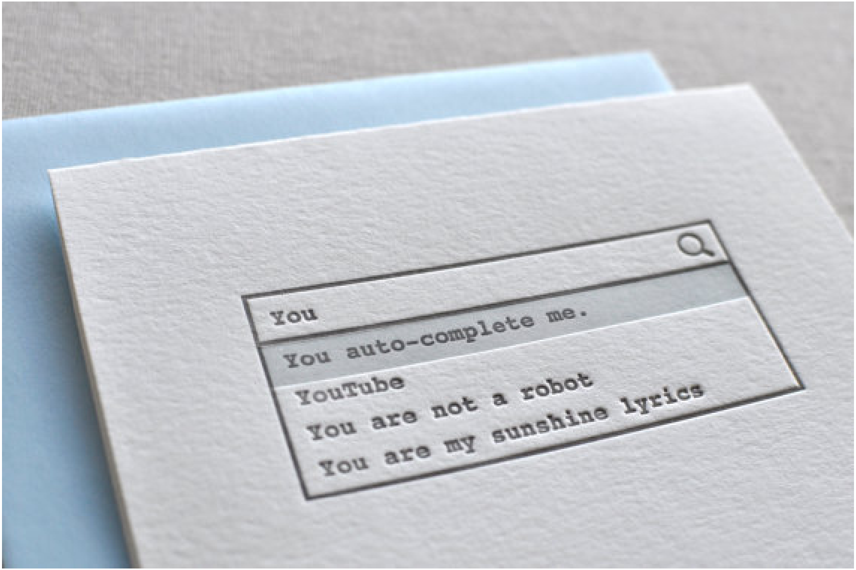
I *(......waiting......)* **traveled** by train to Ulm

Learning from Interpreters

- What tricks do they use?
- How can we teach machines to use them?
- How do we know when to use them?
- Giving back to interpreters

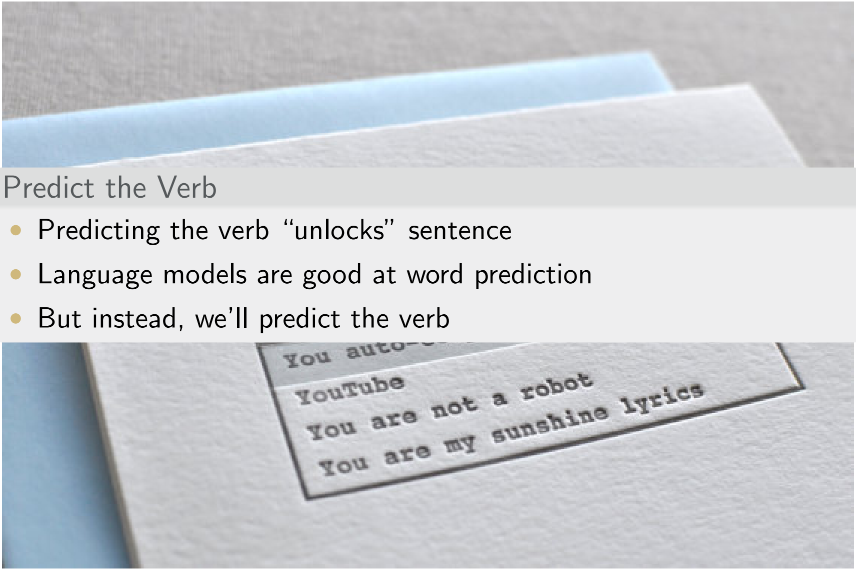
Don't Until the Final Verb Wait: Reinforcement Learning for Simultaneous Machine Translation

Alvin Grissom II, **Jordan Boyd-Graber**, He He, John Morgan, and Hal Daumé **III**. *Empirical Methods in Natural Language Processing*, 2014


You
You auto-complete me.
YouTube
You are not a robot
You are my sunshine lyrics

Predict the Verb

- Predicting the verb "unlocks" sentence
- Language models are good at word prediction
- But instead, we'll predict the verb



Language Models of Verbs

Apple ist zum wertvollsten Konzern aller Zeiten avanciert

Nein, mit dem Virus ist es noch lange nicht getan

Eine vielbefahrene Brücke in New Jersey wurde grundlos gesperrt

Mit Drohen und Interpretieren ist es nicht getan

Frankfurter Flughafen für Passagiere weitgehend gesperrt

Als ruppiger Bad Boy mit Herz namens Daryl ist er zum Superstar der Besetzung avanciert

Language Models of Verbs

Apple ist zum wertvollsten Konzern aller Zeiten avanciert

Nein, mit dem Virus ist es noch lange nicht getan

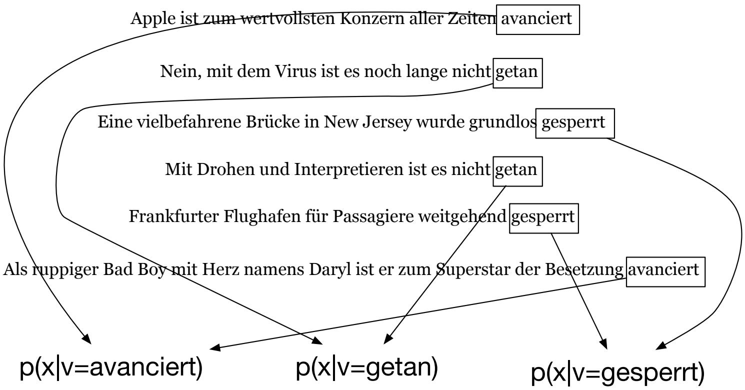
Eine vielbefahrene Brücke in New Jersey wurde grundlos gesperrt

Mit Drohen und Interpretieren ist es nicht getan

Frankfurter Flughafen für Passagiere weitgehend gesperrt

Als ruppiger Bad Boy mit Herz namens Daryl ist er zum Superstar der Besetzung avanciert

Language Models of Verbs



Predicting the Verb

- Build language model for every verb
- Then, for any input text x we can make a prediction of the verb

$$\arg\max_v p(v) \prod_{i=1}^{t} p(x_i \mid v, x_{i-n+1:i-1}) \tag{1}$$

Predicting the Verb

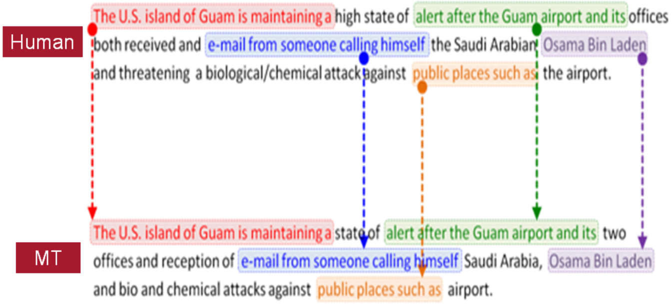
- Build language model for every verb
- Then, for any input text x we can make a prediction of the verb

$$\arg\max_{v} p(v) \prod_{i=1}^{t} p(x_i \mid v, x_{i-n+1:i-1}) \tag{1}$$

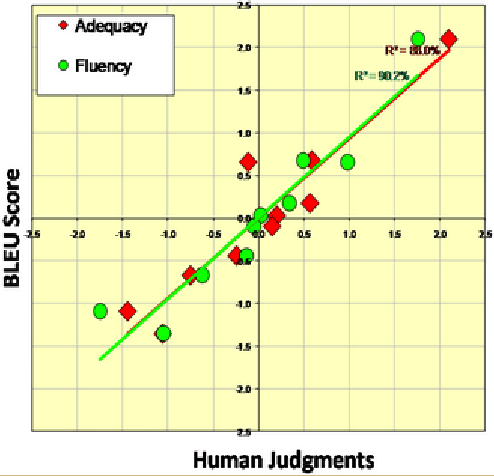
- Most of these predictions will be totally wrong (18% accuracy) ...
- leading to horrible translations

Scoring one Translation

Bilingual Evaluation Understudy (BLEU)

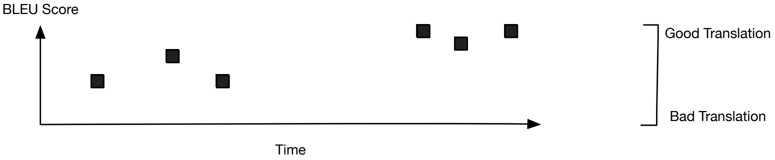


Bilingual Evaluation Understudy (BLEU)



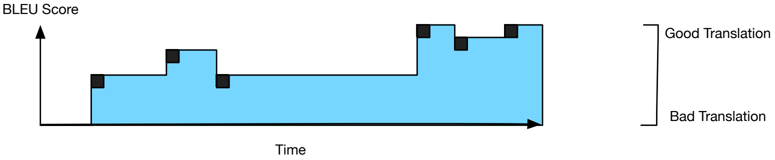
Scoring a series of Translations

Bilingual Evaluation Understudy (BLEU)

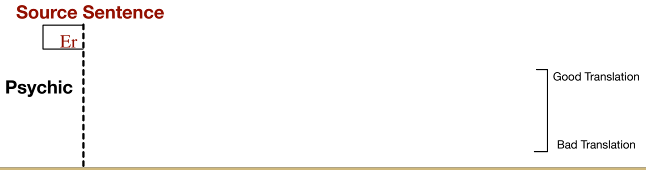


Scoring a series of Translations

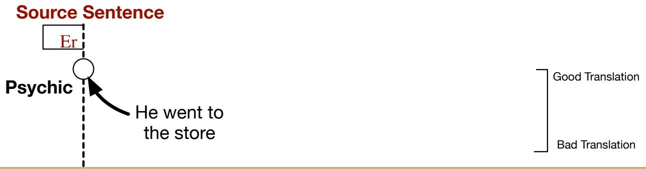
Bilingual Evaluation Understudy (BLEU)



