Algorithms that learn to think on their feet

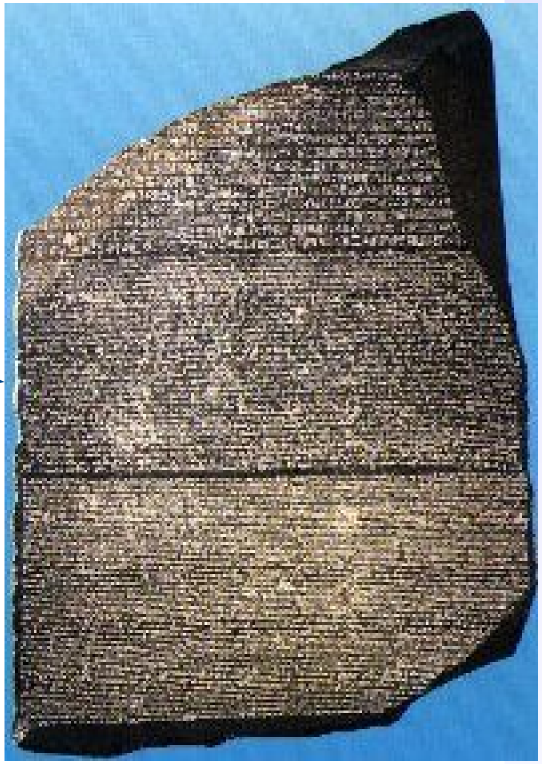
(now, with amazing bonus prize!)

Hal Daumé III | University of Maryland | me@hal3.name | @haldaume3

What is NLP?

- Fundamental goal: deep understanding of text
  - Not just string processing or keyword matching
- End systems that we want to build
  - Simple: Spelling correction, text categorization, etc.
  - Complex: Speech recognition, machine translation, information extraction, dialog interfaces, question answering
  - Unknown: human-level comprehension (more than just NLP?)

Why is language hard?



- Ambiguity abounds (some headlines)
  - Iraqi Head Seeks Arms
  - Teacher Strikes Idle Kids
  - Kids Make Nutritious Snacks
  - Stolen Painting Found by Tree
  - Local HS Dropouts Cut in Half
  - Enraged Cow Injures Farmer with Ax
  - Hospitals are Sued by 7 Foot Doctors
  - Ban on Nude Dancing on Governor's Desk
  - Scientists study whales from space
- Why are these funny?
- What does ambiguity imply about the role of learning?

Despite ambiguity, language is predictable

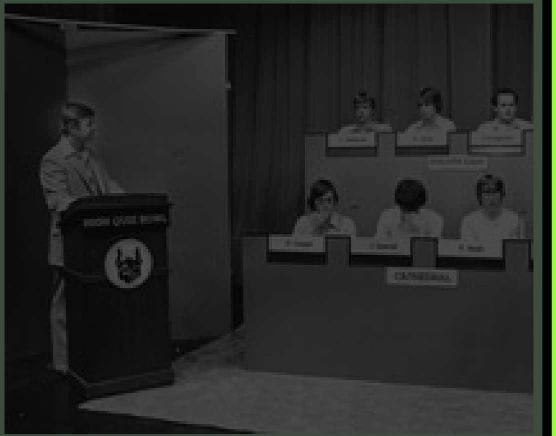
I like my coffee with cream and *asparagus*

This is crummy weather for San *ta Claus*

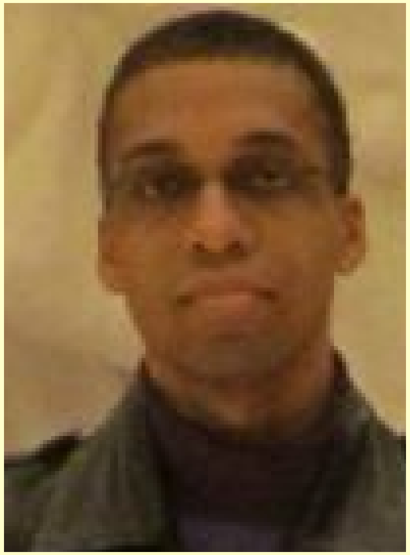
- The brain uses this information!
- Can we use predictability to make decisions *before* all of the input is observed?

YES ! ! !