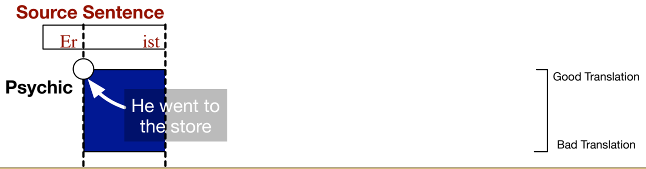
Comparing Policies



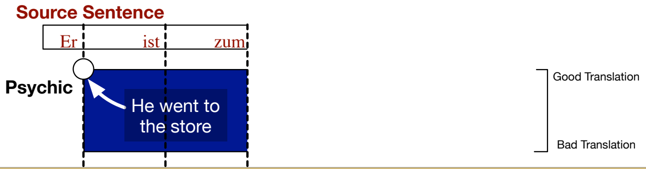
Comparing Policies



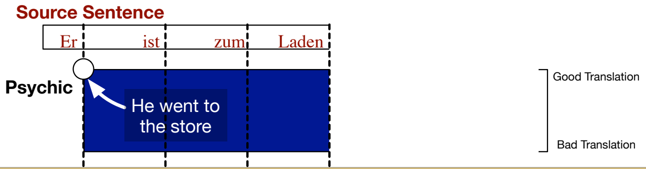
Comparing Policies



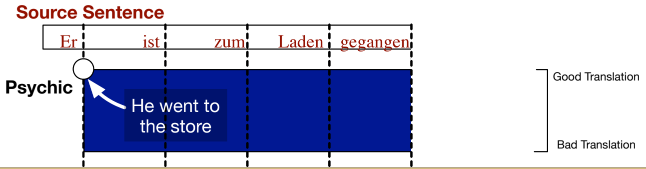
Comparing Policies



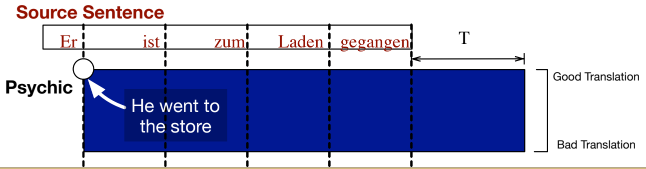
Comparing Policies



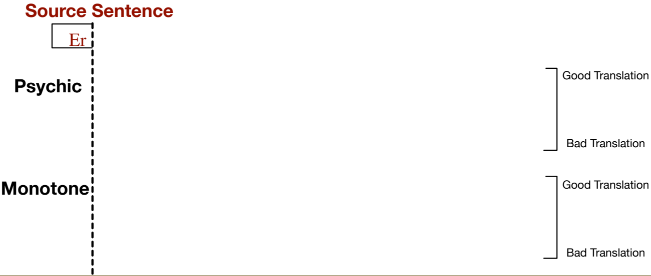
Comparing Policies



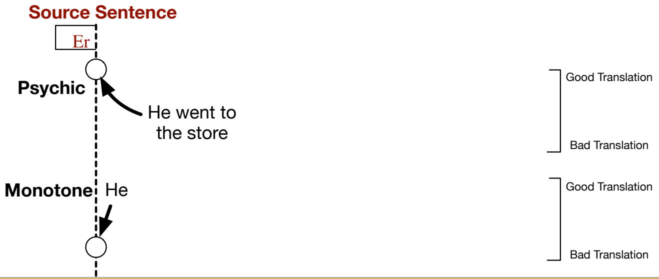
Comparing Policies



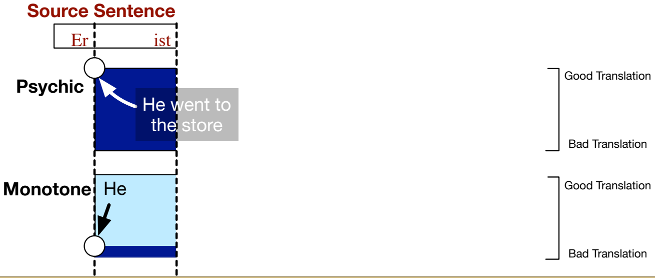
Comparing Policies



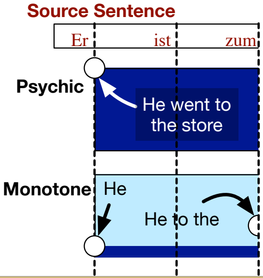
Comparing Policies

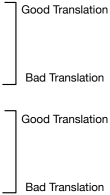


Comparing Policies

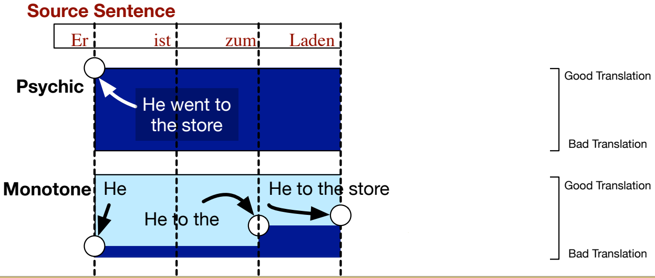


Comparing Policies

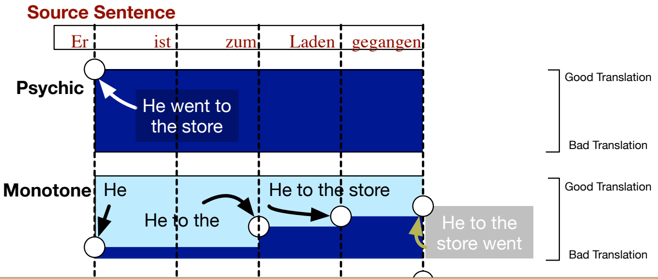




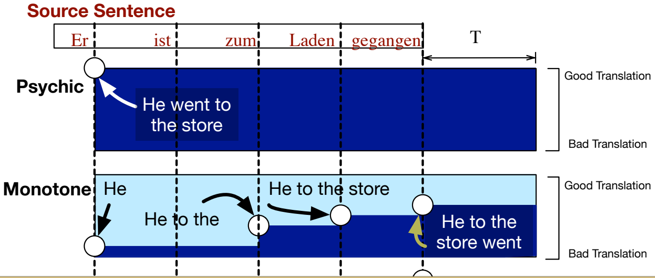
Comparing Policies



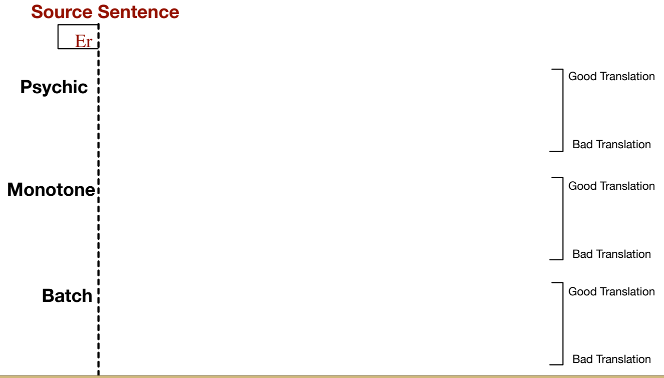
Comparing Policies



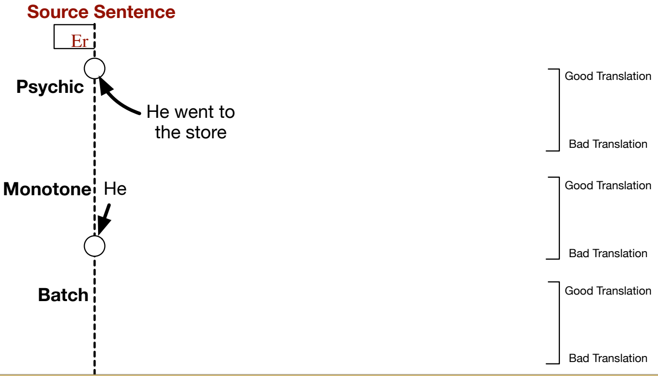
Comparing Policies



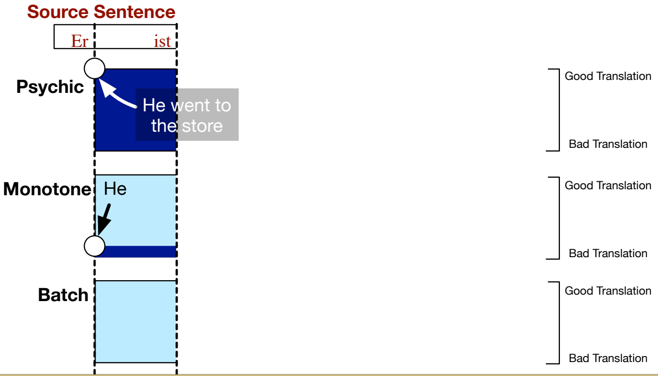
Comparing Policies



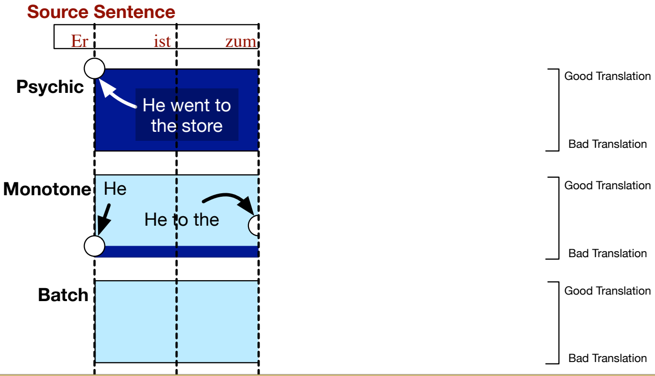
Comparing Policies



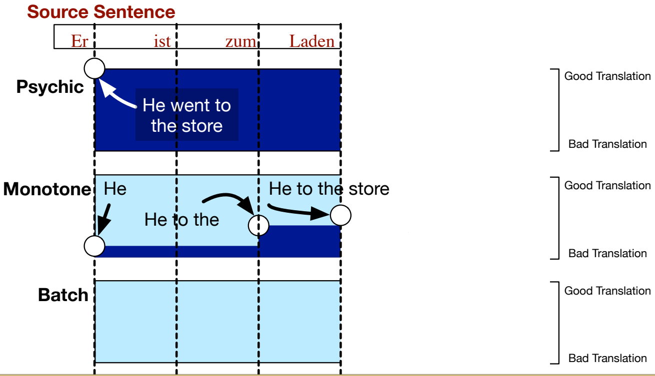
Comparing Policies



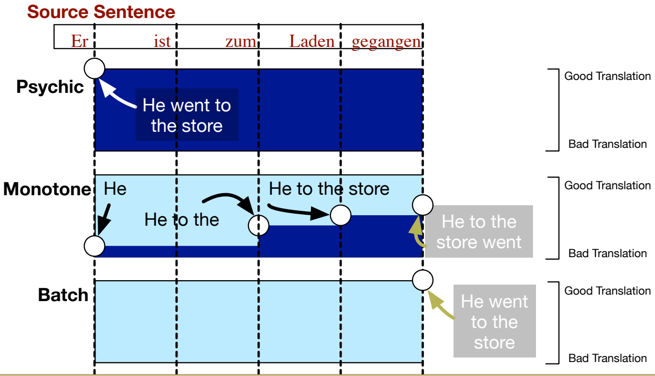
Comparing Policies



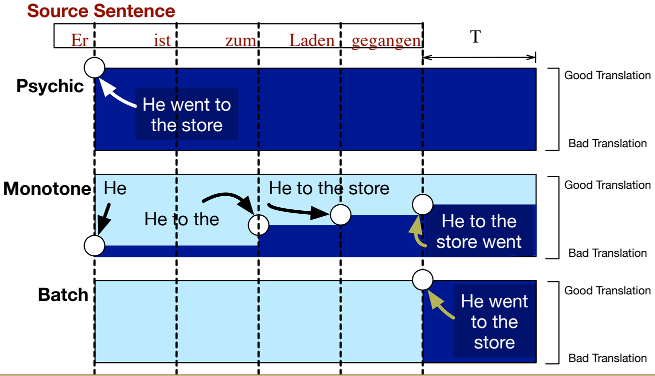
Comparing Policies



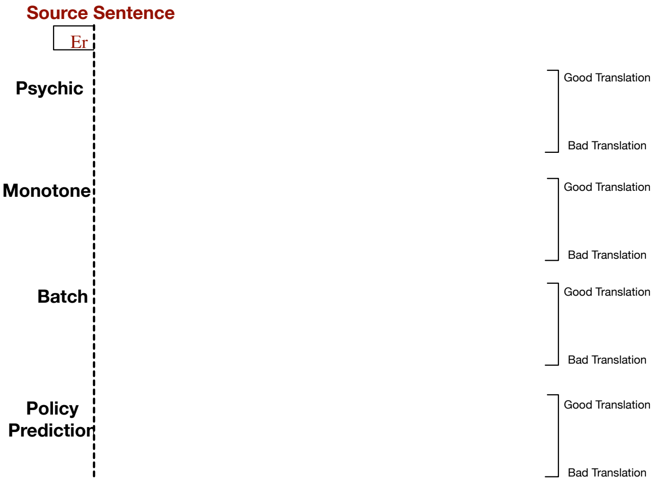
Comparing Policies



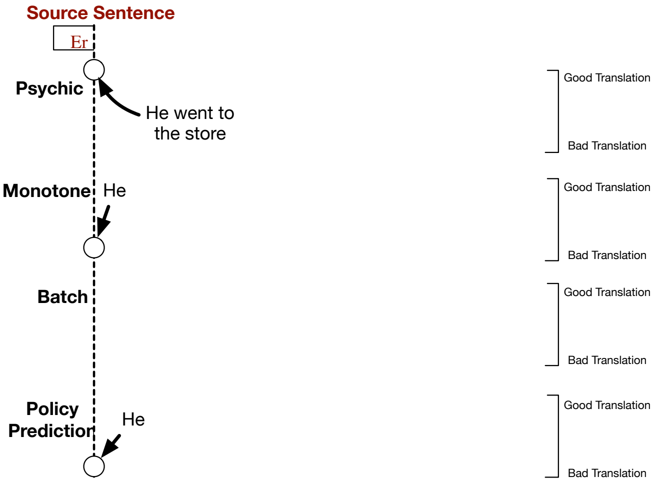
Comparing Policies



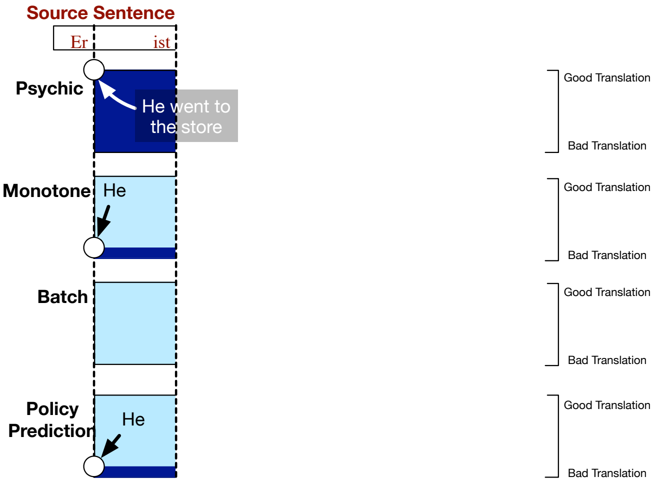
Comparing Policies



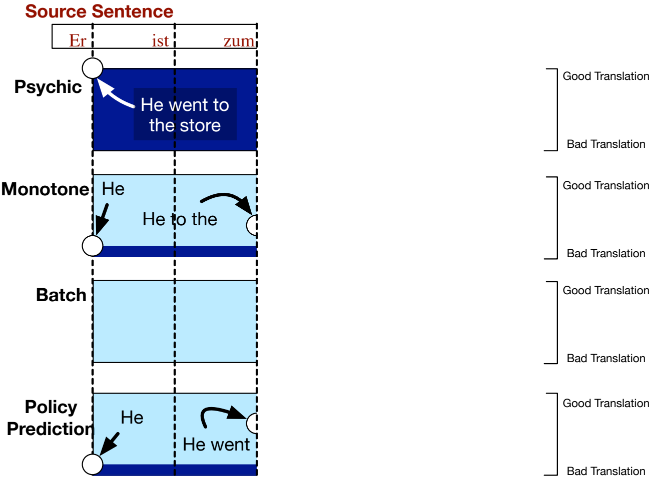
Comparing Policies



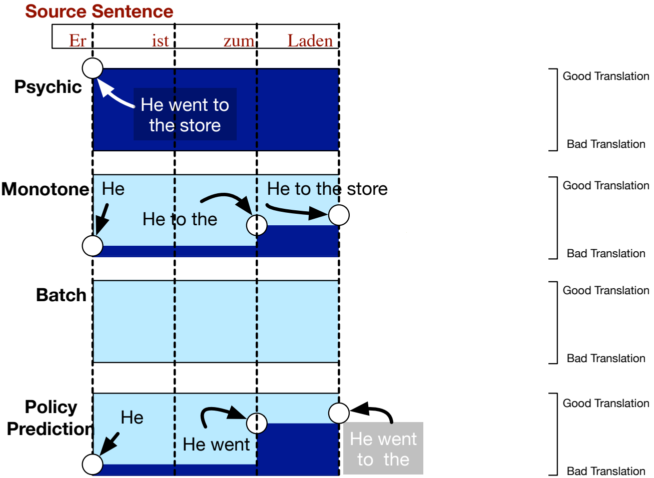
Comparing Policies



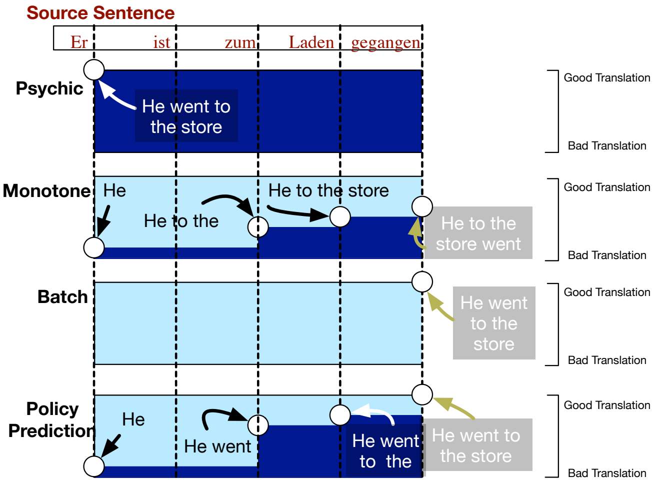
Comparing Policies



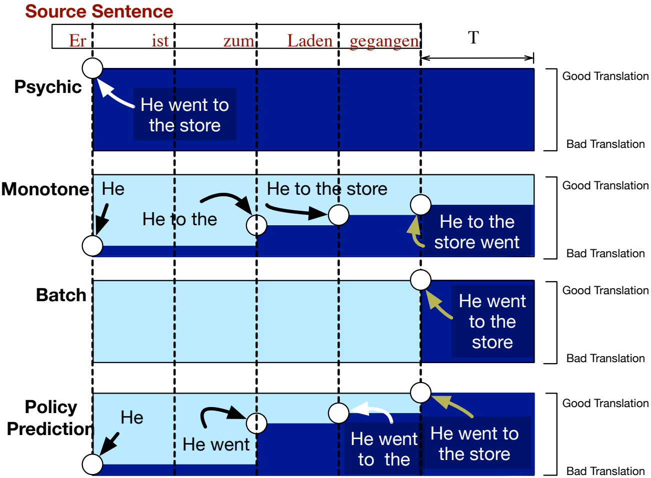
Comparing Policies



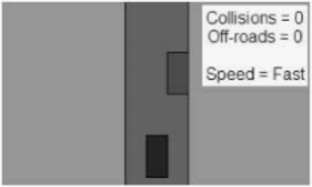
Comparing Policies



Comparing Policies

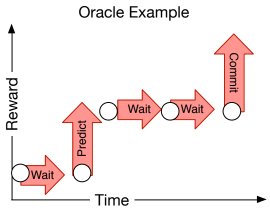


Imitation Learning

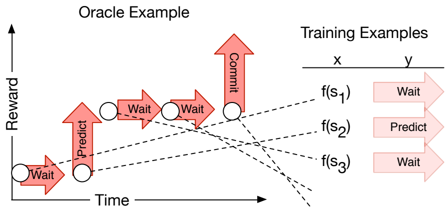


- Given all the predictions that we make (and the resulting translations) . . .
- Discover the optimal in hindsight policies
- Goal: Teach our algorithm to think on its feet
- Challenge: Represent states in a way that will generalize

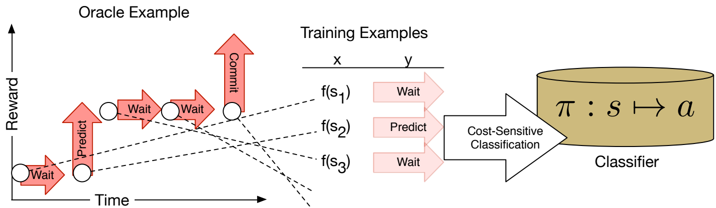
How do we find a good policy?



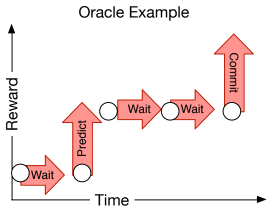
How do we find a good policy?

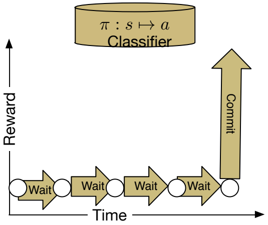


How do we find a good policy?

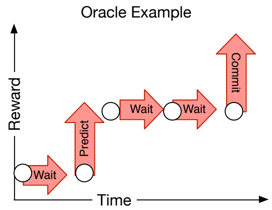


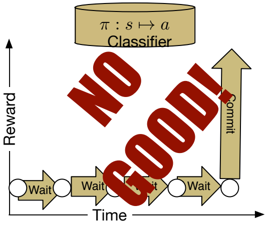
How do we find a good policy?



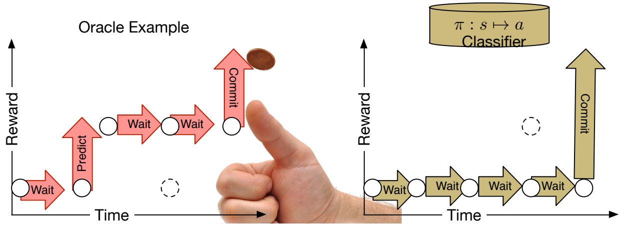


How do we find a good policy?

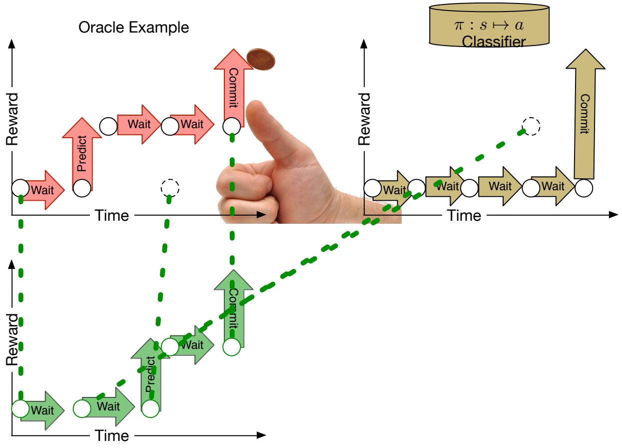




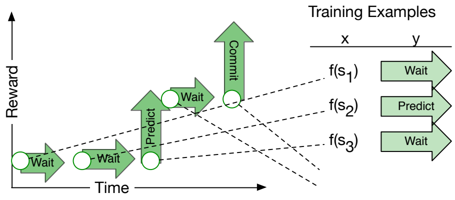
How do we find a good policy?



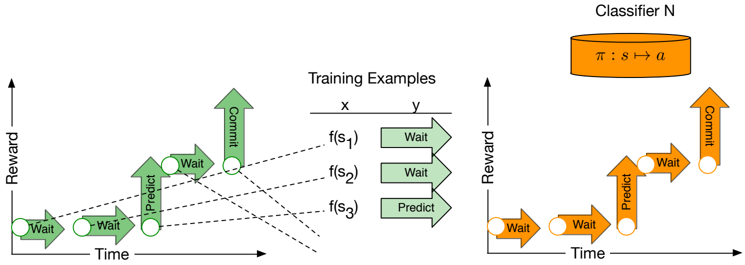
How do we find a good policy?



How do we find a good policy?

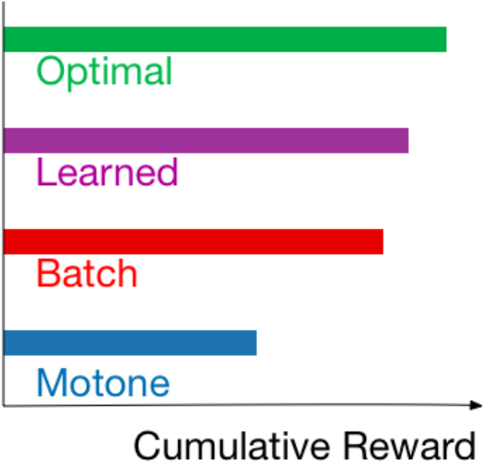


How do we find a good policy?

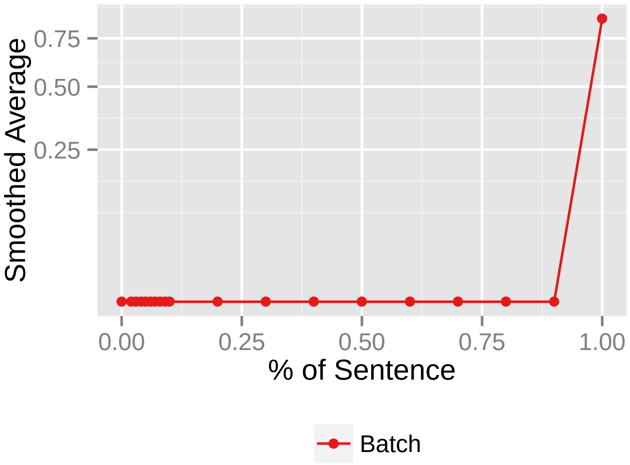


SEARN: Searching to Learn (Daumé & Marcu, 2006)

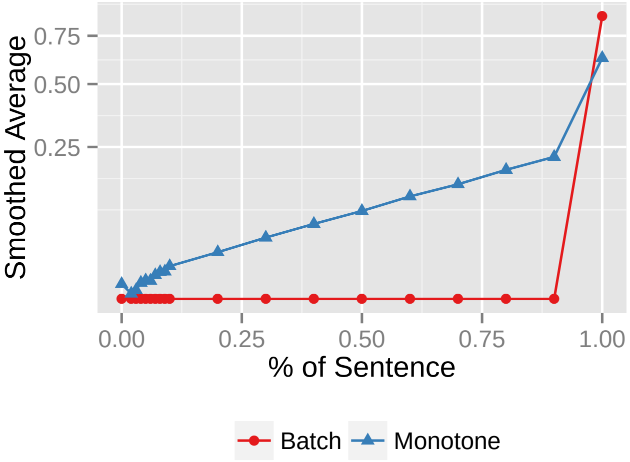
Comparing Policies



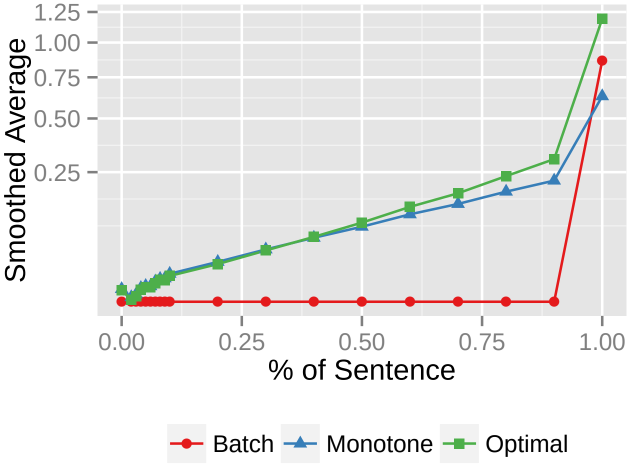
Learned Policy with Accumulated Reward



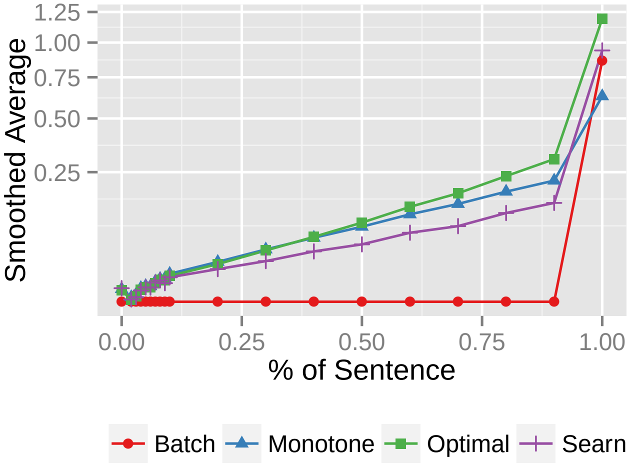
Learned Policy with Accumulated Reward



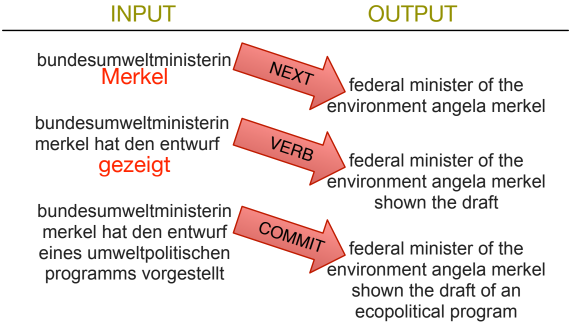
Learned Policy with Accumulated Reward



Learned Policy with Accumulated Reward



Example Sentence



What tricks do interpreters use?

Interpretese vs. Translationese: The Uniqueness of Human Strategies in Simultaneous Interpretation

He He, **Jordan Boyd-Graber**, and Hal Daumé III. *North American Association for Computational Linguistics*, 2016

- Predictions [Levy and Keller 2013, Momma et al. 2015]
- Passivization
- Segmentation [Camayd-Freixas 2011, Shimizu et al. 2013]
- Generalize [Dell and O'Seaghdha 1992, Cuetos et al. 2006]
- Summarize

Syntax-based Rewriting for Simultaneous Machine Translation

He He, Alvin Grissom II, **Jordan Boyd-Graber**, and Hal Daumé III. *Empirical Methods in Natural Language Processing*, 2015

Syntax-based Rewriting for Simultaneous Machine Translation

He He, Alvin Grissom II, **Jordan Boyd-Graber**, and Hal Daumé III. *Empirical Methods in Natural Language Processing*, 2015

	Translation			
	GD	RW	RW+GD	Gold ref
# of verbs	1971	2050	**2224**	2731


TRANSLATOR
Claudio Munoz


початок об 11.00
Олена
Anna
Ushenina
Ukraine
Olena
Boytsun
Ukraine

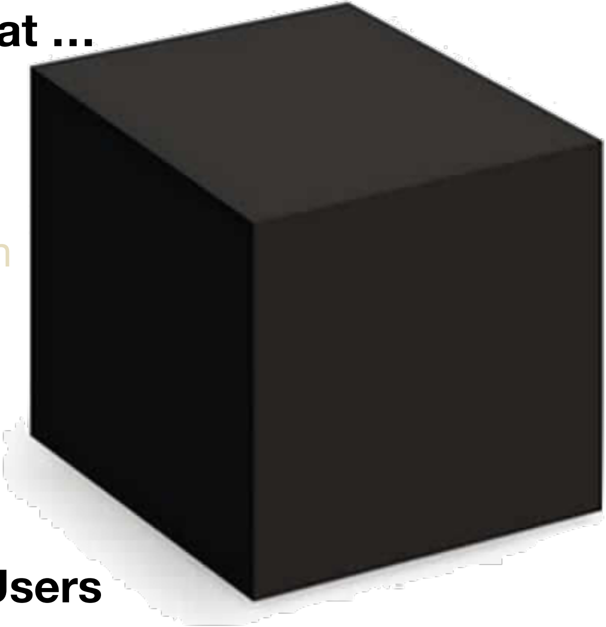
Algorithms that …

Inform

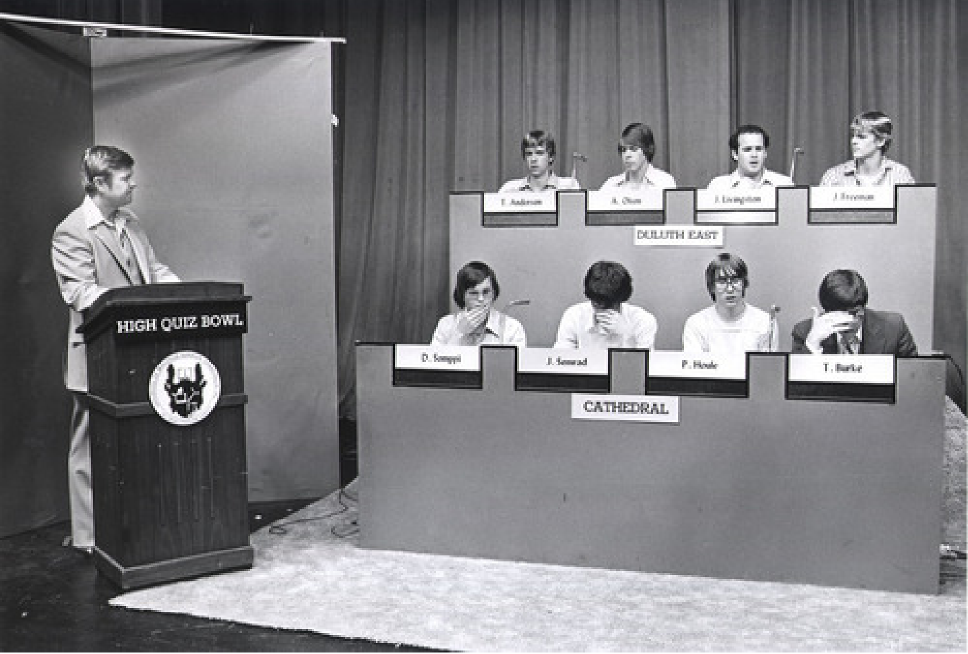
Collaborate with

Compete with

Understand



their Human Users


HIGH QUIZ BOWL
DULUTH EAST
CATHEDRAL
T. Burke

Sample Question

With Leo Szilard, he invented a doubly-eponymous

Sample Question

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of

Sample Question

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so

Sample Question

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sums repeated indices in tensor products. His name is attached to the A and B coefficients

Sample Question

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sums repeated indices in tensor products. His name is attached to the A and B coefficients for spontaneous and stimulated emission, the subject of one of his multiple groundbreaking 1905 papers. He further developed the model of statistics sent to him by

Sample Question

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sums repeated indices in tensor products. His name is attached to the A and B coefficients for spontaneous and stimulated emission, the subject of one of his multiple groundbreaking 1905 papers. He further developed the model of statistics sent to him by Bose to describe particles with integer spin. For 10 points, who is this German physicist best known for formulating the

Sample Question

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sums repeated indices in tensor products. His name is attached to the A and B coefficients for spontaneous and stimulated emission, the subject of one of his multiple groundbreaking 1905 papers. He further developed the model of statistics sent to him by Bose to describe particles with integer spin. For 10 points, who is this German physicist best known for formulating the special and general theories of relativity?

Sample Question

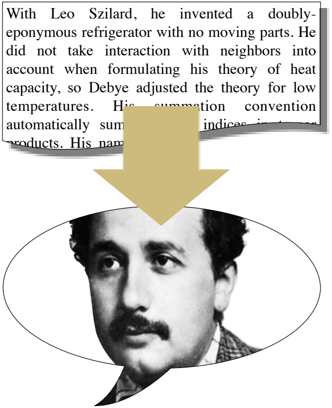
With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sums repeated indices in tensor products. His name is attached to the A and B coefficients for spontaneous and stimulated emission, the subject of one of his multiple groundbreaking 1905 papers. He further developed the model of statistics sent to him by Bose to describe particles with integer spin. For 10 points, who is this German physicist best known for formulating the special and general theories of relativity?

Albert Einstein

This is **not** Jeopardy [Ferrucci et al. 2010]

- Jeopardy: must decide to answer **once**, after complete question
- Quiz Bowl: decide after each word

How to approach this problem . . .



How to approach this problem . . .

With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sum indices products. His nam

A Neural Network for Factoid Question Answering over Paragraphs

Mohit Iyyer, **Jordan Boyd-Graber**, Leonardo Claudino, Richard Socher, and Hal Daumé III. *Empirical Methods in Natural Language Processing*, 2014

Vector Space Model

Qatar

From Wikipedia, the free encyclopedia

For other places with the same name, see Qatar (disambiguation).

Qatar (/ˈkɑːtɑr/, /ˈkɑːtər/ or /kəˈtɑr/;[6] Arabic: قطر *Qatar* [ˈqɑtˤɑr]; local
the **State of Qatar** (Arabic: دولة قطر *Dawlat Qatar*), is a sovereign Arab
the small Qatar Peninsula on the northeastern coast of the Arabian Penir
to the south, with the rest of its territory surrounded by the Persian Gulf. /
from the nearby island kingdom of Bahrain. In 2013, Qatar's total populat
and 1.5 million expatriates.[8]

Vector Space Model

Qatar

From Wikipedia, the free encyclopedia

For other places with the same name, see Qatar (disambiguation).

Qatar (/ˈkɑːtɑr/, /ˈkɑːtər/ or /kəˈtɑr/;[6] Arabic: قطر *Qatar* [ˈqɑtˤɑr]; local
the **State of Qatar** (Arabic: دولة قطر *Dawlat Qatar*), is a sovereign Arab
the small Qatar Peninsula on the northeastern coast of the Arabian Penir
to the south, with the rest of its territory surrounded by the Persian Gulf. /
from the nearby island kingdom of Bahrain. In 2013, Qatar's total populat
and 1.5 million expatriates.[8]

arabian

persian

gulf

kingdom

expatriates

Vector Space Model

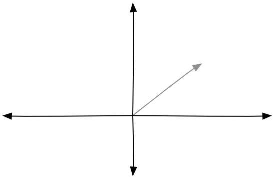
This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.

Vector Space Model

This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.

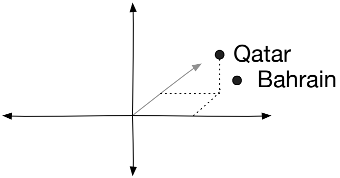
Vector Space Model

This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.



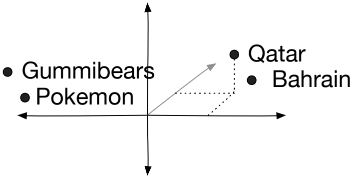
Vector Space Model

This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.



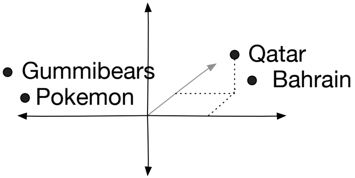
Vector Space Model

This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.



Vector Space Model

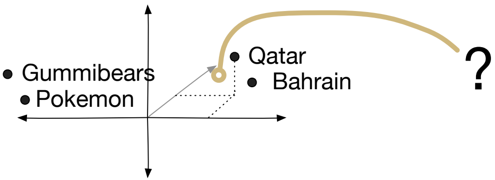
This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.



?

Vector Space Model

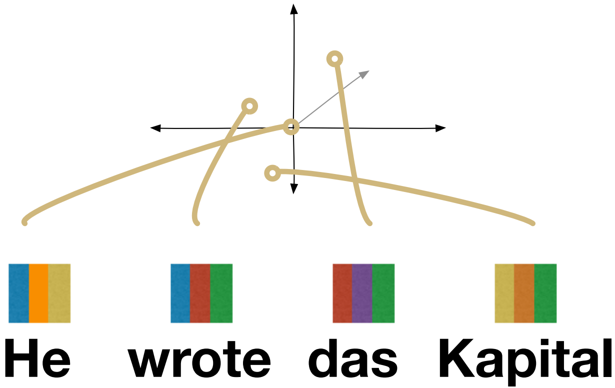
This country invested heavily in liquefied natural gas technologies, which it exports from its undersea North Dome field. This home of CENTCOM is currently led by a man who took power in a 1995 familial coup, Sheik Hamad bin Khalifa al-Thani. Wikileaks revealed that this country may have used its control over television programming as a diplomatic bargaining chip and this country pledged to use solar power to cool stadiums en route to being awarded a bid by FIFA in December 2010. For 10 points, identify this country home to Al-Jazeera which is near Bahrain and juts into the Persian Gulf.



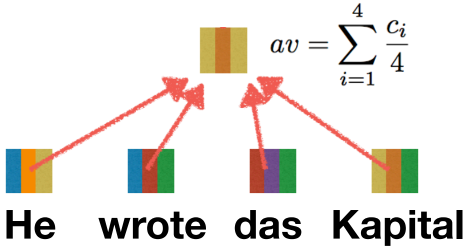
How can we do better?

- Use relationship between questions ("China" and "Taiwan")
- Use learned features and dimensions, not the words we start with

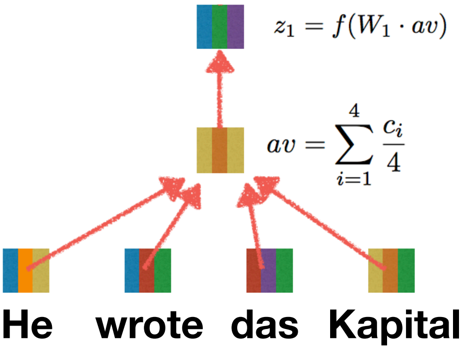
Deep Averaging Networks



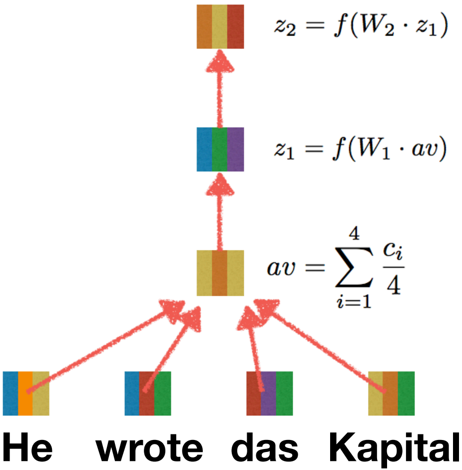
Deep Averaging Networks



Deep Averaging Networks

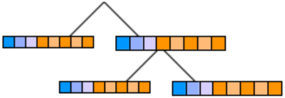


Deep Averaging Networks



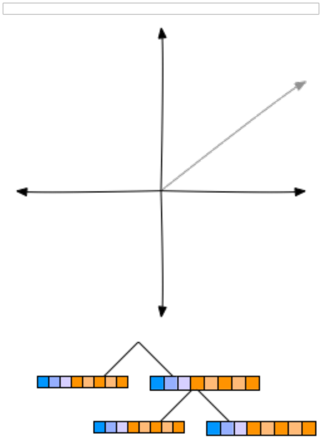
Training

- Initialize embeddings from WORD2VEC
- Randomly initialize composition matrices
- Update using WARP
  - Randomly choose an instance



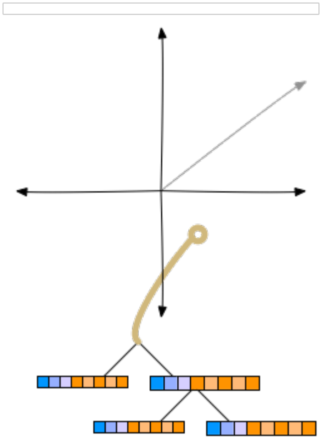
Training

- Initialize embeddings from WORD2VEC
- Randomly initialize composition matrices
- Update using WARP
  - Randomly choose an instance
  - Look where it lands



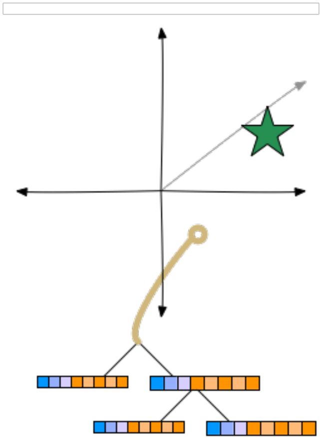
Training

- Initialize embeddings from WORD2VEC
- Randomly initialize composition matrices
- Update using WARP
  - Randomly choose an instance
  - Look where it lands



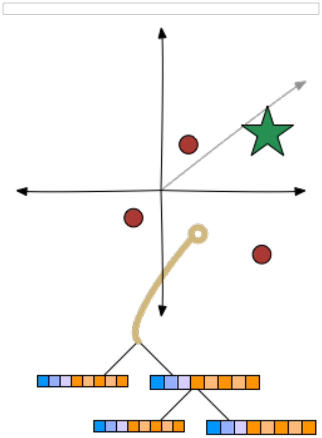
Training

- Initialize embeddings from WORD2VEC
- Randomly initialize composition matrices
- Update using WARP
  - Randomly choose an instance
  - Look where it lands
  - Has a correct answer



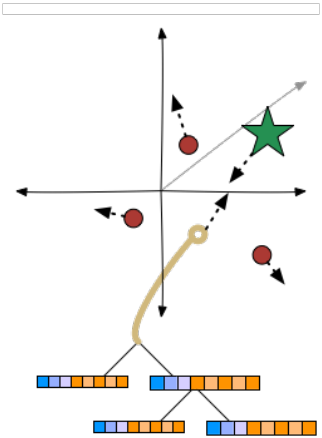
Training

- Initialize embeddings from WORD2VEC
- Randomly initialize composition matrices
- Update using WARP
  - Randomly choose an instance
  - Look where it lands
  - Has a correct answer
  - Wrong answers may be closer

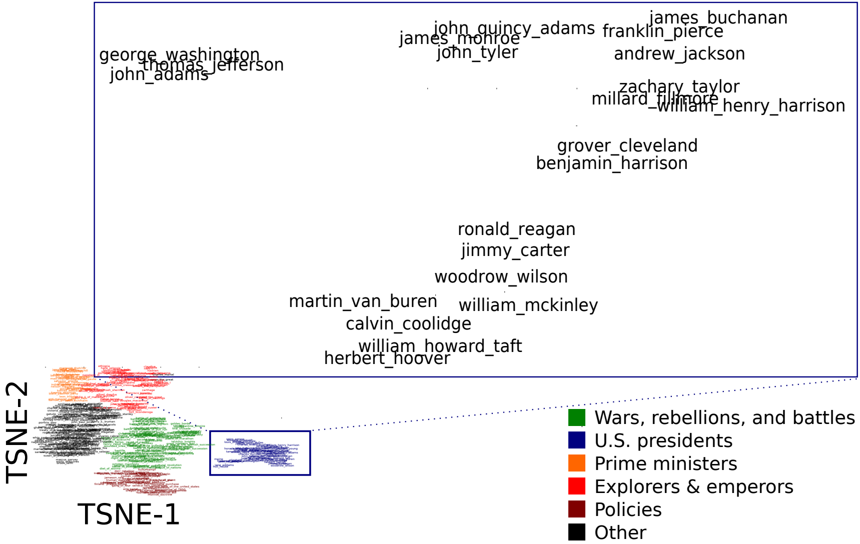


Training

- Initialize embeddings from WORD2VEC
- Randomly initialize composition matrices
- Update using WARP
  - Randomly choose an instance
  - Look where it lands
  - Has a correct answer
  - Wrong answers may be closer
  - Push away wrong answers
  - Bring correct answers closer



Embedding


george_washington
thomas_jefferson
john_adams
john_quincy_adams
james_monroe
john_tyler
james_buchanan
franklin_pierce
andrew_jackson
zachary_taylor
millard_fillmore
william_henry_harrison
grover_cleveland
benjamin_harrison
ronald_reagan
jimmy_carter
woodrow_wilson
martin_van_buren
william_mckinley
calvin_coolidge
william_howard_taft
herbert_hoover
TSNE-2
TSNE-1
Wars, rebellions, and battles
U.S. presidents
Prime ministers
Explorers & emperors
Policies
Other

How to approach this problem . . .