Outline



Quizbowl (Incremental Question Answering)



Alvin Grissom II

He He

train to Ulm

taneous Machine retation

Simultaneous (machine) interpretation

Nuremburg Trials

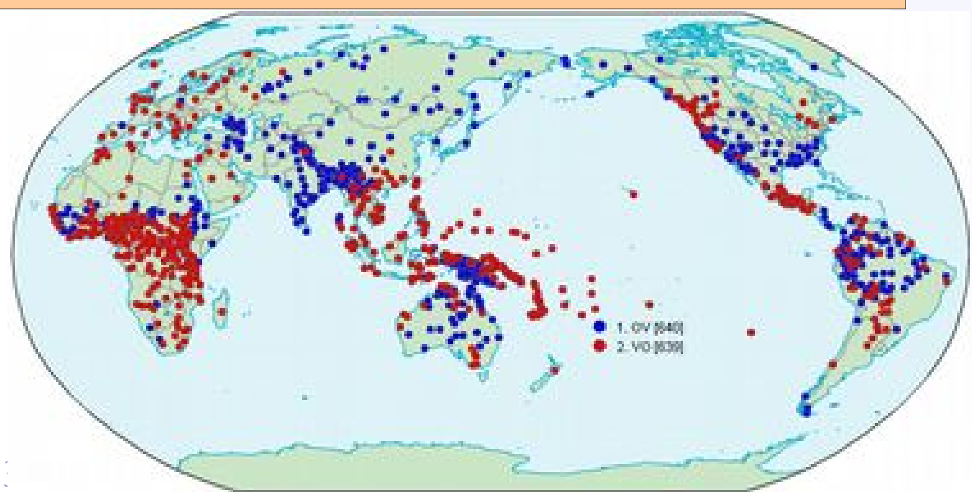
- Dozens of defendants
- Judges from four nations (three languages)
- Status quo: speak, then translate
- After Nuremberg, simultaneous translations became the norm
- Long wait → bad conversation

Why simultaneous interpretation is hard

- Human languages have vastly different word orders
  - About half are OV, the other half are VO
  - This comes with a lot more baggage than just verb-final

Running (German/English) Example:

Ich	bin	mit	dem	Zug	nach	Ulm	gefahren	
I	am	with	the	train	to	Ulm	traveled	
I	*(……waiting……)*						traveled by train to Ulm	



Model for interpretation decisions

- We have a set of actions (predict / translate)
  - Wait
  - Predict clause-verb
  - Predict next word
  - Commit ("speak")
- In a changing environment (state)
  - The words we've seen so far
  - Our models' internal predictions
- With a well defined notion of "optimal action" at training time

Example of interpretation trajectory

Ich	bin	mit	dem	Zug	nach	Ulm	gefahren	
I	am	with	the	train	to	Ulm	traveled	
I	*(……waiting……)*						traveled by train to Ulm	

DAgger: Dataset Aggregation

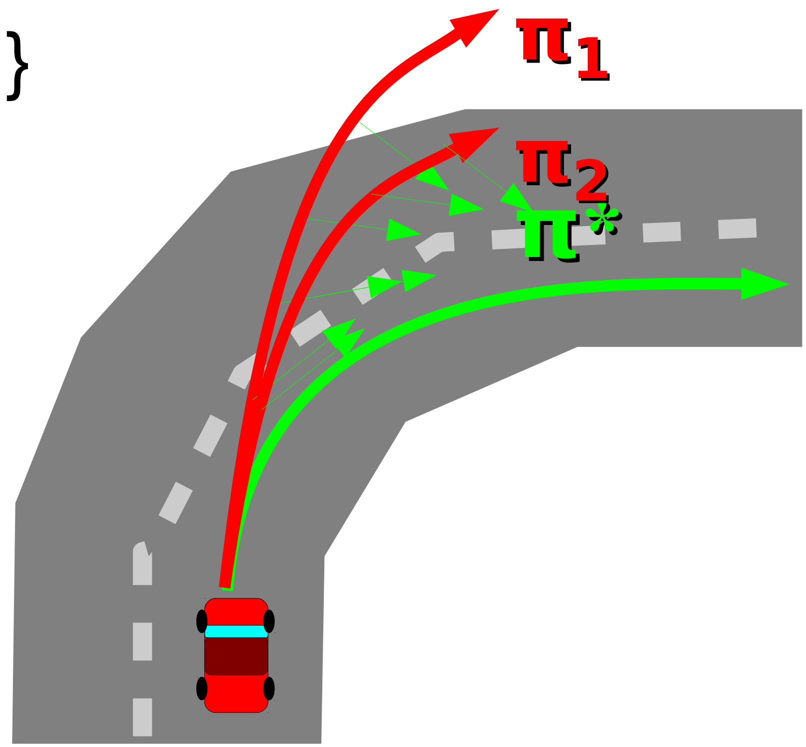
- Collect trajectories from expert π
- Dataset $\mathbf{D_0} = \{ (s, \pi^*(s)) \mid s \sim \pi^* \}$
- Train π_1 on $\mathbf{D_0}$
- Collect new trajectories from π_1
  - But let the *expert* steer!
- Dataset $\mathbf{D_1} = \{ (s, \pi^*(s)) \mid s \sim \pi_1 \}$
- Train π_2 on $\mathbf{D_0} \cup \mathbf{D_1}$