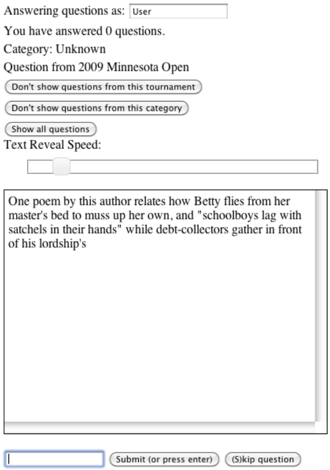
With Leo Szilard, he invented a doubly-eponymous refrigerator with no moving parts. He did not take interaction with neighbors into account when formulating his theory of heat capacity, so Debye adjusted the theory for low temperatures. His summation convention automatically sum indices products. His nam

How to approach this problem . . .

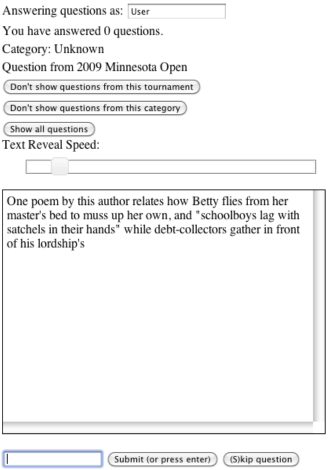
Besting the Quiz Master: Crowdsourcing Incremental Classification Games

Jordan Boyd-Graber, He He, and Hal Daumé III. *Empirical Methods in Natural Language Processing*, 2012

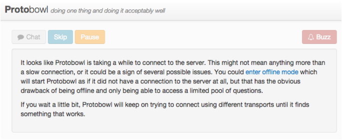
Interface


Answering questions as: User
You have answered 0 questions.
Category: Unknown
Question from 2009 Minnesota Open
Don't show questions from this tournament
Don't show questions from this category
Show all questions
Text Reveal Speed:
One poem by this author relates how Betty flies from her master's bed to muss up her own, and "schoolboys lag with satchels in their hands" while debt-collectors gather in front of his lordship's
Submit (or press enter)
(S)kip question

Interface



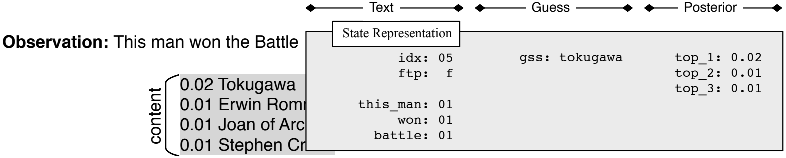
- 7000 questions: first day
- 43000 questions: two weeks
- 461 unique users
- Imitated . . .



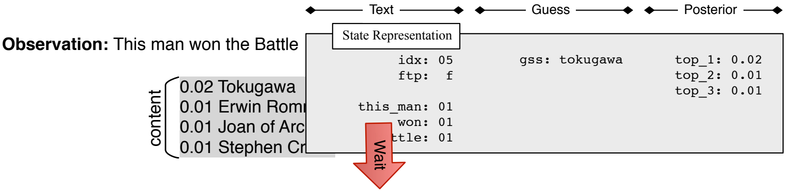
Observation: This man won the Battle

content

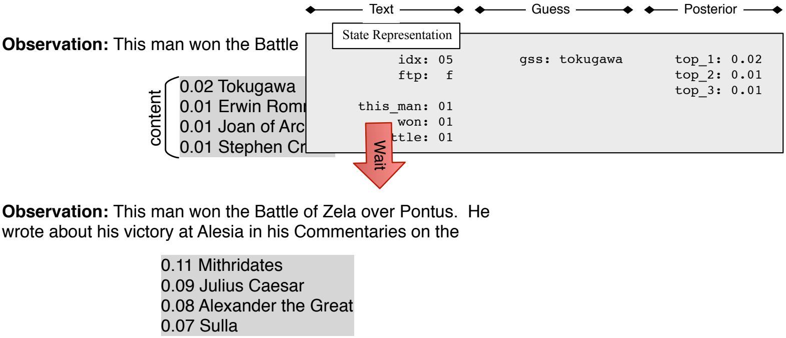
0.02 Tokugawa
0.01 Erwin Rommel
0.01 Joan of Arc
0.01 Stephen Crane

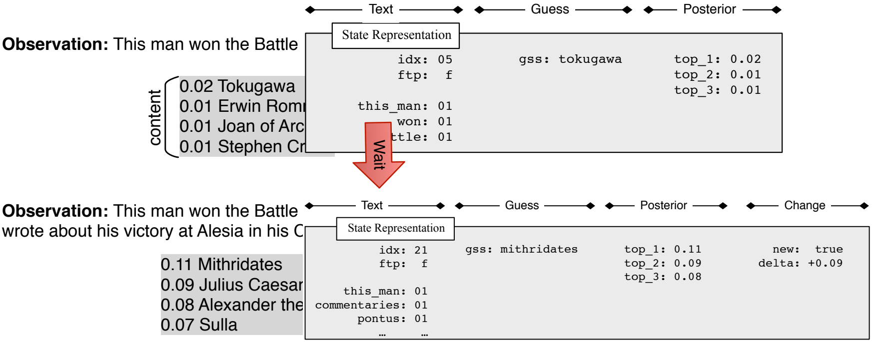


Text
Guess
Posterior
State Representation
Observation: This man won the Battle
idx: 05
ftp: f
this_man: 01
won: 01
battle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
content
0.02 Tokugawa
0.01 Erwin Romr
0.01 Joan of Arc
0.01 Stephen Cr

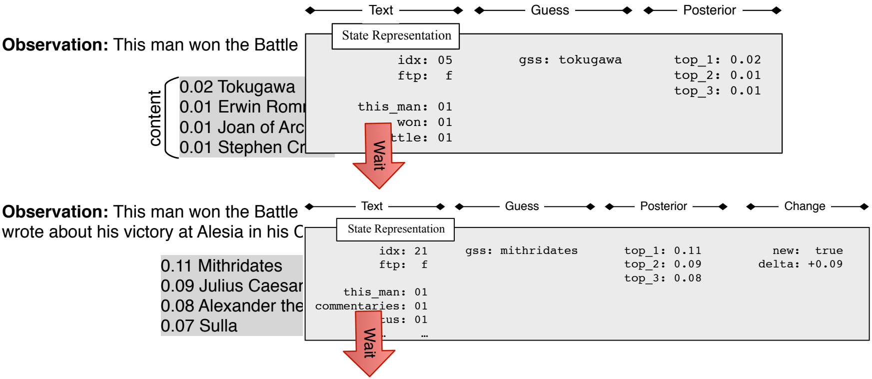


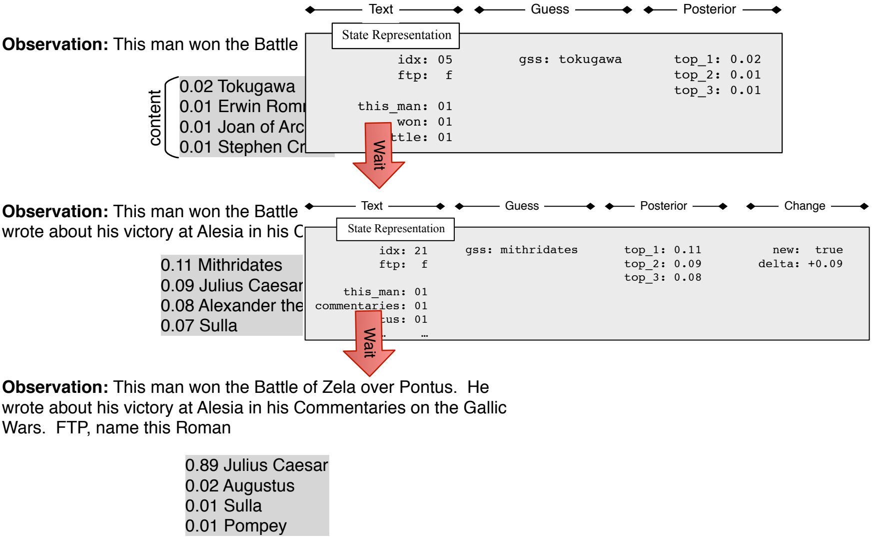
Text
Guess
Posterior
State Representation
Observation: This man won the Battle
content
0.02 Tokugawa
0.01 Erwin Romr
0.01 Joan of Arc
0.01 Stephen Cr
idx: 05
ftp: f
this_man: 01
won: 01
ttle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
Wait

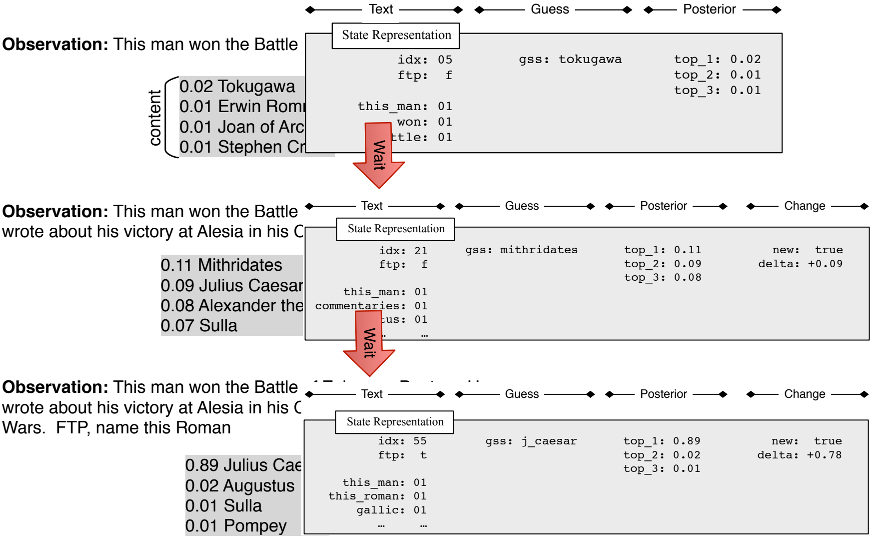

Text
Guess
Posterior
Observation: This man won the Battle
State Representation
idx: 05
ftp: f
this_man: 01
won: 01
ttle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
content
0.02 Tokugawa
0.01 Erwin Romm
0.01 Joan of Arc
0.01 Stephen Cr
Wait
Observation: This man won the Battle of Zela over Pontus. He wrote about his victory at Alesia in his Commentaries on the
0.11 Mithridates
0.09 Julius Caesar
0.08 Alexander the Great
0.07 Sulla

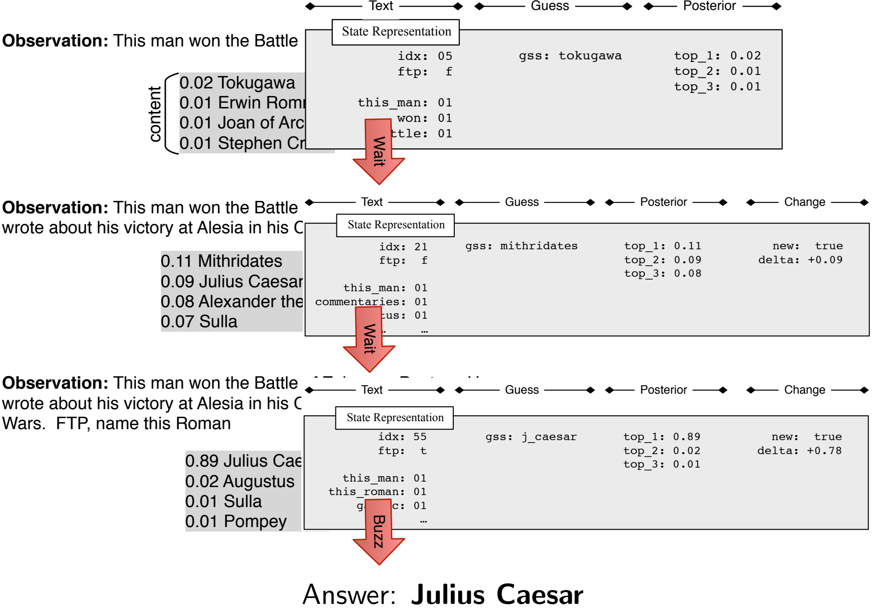


Text
Guess
Posterior
Observation: This man won the Battle
State Representation
idx: 05
ftp: f
this_man: 01
won: 01
ttle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
content
0.02 Tokugawa
0.01 Erwin Romi
0.01 Joan of Arc
0.01 Stephen Cr
Wait
Text
Guess
Posterior
Change
Observation: This man won the Battle
wrote about his victory at Alesia in his C
State Representation
idx: 21
ftp: f
this_man: 01
commentaries: 01
pontus: 01
… …
gss: mithridates
top_1: 0.11
top_2: 0.09
top_3: 0.08
new: true
delta: +0.09
0.11 Mithridates
0.09 Julius Caesar
0.08 Alexander the
0.07 Sulla

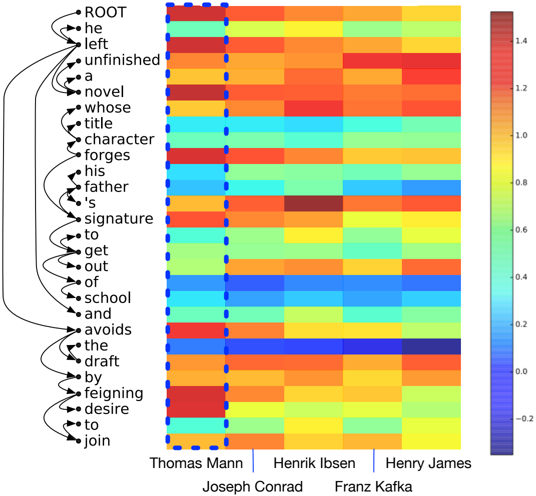

Observation: This man won the Battle
content
0.02 Tokugawa
0.01 Erwin Romm
0.01 Joan of Arc
0.01 Stephen Cr
Text
Guess
Posterior
State Representation
idx: 05
ftp: f
this_man: 01
won: 01
ttle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
Wait
Observation: This man won the Battle
wrote about his victory at Alesia in his C
0.11 Mithridates
0.09 Julius Caesar
0.08 Alexander the
0.07 Sulla
Text
Guess
Posterior
Change
State Representation
idx: 21
ftp: f
this_man: 01
commentaries: 01
us: 01
gss: mithridates
top_1: 0.11
top_2: 0.09
top_3: 0.08
new: true
delta: +0.09
Wait


Text
Guess
Posterior
State Representation
Observation: This man won the Battle
content
0.02 Tokugawa
0.01 Erwin Romi
0.01 Joan of Arc
0.01 Stephen Cr
idx: 05
ftp: f
this_man: 01
won: 01
ttle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
Wait
Text
Guess
Posterior
Change
State Representation
Observation: This man won the Battle
wrote about his victory at Alesia in his C
0.11 Mithridates
0.09 Julius Caesar
0.08 Alexander the
0.07 Sulla
idx: 21
ftp: f
this_man: 01
commentaries: 01
tus: 01
…
gss: mithridates
top_1: 0.11
top_2: 0.09
top_3: 0.08
new: true
delta: +0.09
Wait
Observation: This man won the Battle of Zela over Pontus. He wrote about his victory at Alesia in his Commentaries on the Gallic Wars. FTP, name this Roman
0.89 Julius Caesar
0.02 Augustus
0.01 Sulla
0.01 Pompey


Text
Guess
Posterior
State Representation
Observation: This man won the Battle
content
0.02 Tokugawa
0.01 Erwin Romi
0.01 Joan of Arc
0.01 Stephen Cr
idx: 05
ftp: f
this_man: 01
won: 01
ttle: 01
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
Wait
Observation: This man won the Battle
wrote about his victory at Alesia in his C
0.11 Mithridates
0.09 Julius Caesar
0.08 Alexander the
0.07 Sulla
Text
Guess
Posterior
Change
State Representation
idx: 21
ftp: f
this_man: 01
commentaries: 01
cus: 01
gss: mithridates
top_1: 0.11
top_2: 0.09
top_3: 0.08
new: true
delta: +0.09
Wait
Observation: This man won the Battle
wrote about his victory at Alesia in his C
Wars. FTP, name this Roman
0.89 Julius Cae
0.02 Augustus
0.01 Sulla
0.01 Pompey
Text
Guess
Posterior
Change
State Representation
idx: 55
ftp: t
this_man: 01
this_roman: 01
gallic: 01
gss: j_caesar
top_1: 0.89
top_2: 0.02
top_3: 0.01
new: true
delta: +0.78


Text
Guess
Posterior
State Representation
idx: 05
ftp: f
gss: tokugawa
top_1: 0.02
top_2: 0.01
top_3: 0.01
this_man: 01
won: 01
ttle: 01
Observation: This man won the Battle
content
0.02 Tokugawa
0.01 Erwin Romi
0.01 Joan of Arc
0.01 Stephen Cr
Wait
Text
Guess
Posterior
Change
Observation: This man won the Battle
wrote about his victory at Alesia in his C
State Representation
idx: 21
ftp: f
gss: mithridates
top_1: 0.11
top_2: 0.09
top_3: 0.08
new: true
delta: +0.09
this_man: 01
commentaries: 01
us: 01
0.11 Mithridates
0.09 Julius Caesar
0.08 Alexander the
0.07 Sulla
Wait
Observation: This man won the Battle
wrote about his victory at Alesia in his C
Wars. FTP, name this Roman
Text
Guess
Posterior
Change
State Representation
idx: 55
ftp: t
gss: j_caesar
top_1: 0.89
top_2: 0.02
top_3: 0.01
new: true
delta: +0.78
this_man: 01
this_roman: 01
c: 01
0.89 Julius Cae
0.02 Augustus
0.01 Sulla
0.01 Pompey
Buzz
Answer: Julius Caesar

Examining vectors



Experiment 1

Colby Burnett: $375,000

Ben Ingram: $427,534

Alex Jacobs: $151,802

Kristin Sausville: $95,201

Experiment 1

Colby Burnett: $375,000

Ben Ingram: $427,534

Alex Jacobs: $151,802

Kristin Sausville: $95,201

End result: 200-200 tie!

23. October 2015, Seattle

Humans 345-145

Humans 190-155

Where we have problems

Out of Date

Although he won the California primary in 2000, he distanced himself from fellow reform presidential candidate Pat Buchanan by comparing him to Attila the Hun. After being called a jackass, he prompted Lindsey Graham to destroy his phone by giving out his number during a speech. The slogan (*) Make America Great Again has been used by this politician, who claimed he didn't like people who were captured as a slight to John McCain and kicked off his 2016 presidential bid with some inflammatory remarks about Mexicans. For 10 points, name this Republican candidate and real estate mogul.

Where we have problems

Out of Date

Although he won the California primary in 2000, he distanced himself from fellow reform presidential candidate Pat Buchanan by comparing him to Attila the Hun. After being called a jackass, he prompted Lindsey Graham to destroy his phone by giving out his number during a speech. The slogan (*) Make America Great Again has been used by this politician, who claimed he didn't like people who were captured as a slight to John McCain and kicked off his 2016 presidential bid with some inflammatory remarks about Mexicans. For 10 points, name this Republican candidate and real estate mogul.

Chris Christie?

Where we have problems

Out of Touch

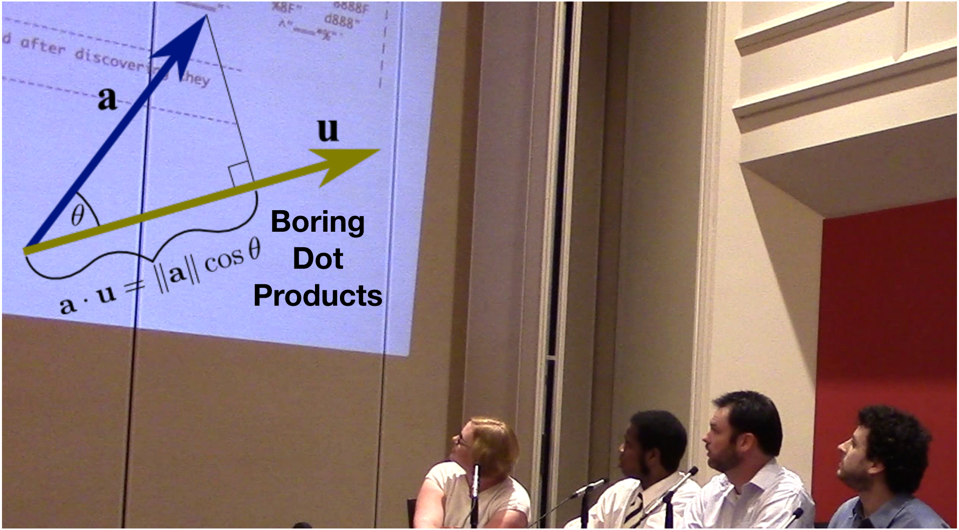
This singer recently cancelled the Great Escape Tour, and, in one song, she claims that she will be "Eating crumpets with the sailors / On acres without the neighbors." She collaborated with Jennifer (*) Hudson on the song "Trouble," which was issued in her album update Reclassified. This artist of "Change Your Life" was inspired by scenes from the movie Clueless to make the music video for a song in which she collaborated with Charli XCX. For 10 points, name this Australian rapper whose album The New Classic contained "Fancy."

Where we have problems

Out of Touch

This singer recently cancelled the Great Escape Tour, and, in one song, she claims that she will be "Eating crumpets with the sailors / On acres without the neighbors." She collaborated with Jennifer (*) Hudson on the song "Trouble," which was issued in her album update Reclassified. This artist of "Change Your Life" was inspired by scenes from the movie Clueless to make the music video for a song in which she collaborated with Charli XCX. For 10 points, name this Australian rapper whose album The New Classic contained "Fancy."

Bruce Springsteen?


a
u
θ
a · u = ||a|| cos θ
Boring
Dot
Products

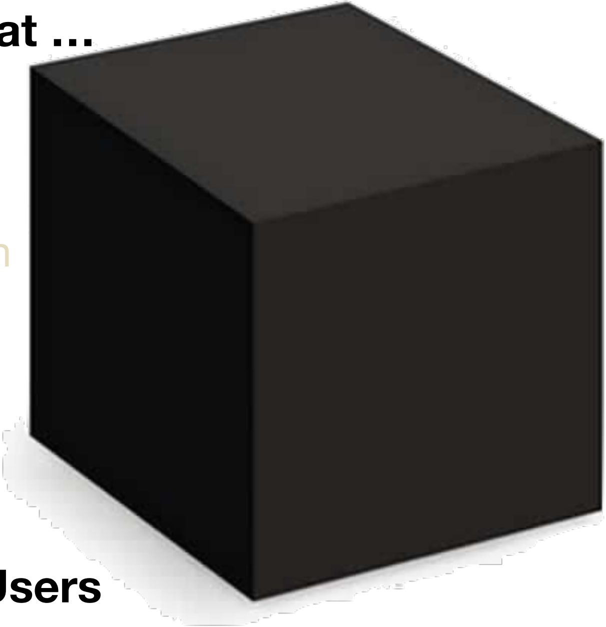
Algorithms that …

Inform

Collaborate with

Compete with

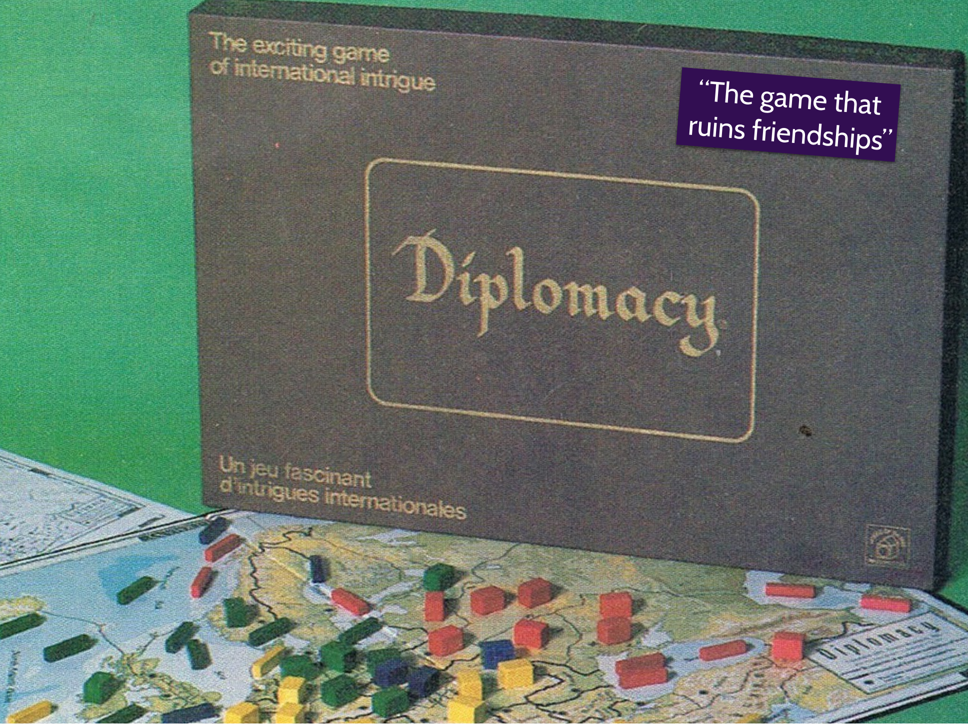
Understand



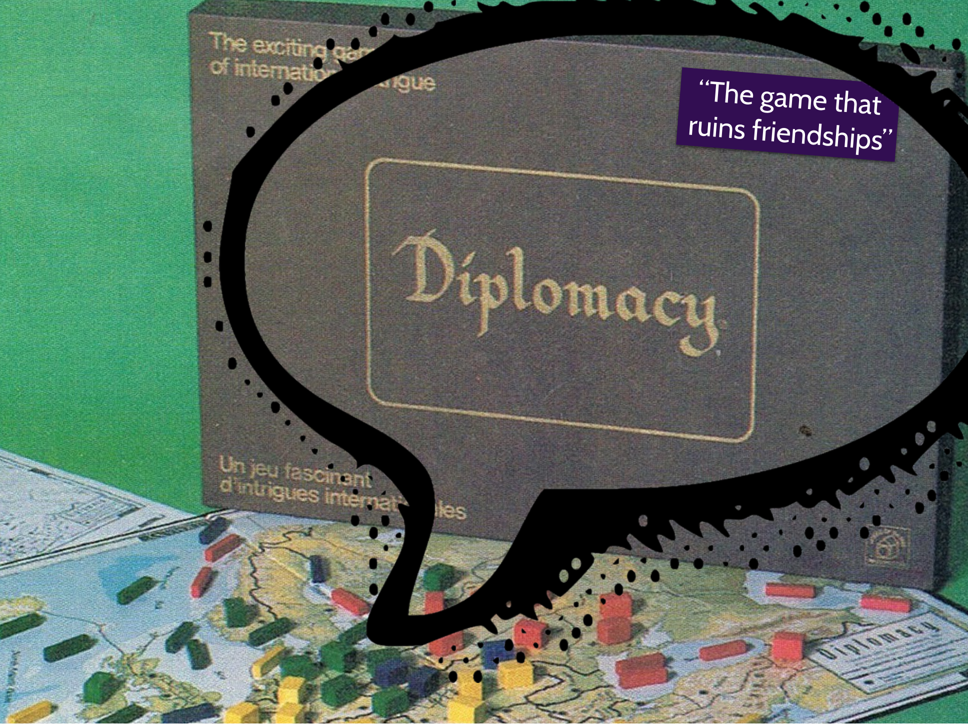
their Human Users

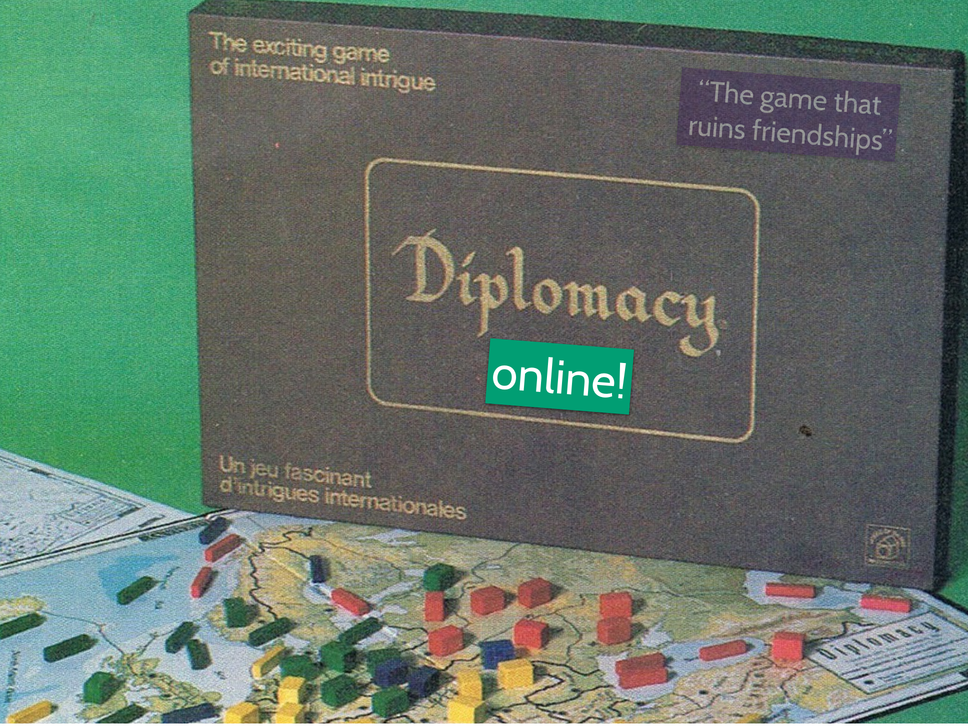
Linguistic Harbingers of Betrayal: A Case Study on an Online Strategy Game

Vlad Niculae, Srijan Kumar, Jordan Boyd-Graber, and Cristian Danescu-Niculescu-Mizil. *Association for Computational Linguistics*, 2015


The exciting game
of international intrigue
Diplomacy
Un jeu fascinant
d'intrigues internationales
"The game that
ruins friendships"


The exciting game
of international intrigue
“The game that
uins friendships”
Un jeu fascinant
Diplomacy


“The game that ruins friendships”
Diplomacy
Un jeu fascinant


The exciting game
of international intrigue
“The game that
ruins friendships”
Diplomacy
online!
Un jeu fascinant
d'intrigues internationales


The exciting game of international intrigue
“The game that ruins friendships”
Diplomacy
online!
249 games
~6 months/game
145k messages
diplom.org; usak.asciiking.com

Iceland
North Atlantic
NAt
NWY
Sweden
BOT
NC
StP
Siberia
NTH
Denmark
Helgoland Bight
Holland
Kiel
Livonia
LVN
Ruhr
Belgium
English Channel
Ukraine
Mid Atlantic
Munich
Tyrolia
Budapest
Rumania
Black Sea
EC
Portugal
NC
Spain
SC
GoL
Tuscany
Roma
West Mediterranean
Tyrrhenian Sea
TYN
Greece
Ankara
Armenia
Smyrna
Syria
Diplomacy
by Allan B. Calhamer
Copyright 1999, Avalon Hill
Map by J. Fatula, III


Berlin
Ukraine
Bohemia
Galicia
Vienna
Budapest
Trieste
Rumania
Serbia
EC
Bulgaria

Kiel
Berlin
Ruhr
Munich
Bohemia
Galicia
Ukraine
Vienna
Switz.
Budapest
Venezia
Trieste
Tuscany
Rumania
Serbia
EC
Bulgaria


Kiel
Berlin
Ruhr
Munich
Bohemia
Galicia
Ukraine
Vienna
Switz.
Budapest
Venezia
Tuscany
Trieste
Rumania
Serbia
EC
Bulgaria


Kiel
Berlin
Ruhr
Galicia
Switz.
Budapest
Venezia
Tuscany
Trieste
Rumania
Serbia
EC
Bulgaria


help?
Kiel
Berlin
Ruhr
Munich
Ukraine
Galicia
Budapest
Switz.
Venezia
Tuscany
Trieste
Rumania
Serbia
Bulgaria
EC

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

E

...

stabs !


NOW STAND BACK,
I GOTTA PRACTICE MY STABBIN'

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

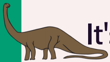
E

...

stabs !

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.



It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

...

E

 stabs  !

F

What I would like you to do is keep Turkey busy and somehow get Russia and Turkey to engage. Meanwhile we need to take VIE, suggest you support me in there.

It's a sensible plan. I'll support you as requested. Please be sure to simultaneously attack SWE.

F

...

E

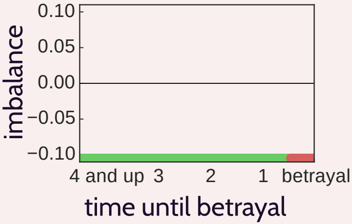
stabs !

Not really sure what to say, except that I regret you did what you did.


Curse your sudden but inevitable betrayal!

(Im)balance Over Time

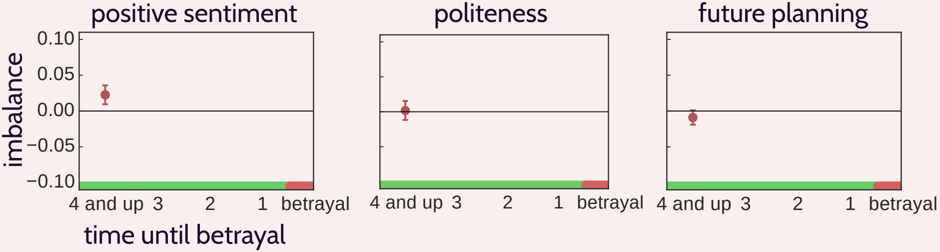
Imbalance: *f*(betrayer) - *f*(victim). Looking only at betrayals.



(Error bars show standard error.)

(Im)balance Over Time

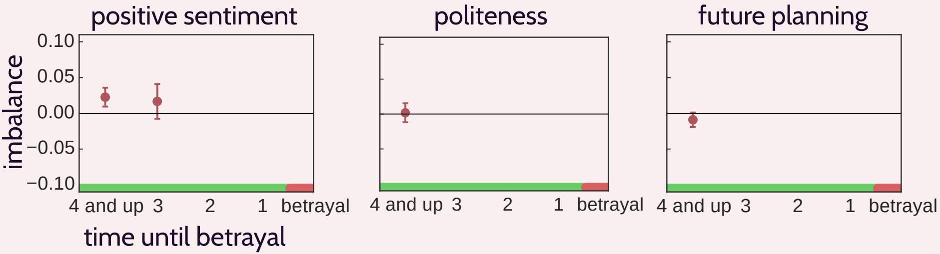
Imbalance: *f(*betrayer*)* - *f(*victim*)*



(Error bars show standard error.)

(Im)balance Over Time

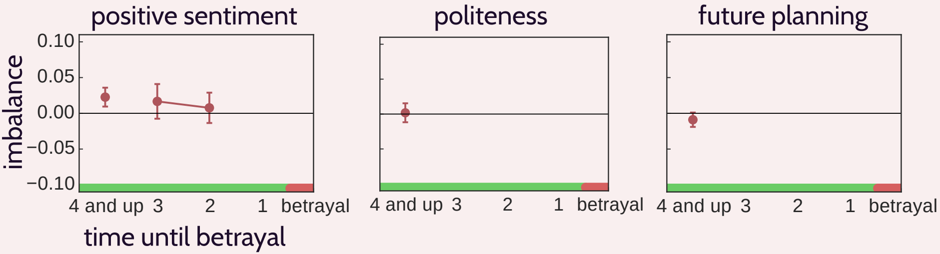
Imbalance: *f*(betrayer) - *f*(victim)



(Error bars show standard error.)

(Im)balance Over Time

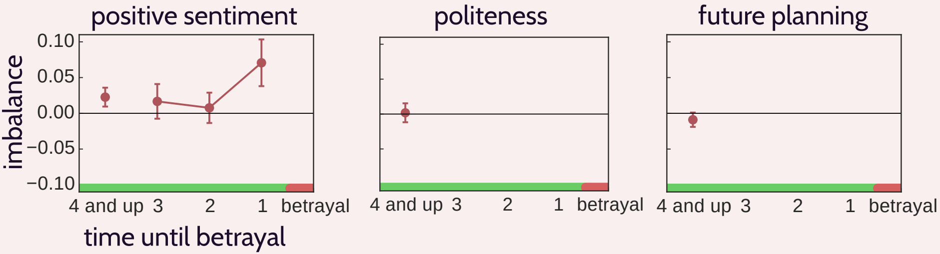
Imbalance: *f*(betrayer) - *f*(victim)



(Error bars show standard error.)

(Im)balance Over Time

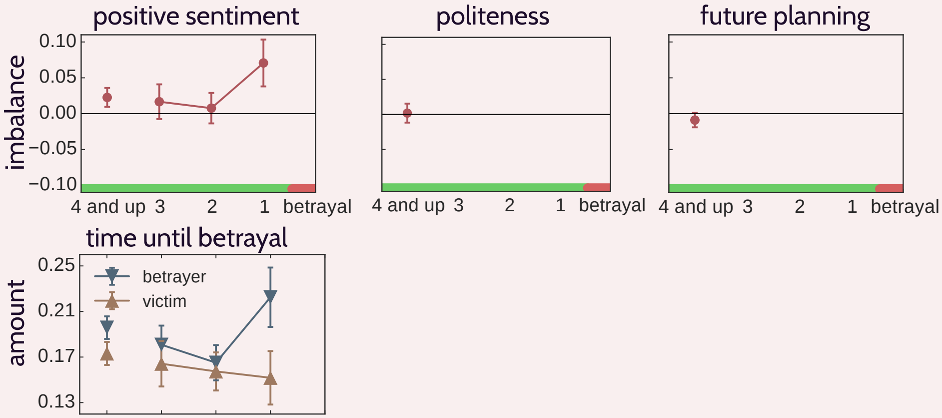
Imbalance: *f*(betrayer) - *f*(victim)



(Error bars show standard error.)

(Im)balance Over Time

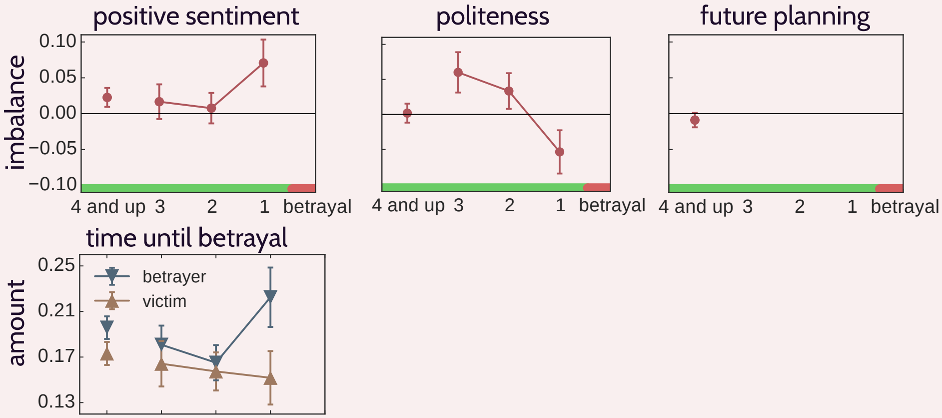
Imbalance: *f*(betrayer) - *f*(victim)



(Error bars show standard error.)

(Im)balance Over Time

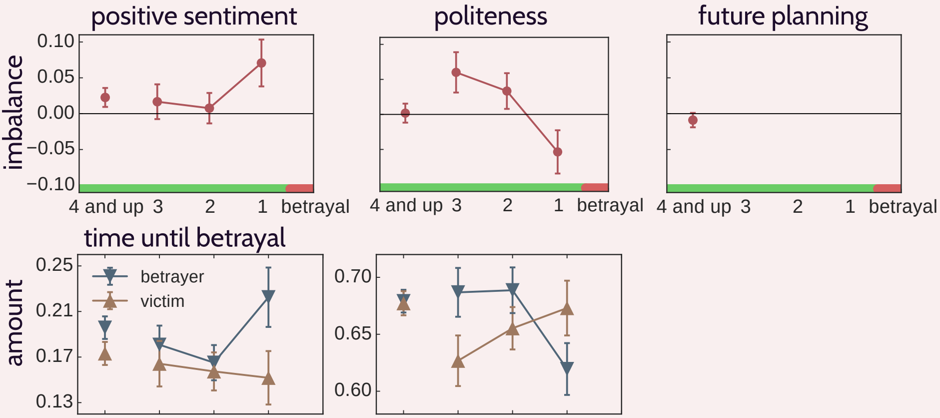
Imbalance: *f*(betrayer) - *f*(victim)



(Error bars show standard error.)

(Im)balance Over Time

Imbalance: *f*(betrayer) - *f*(victim)

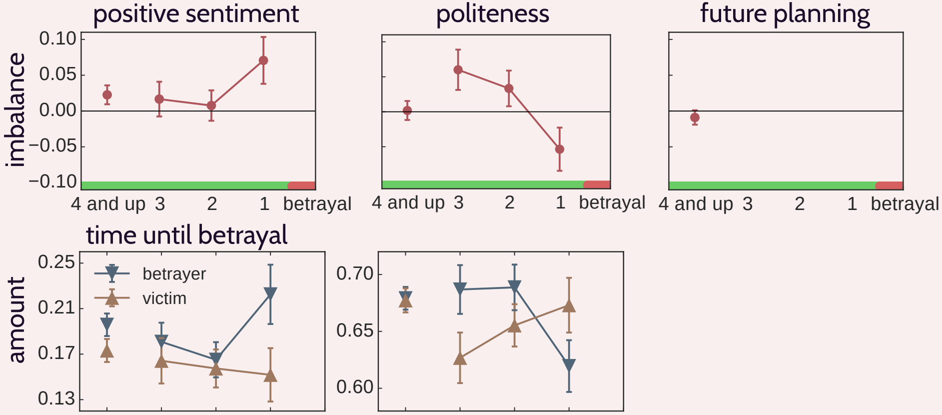


(Error bars show standard error.)

(Im)balance Over Time

Imbalance: *f*(betrayer) - *f*(victim)

Demand-Withdraw pattern pre-divorce.
(Gottman & Levenson, 2000)

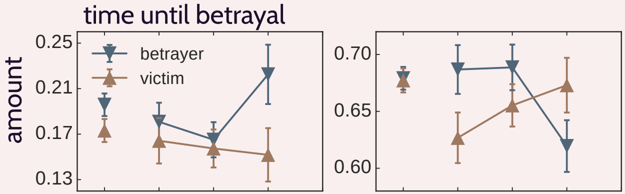


(Error bars show standard error.)

(Im)balance Over Time

Demand-Withdraw pattern pre-divorce.
(Gottman & Levenson, 2000)

Imbalance: *f(*betrayer*)* - *f(*victim*)*

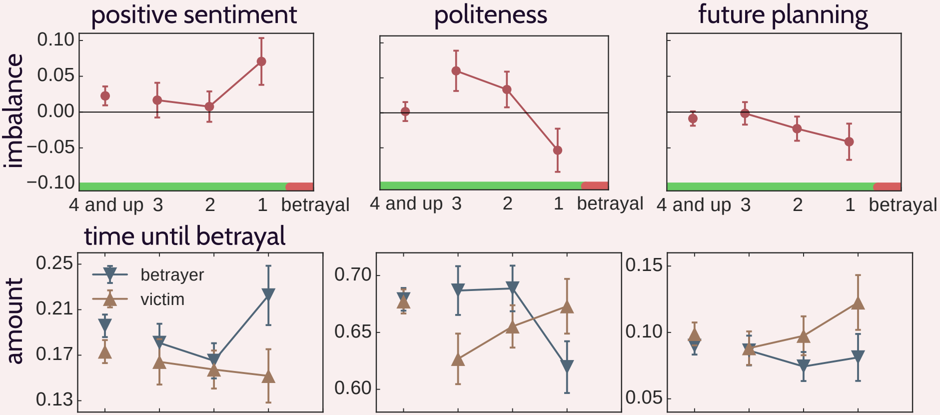


(Error bars show standard error.)

(Im)balance Over Time

Imbalance: *f*(betrayer) - *f*(victim)

Demand-Withdraw pattern pre-divorce.
(Gottman & Levenson, 2000)



(Error bars show standard error.)

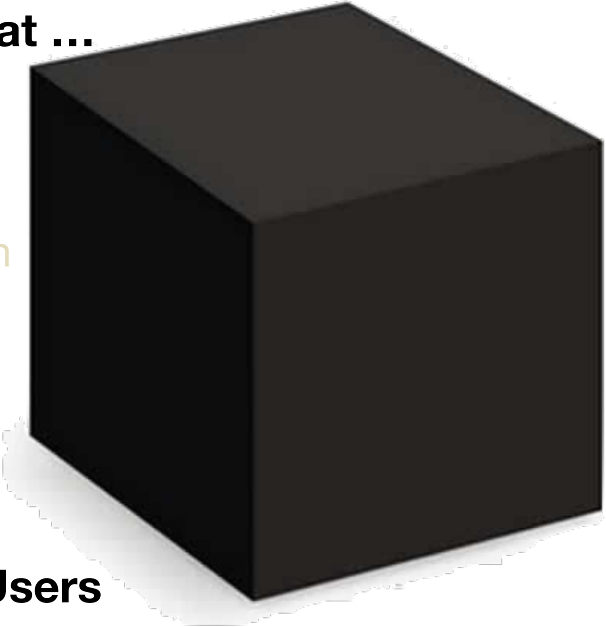
Algorithms that …

Inform

Collaborate with

Compete with

Understand



their Human Users

Tea Party in the House: A Hierarchical Ideal Point Topic Model and Its Application to Republican Legislators in the 112th Congress

Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, and Kristina Miler. *Association for Computational Linguistics*, 2015

Evaluation: Tea Party in the House

The Tea Party

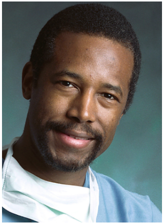
- American political movement for freedom, small government, lower tax
- Disrupting Republican Party and recent elections
- Organizations:
  - Institutional: Tea Party Caucus
  - Other: Tea Party Express, Tea Party Patriots, Freedom Works
- "**Conventional views of ideology as a single–dimensional, left-right spectrum experience great difficulty in understanding or explaining the Tea Party.**"

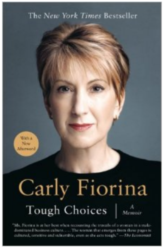
[Carmines and D'Amico 2015, ARPS]

Goal

- Explain Tea Partiers in terms of issues and votes
- Identify Tea Partiers from their rhetoric

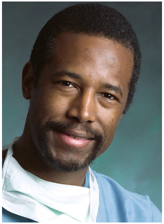
Not everyone has a voting record

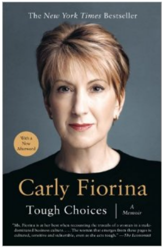




- Ideal points estimated based on voting record
- Not all candidates have a voting record
  - Governors
  - Entertainers
  - CEOs

Not everyone has a voting record





- Ideal points estimated based on voting record
- Not all candidates have a voting record
  - Governors
  - Entertainers
  - CEOs
- But all politicians—by definition—talk

Let's use whatever data we have

A single model that uses:

- Bill text
- Votes
- Commentary

to map political actors to the same continuous space.

Let's use whatever data we have

A single model that uses:

- Bill text
- Votes
- Commentary

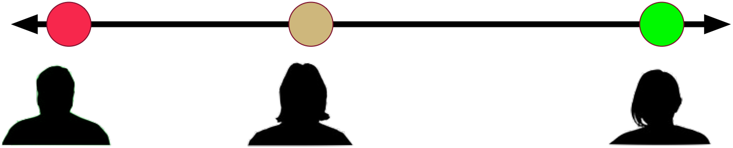
to map political actors to the same continuous space. This work: congressional floor speeches

Hierarchical Ideal Point Topic Model: Intuition

What are your thoughts on the issue of **immigration**?

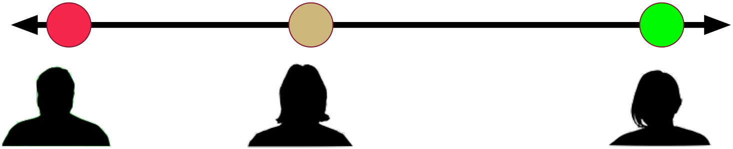
Hierarchical Ideal Point Topic Model

Hierarchical Ideal Point Topic Model



Hierarchical Ideal Point Topic Model

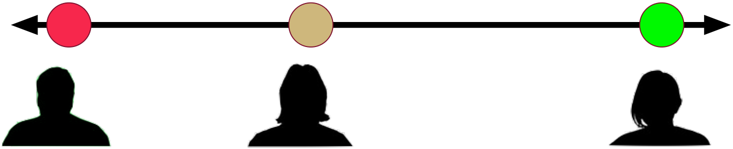
Issue: Healthcare



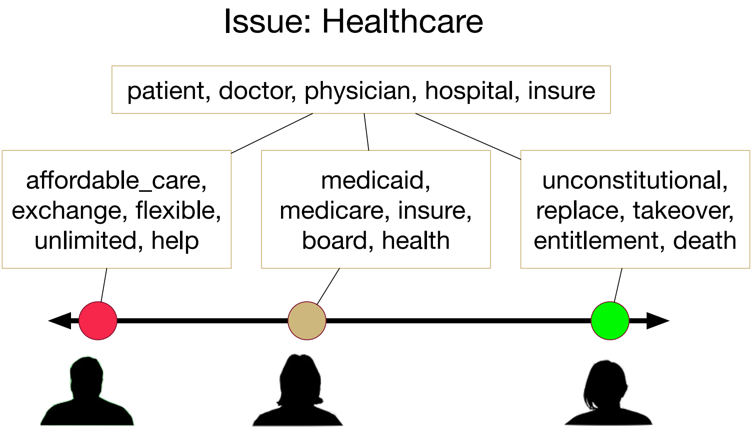
Hierarchical Ideal Point Topic Model

Issue: Healthcare

patient, doctor, physician, hospital, insure



Hierarchical Ideal Point Topic Model



Tea Party Caucus Membership Prediction

Experiment setup

- Task: Binary classification of whether a legislator is a member of the Tea Party Caucus
- Evaluation metric: AUC-ROC
- Classifier: SVM^{light}
- Five-fold stratified cross-validation

Tea Party Caucus Membership Prediction

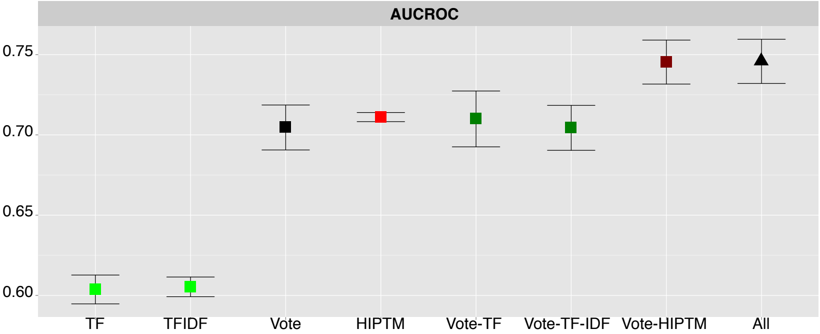
Experiment setup

- Task: Binary classification of whether a legislator is a member of the Tea Party Caucus
- Evaluation metric: AUC-ROC
- Classifier: SVM^{light}
- Five-fold stratified cross-validation

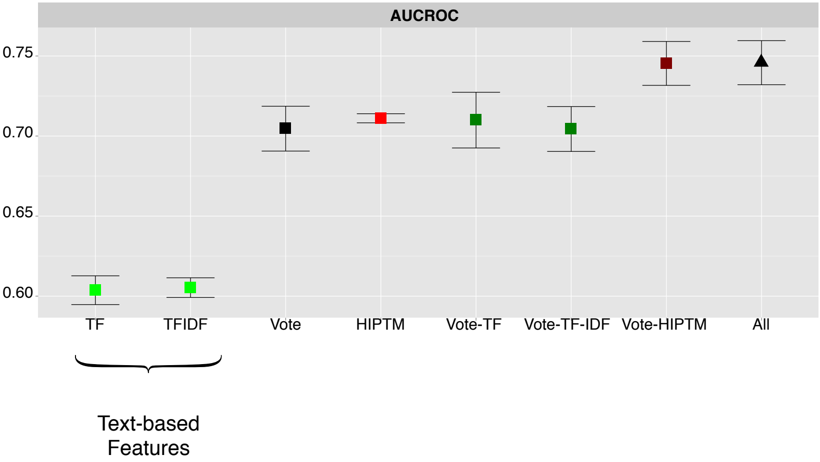
Features

- Text-based features: normalized term frequency (**TF**) and **TF-IDF**
- **Vote**: binary features
- **HIPTM**: features extracted from our model including
  - K-dim ideal point $u_{a,k}$ estimated from both votes and text
  - K-dim ideal point estimated from text only $\boldsymbol{\eta}_k^T \hat{\psi}_{a,k}$
  - B probabilities estimating a's votes $\Phi(x_b \sum_{k=1}^{K} \vartheta_{b,k} u_{a,k} + y_b)$

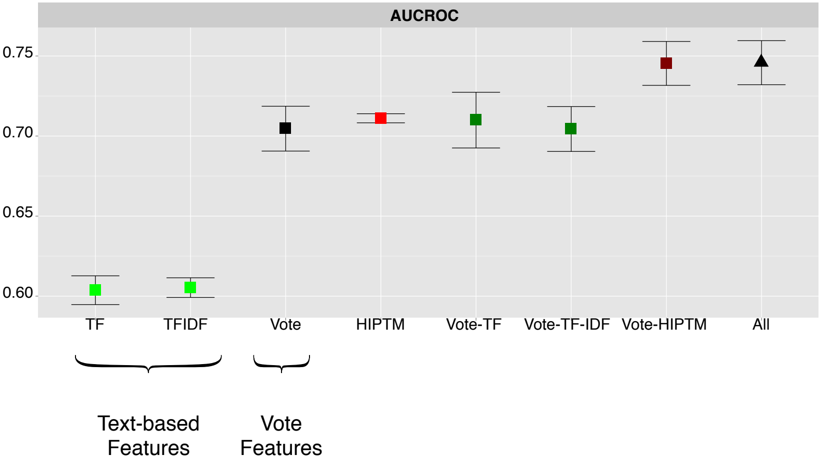
Tea Party Caucus Membership Prediction: Votes & Text



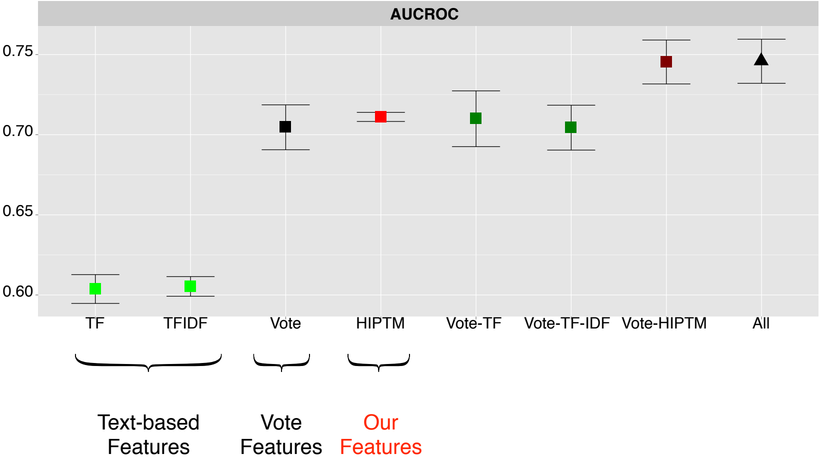
Tea Party Caucus Membership Prediction: Votes & Text



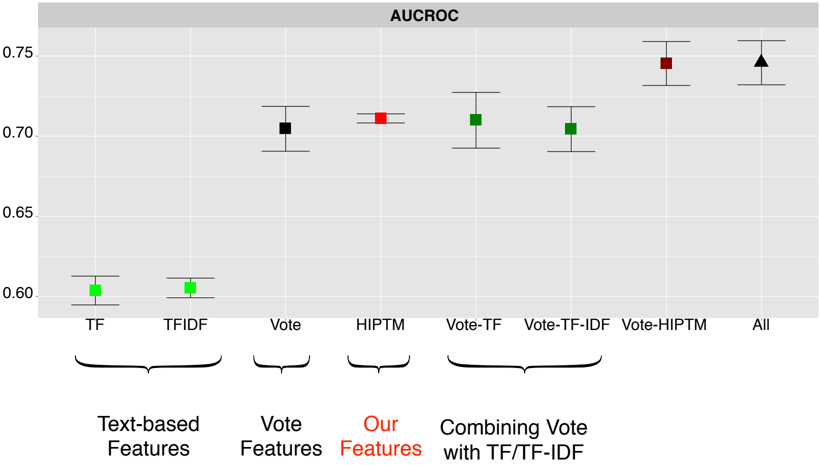
Tea Party Caucus Membership Prediction: Votes & Text



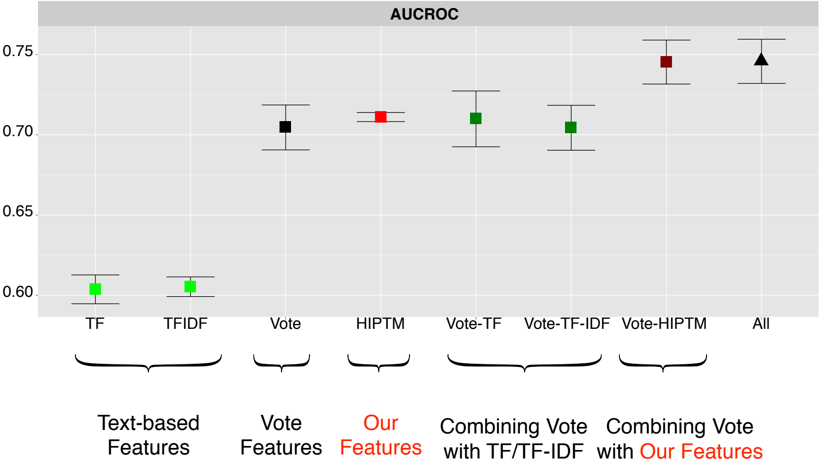
Tea Party Caucus Membership Prediction: Votes & Text



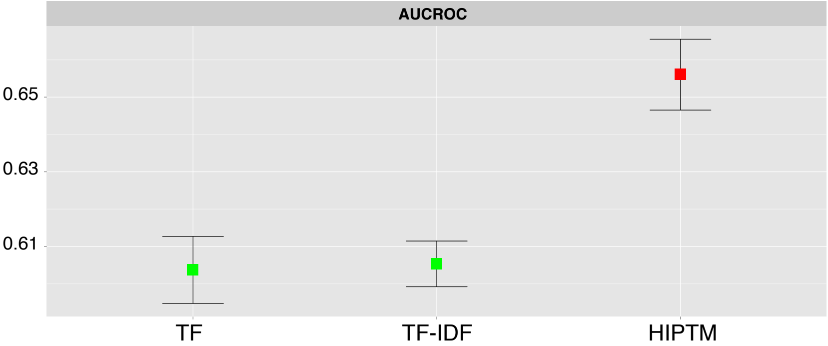
Tea Party Caucus Membership Prediction: Votes & Text



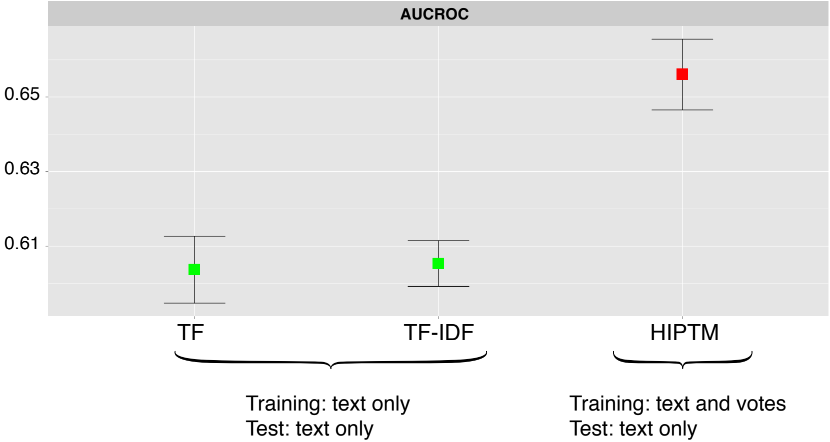
Tea Party Caucus Membership Prediction: Votes & Text



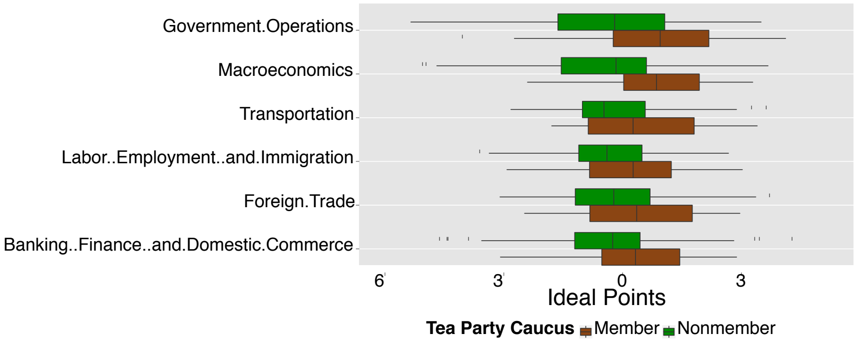
Tea Party Caucus Membership Prediction: Text Only



Tea Party Caucus Membership Prediction: Text Only

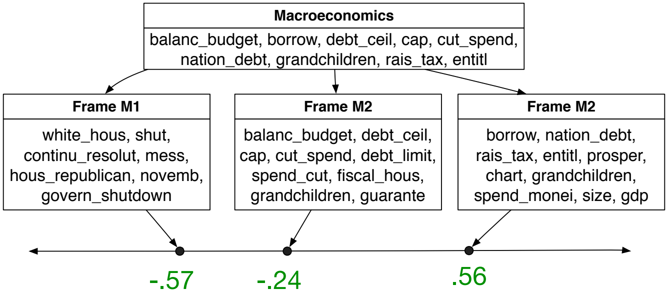


Multi-dimensional Ideal Points



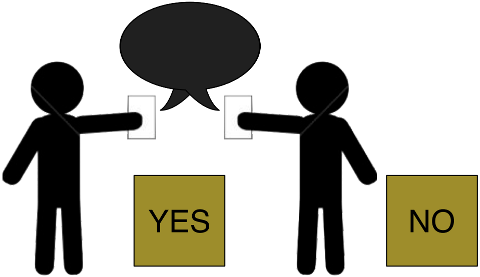
Most highly polarized dimensions are about government spending

Framing Macroeconomics



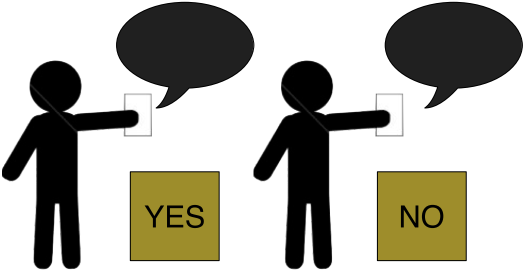
Polarization

Distribution of Issue Frames	Ideal Point Distributions	Not	Polarized
	Not	**Civil Rights, Minority Issues, Civil Liberties**	Banking and Finance; Transportation
	Polarized	Health; Public Lands and Water Management	Macroeconomics; Government Operations



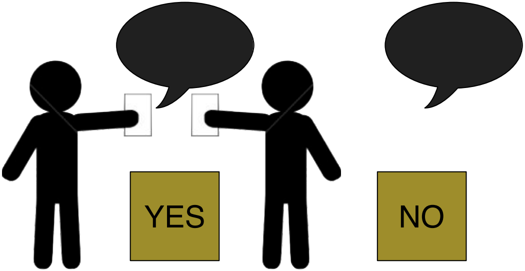
Polarization

Ideal Point Distributions		Not	Polarized
Distribution of Issue Frames	Not	Civil Rights, Minority Issues, Civil Liberties	Banking and Finance Transportation
	Polarized	Health; Public Lands and Water Management	**Macroeconomics; Government Operations**



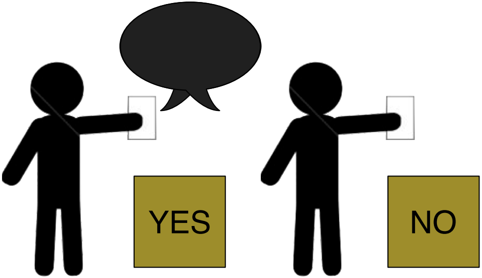
Polarization

Ideal Point Distributions		Not	Polarized
Distribution of Issue Frames	Not	Civil Rights, Minority Issues, Civil Liberties	Banking and Finance Transportation
	Polarized	**Health; Public Lands and Water Management**	Macroeconomics; Government Operations



Polarization

Ideal Point Distributions		Not	Polarized
Distribution of Issue Frames	Not	Civil Rights, Minority Issues, Civil Liberties	**Banking and Finance Transportation**
	Polarized	Health; Public Lands and Water Management	Macroeconomics; Government Operations



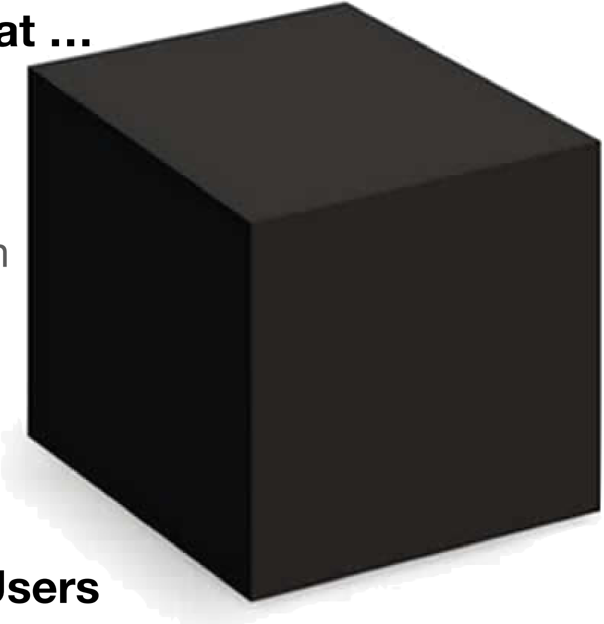
Algorithms that …

Inform

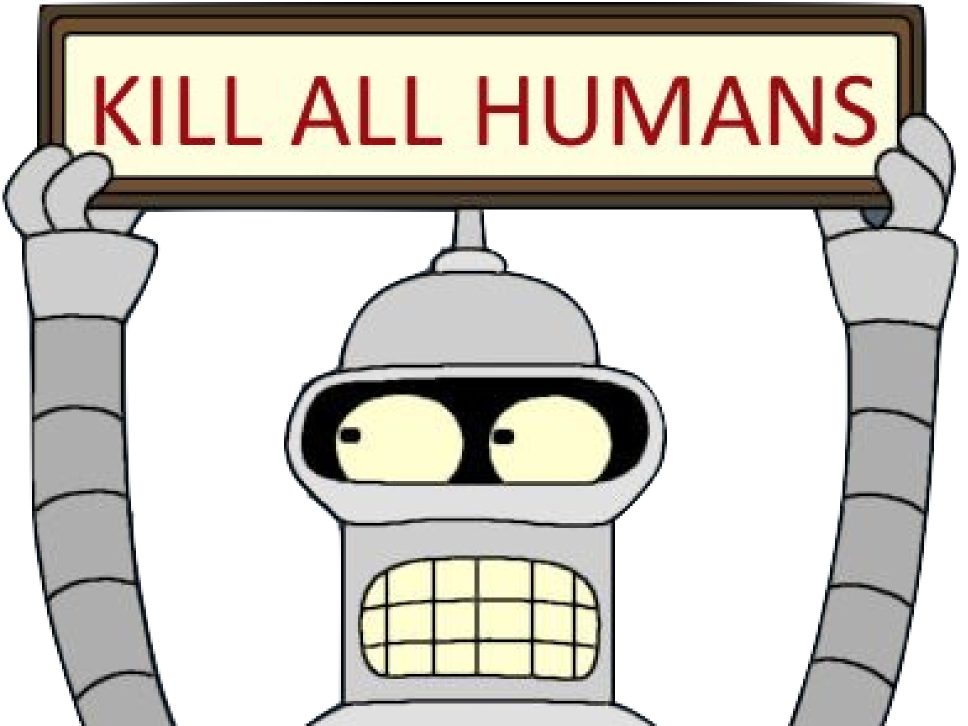
Collaborate with

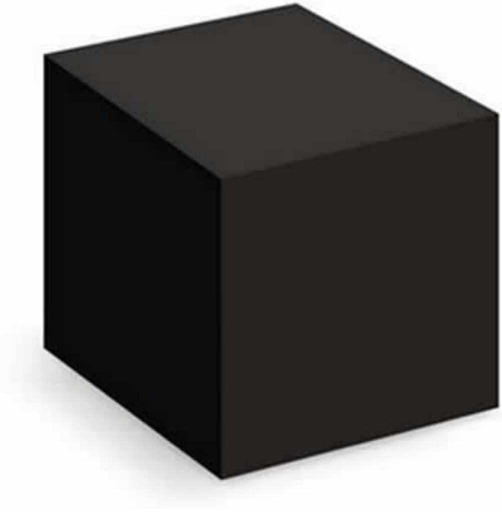
Compete with

Understand



their Human Users


KILL ALL HUMANS



We need ML that understands our gratitude and our fears

Thanks

Collaborators

NAQT, Hal Daumé III (UMD), Philip Resnik (UMD), Cristian Danescu-Niculescu-Mizil (Cornell), Leah Findlater (UMD), Kevin Seppi (BYU), Eric Ringger (BYU)

Funders

Supporters

References

David M. Blei, Andrew Ng, and Michael Jordan.
2003.
Latent Dirichlet allocation.
Journal of Machine Learning Research, 3.

Erik Camayd-Freixas.
2011.
Cognitive theory of simultaneous interpreting and training.
In *Proceedings of the 52nd Conference of the American Translators Association*.

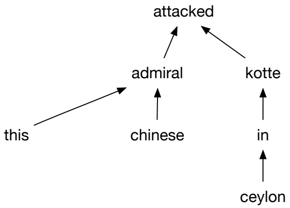
Edward G Carmines and Nicholas J D'Amico.
2015.
The new look in political ideology research.
Annual Review of Political Science, 18(4).

F. Cuetos, B. Alvarez B, M. González-Nosti, A. Méot, and P. Bonin.
2006.
Determinants of lexical access in speech production: role of word frequency and age of acquisition.
Mem Cognit, 34.

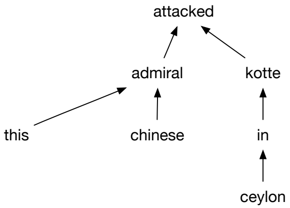
Hal Daumé III.
2004.
Notes on CG and LM-BFGS optimization of logistic regression.
Paper available at `http://pub.hal3.name/~daume04cg-bfgs`, implementation available at `http://hal3.name/megam/`.

G.S. Dell and P.G. O'Seaghdha.
1992.
Stages of lexical access in language production.

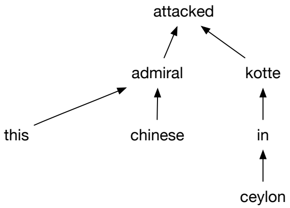
Using Compositionality

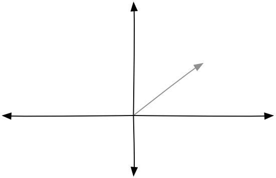


Using Compositionality

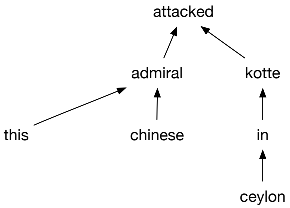


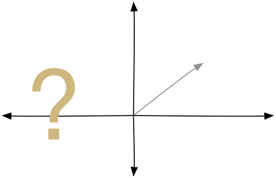
Using Compositionality



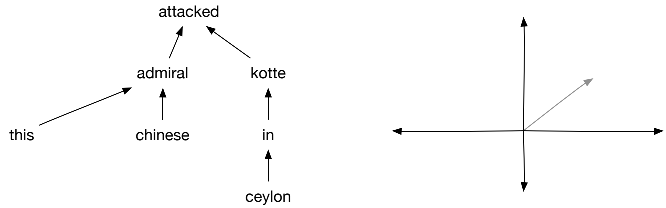


Using Compositionality

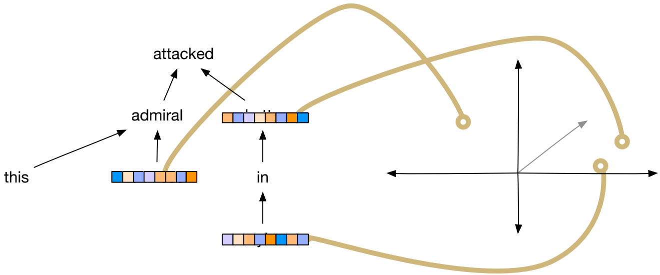




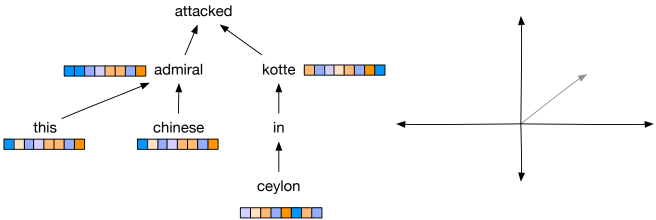
Using Compositionality



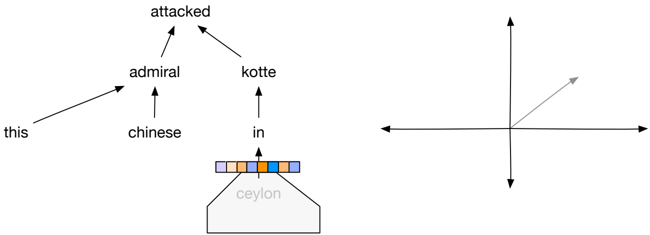
Using Compositionality



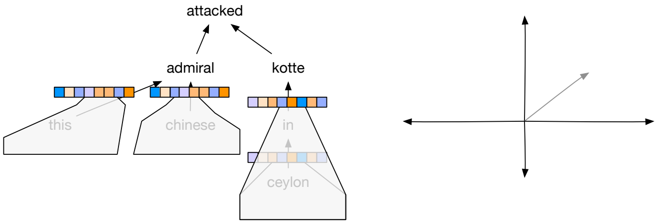
Using Compositionality



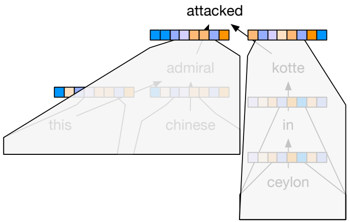
Using Compositionality

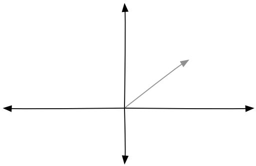


Using Compositionality



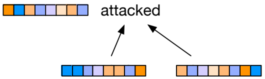
Using Compositionality

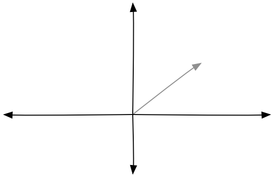




Using Compositionality

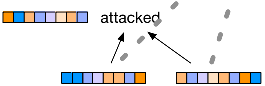
$$f(W_v \cdot x_w + b + \sum_{k \in K(n)} W_{R(n,k)} \cdot h_k) \;=$$

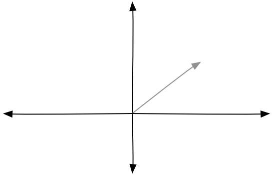




Using Compositionality

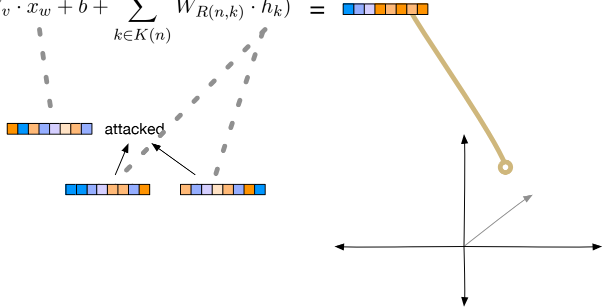
$$f(W_v \cdot x_w + b + \sum_{k \in K(n)} W_{R(n,k)} \cdot h_k) =$$





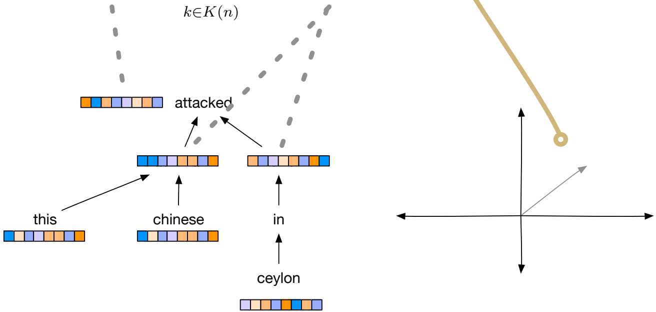
Using Compositionality

$$f(W_v \cdot x_w + b + \sum_{k \in K(n)} W_{R(n,k)} \cdot h_k) =$$



Using Compositionality

$$f(W_v \cdot x_w + b + \sum_{k \in K(n)} W_{R(n,k)} \cdot h_k) =$$



Learning which Features are Useful

- Use how humans these data as a prior for supervised maxent model [Daumé III 2004]
- Prior for label a and feature f is a function of the number of buzzes b and tf-idf [Salton 1968]

$$\left[\alpha \mathbb{I}\left[b(a,f) > 0\right] + \beta b(a,f) + \gamma\right] \text{tf-idf}(a,f). \qquad (2)$$

  - α, β, and $\gamma = 0$: naïve zero prior
  - α and $\beta = 0$: linear transformation of the mean
  - α and $\gamma = 0$: number of buzzes times tf-idf value of the features

Learning which Features are Useful

- Use how humans these data as a prior for supervised maxent model [Daumé III 2004]
- Prior for label a and feature f is a function of the number of buzzes b and tf-idf [Salton 1968]

$$\left[\alpha \mathbb{I}\left[b(a,f) > 0\right] + \beta b(a,f) + \gamma\right] \text{tf-idf}(a,f). \tag{2}$$

  - α, β, and $\gamma = 0$: naïve zero prior
  - α and $\beta = 0$: linear transformation of the mean
  - α and $\gamma = 0$: number of buzzes times tf-idf value of the features

Learning which Features are Useful

- Use how humans these data as a prior for supervised maxent model [Daumé III 2004]
- Prior for label a and feature f is a function of the number of buzzes b and tf-idf [Salton 1968]

$$\left[\alpha \mathbb{I}\left[b(a,f) > 0\right] + \beta b(a,f) + \gamma\right] \text{tf-idf}(a,f). \tag{2}$$

  - α, β, and $\gamma = 0$: naïve zero prior
  - α and $\beta = 0$: linear transformation of the mean
  - α and $\gamma = 0$: number of buzzes times tf-idf value of the features

Learning which Features are Useful

- Use how humans these data as a prior for supervised maxent model [Daumé III 2004]
- Prior for label a and feature f is a function of the number of buzzes b and tf-idf [Salton 1968]

$$\left[\alpha \mathbb{1}\left[b(a,f) > 0\right] + \beta b(a,f) + \gamma\right] \text{tf-idf}(a,f). \tag{2}$$

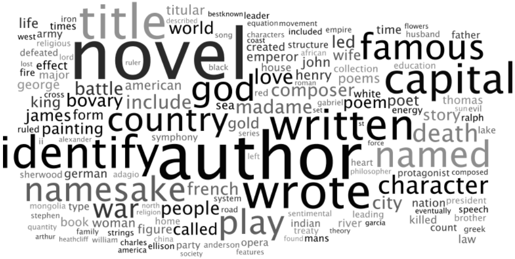
  - α, β, and $\gamma = 0$: naïve zero prior
  - α and $\beta = 0$: linear transformation of the mean
  - α and $\gamma = 0$: number of buzzes times tf-idf value of the features

Using buzzes as a prior

$$\left[\alpha \mathbb{I}\left[b(a,f) > 0\right] + \beta b(a,f) + \gamma\right] \text{tf-idf}(a,f).$$

Answers	Weighting	α	β	γ	Error[1]
	zero	-	-	-	0.22
	tf-idf	-	-	8.3	0.08
100	buzz-binary	10.7	-	-	**0.06**
	buzz-linear	-	1.1	-	0.10
	buzz-tier	-	1.6	0.5	0.07

[1] Buzz and tf-idf computed on training data; grid search on dev data; error on test data

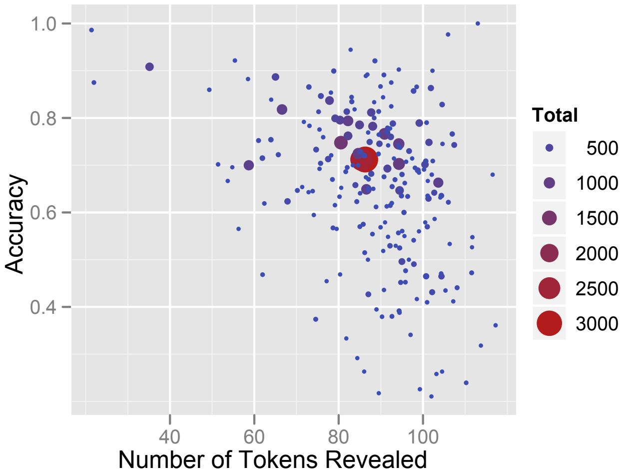


(a) Buzzes over all Questions

(b) Wuthering Heights Question Text

(c) Buzzes on Wuthering Heights

Accuracy vs. Speed



How we could translate a sentence

Observation

1. Mit dem Zug

How we could translate a sentence

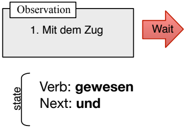
Observation

1. Mit dem Zug

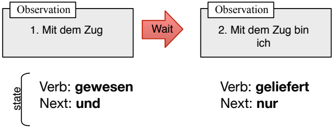
state

Verb: **gewesen**
Next: **und**

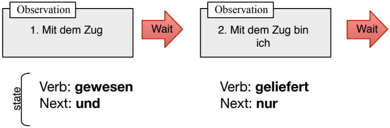
How we could translate a sentence



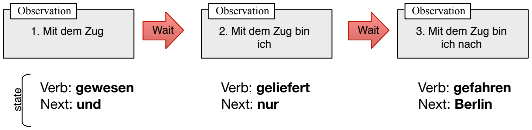
How we could translate a sentence



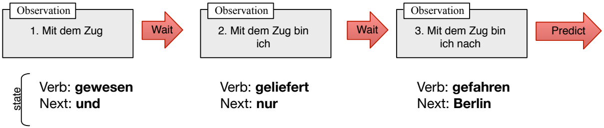
How we could translate a sentence



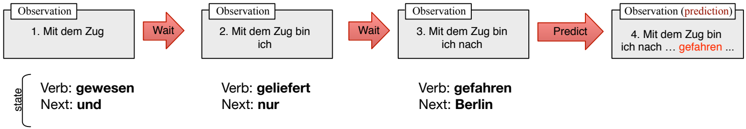
How we could translate a sentence



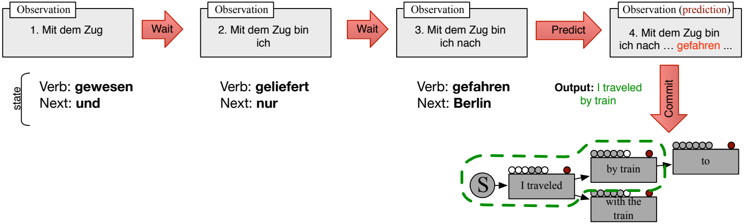
How we could translate a sentence



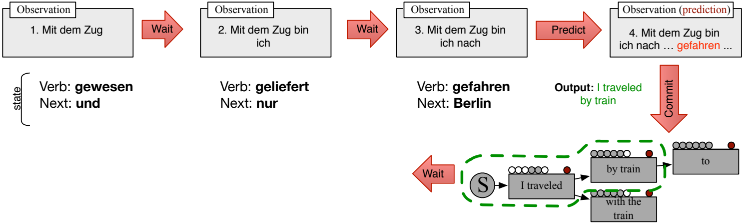
How we could translate a sentence



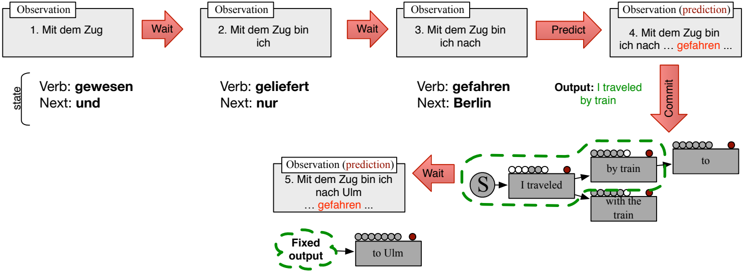
How we could translate a sentence



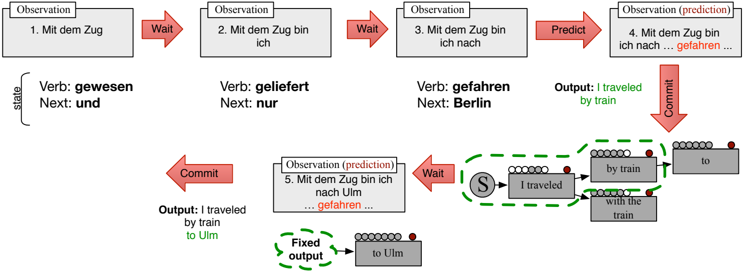
How we could translate a sentence



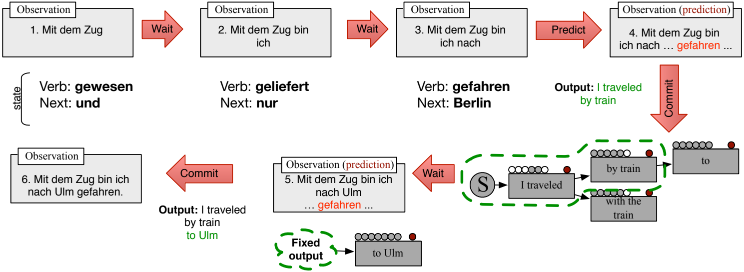
How we could translate a sentence



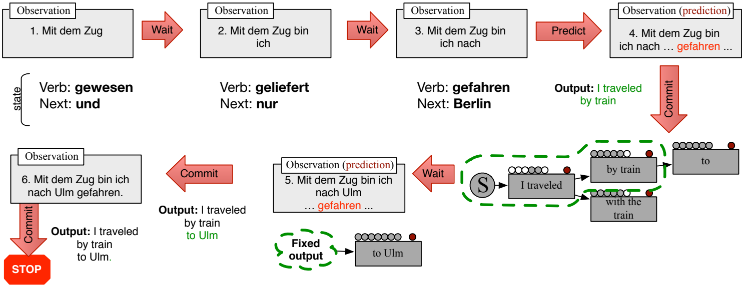
How we could translate a sentence



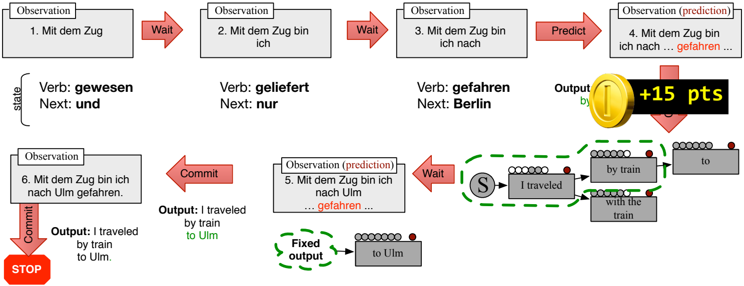
How we could translate a sentence



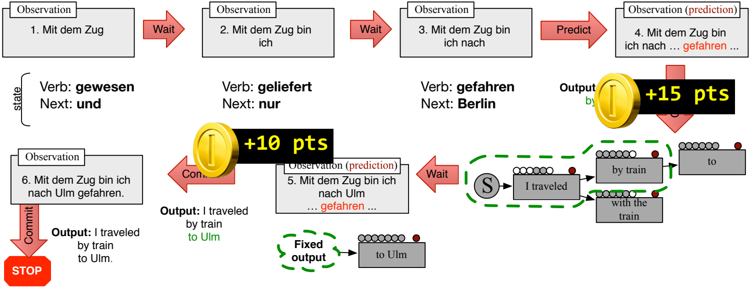
How we could translate a sentence



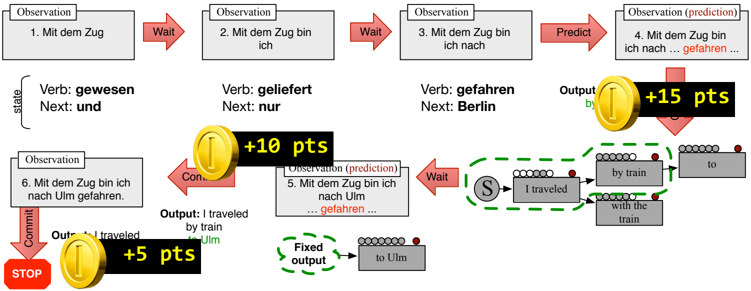
How we could translate a sentence



How we could translate a sentence



How we could translate a sentence



Adding meaning to topic models

Traditional Topic Models

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n}|\phi_{z_{d,n}}) \underbrace{p(z_{d,n}|\theta_d)}_{\text{topic}} \right) p(\theta_d|\alpha) \underbrace{\prod_k^K p(\phi_k|\eta)}_{\text{topic to words}}$$

Our Model

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n}|\pi_{l_{d,n}}) \underbrace{p(l_{d,n}|\phi_{d,n}) p(z_{d,n}|\theta_d)}_{\text{meaning and topic}} \right) p(\theta_d|\alpha)$$

$$\underbrace{\prod_k^K p(\phi_k|\eta)}_{\text{topic to concept}} \underbrace{\prod_c^C (p(\pi_{k,c}|\tau))}_{\text{concept to word}}$$

Adding meaning to topic models

Traditional Topic Models

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n}|\phi_{z_{d,n}}) \underbrace{p(z_{d,n}|\theta_d)}_{\text{topic}} \right) p(\theta_d|\alpha) \underbrace{\prod_k^K p(\phi_k|\eta)}_{\text{topic to words}}$$

Our Model

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n}|\pi_{l_{d,n}}) \underbrace{p(l_{d,n}|\phi_{d,n}) p(z_{d,n}|\theta_d)}_{\text{meaning and topic}} \right) p(\theta_d|\alpha)$$

$$\underbrace{\prod_k^K p(\phi_k|\eta)}_{\text{topic to concept}} \underbrace{\prod_c^C (p(\pi_{k,c}|\tau))}_{\text{concept to word}}$$

Adding meaning to topic models

Traditional Topic Models

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n}|\phi_{z_{d,n}}) \underbrace{p(z_{d,n}|\theta_d)}_{\text{topic}} \right) p(\theta_d|\alpha) \underbrace{\prod_k^K p(\phi_k|\eta)}_{\text{topic to words}}$$

Our Model

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n}|\pi_{l_{d,n}}) \underbrace{p(l_{d,n}|\phi_{d,n})p(z_{d,n}|\theta_d)}_{\text{meaning and topic}} \right) p(\theta_d|\alpha)$$

$$\underbrace{\prod_k^K p(\phi_k|\eta)}_{\text{topic to concept}} \underbrace{\prod_c^C (p(\pi_{k,c}|\tau))}_{\text{concept to word}}$$

Adding meaning to topic models

Traditional Topic Models

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n} | \phi_{z_{d,n}}) \underbrace{p(z_{d,n} | \theta_d)}_{\text{topic}} \right) p(\theta_d | \alpha) \underbrace{\prod_k^K p(\phi_k | \eta)}_{\text{topic to words}}$$

Our Model

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n} | \pi_{l_{d,n}}) \underbrace{p(l_{d,n} | \phi_{d,n}) p(z_{d,n} | \theta_d)}_{\text{meaning and topic}} \right) p(\theta_d | \alpha)$$

$$\underbrace{\prod_k^K p(\phi_k | \eta)}_{\text{topic to concept}} \underbrace{\prod_c^C (p(\pi_{k,c} | \tau))}_{\text{concept to word}}$$

Adding meaning to topic models

Traditional Topic Models

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n} | \phi_{z_{d,n}}) \underbrace{p(z_{d,n} | \theta_d)}_{\text{topic}} \right) p(\theta_d | \alpha) \underbrace{\prod_k^K p(\phi_k | \eta)}_{\text{topic to words}}$$

Our Model

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n} | \pi_{l_{d,n}}) \underbrace{p(l_{d,n} | \phi_{d,n}) p(z_{d,n} | \theta_d)}_{\text{meaning and topic}} \right) p(\theta_d | \alpha)$$

$$\underbrace{\prod_k^K p(\phi_k | \eta)}_{\text{topic to concept}} \underbrace{\prod_c^C (p(\pi_{k,c} | \tau))}_{\text{concept to word}}$$

Adding meaning to topic models

Traditional Topic Models

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n} | \phi_{z_{d,n}}) \underbrace{p(z_{d,n} | \theta_d)}_{\text{topic}} \right) p(\theta_d | \alpha) \underbrace{\prod_k^K p(\phi_k | \eta)}_{\text{topic to words}}$$

Our Model

$$p(w) = \prod_d \prod_n^{N_d} \left(p(w_{d,n} | \pi_{l_{d,n}}) \underbrace{p(l_{d,n} | \phi_{d,n}) p(z_{d,n} | \theta_d)}_{\text{meaning and topic}} \right) p(\theta_d | \alpha)$$

$$\underbrace{\prod_k^K p(\phi_k | \eta)}_{\text{topic to concept}} \underbrace{\prod_c^C (p(\pi_{k,c} | \tau))}_{\text{concept to word}}$$