- In general:
  - $\mathbf{D_n} = \{ (s, \pi^*(s)) \mid s \sim \pi_n \}$
  - Train π_n on $\bigcup_{i<n} \mathbf{D_i}$

If $N = T \log T$,

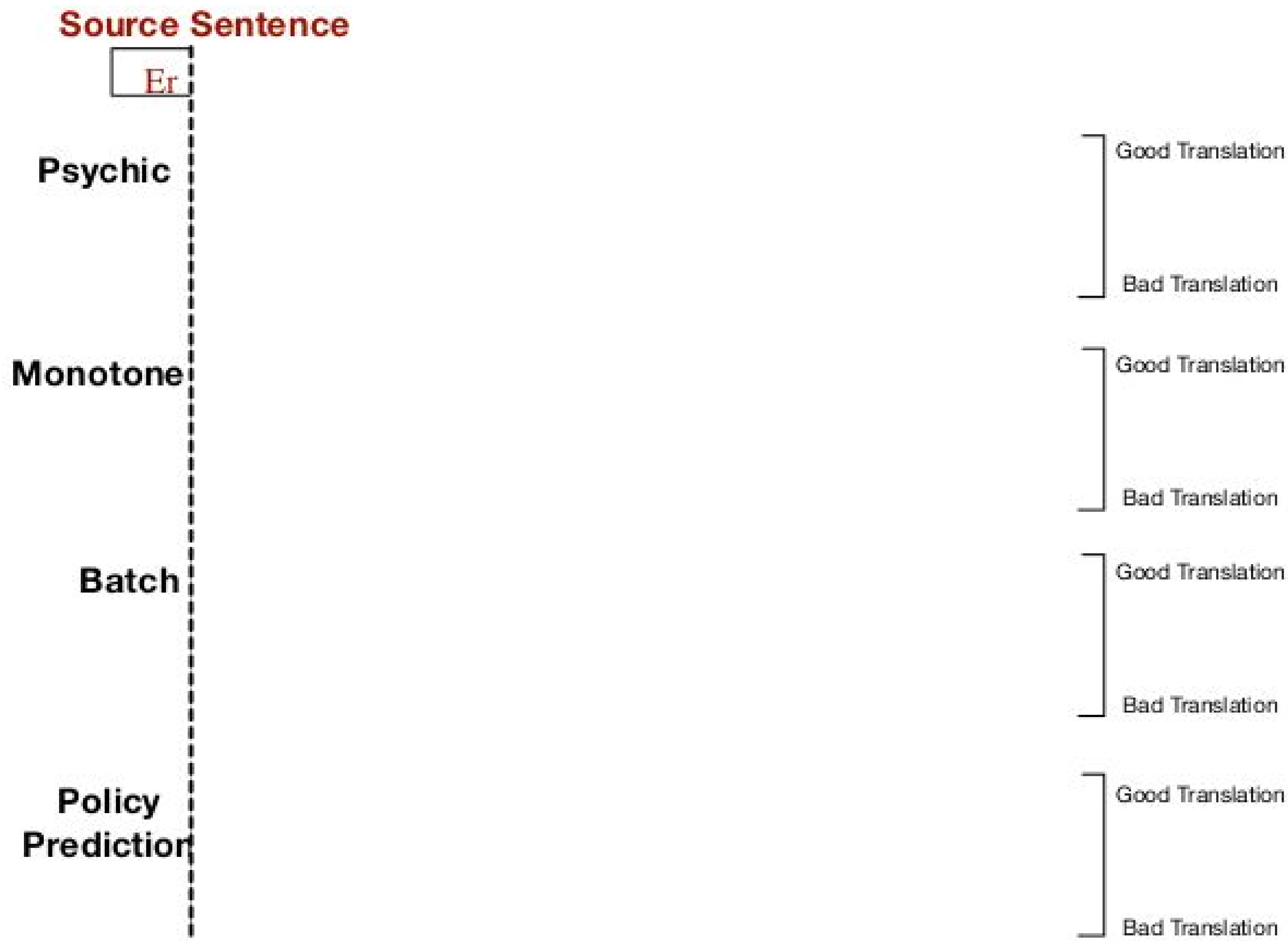
$$\mathbf{L(\pi_n) < T\, \epsilon_N + O(1)}$$

for some n



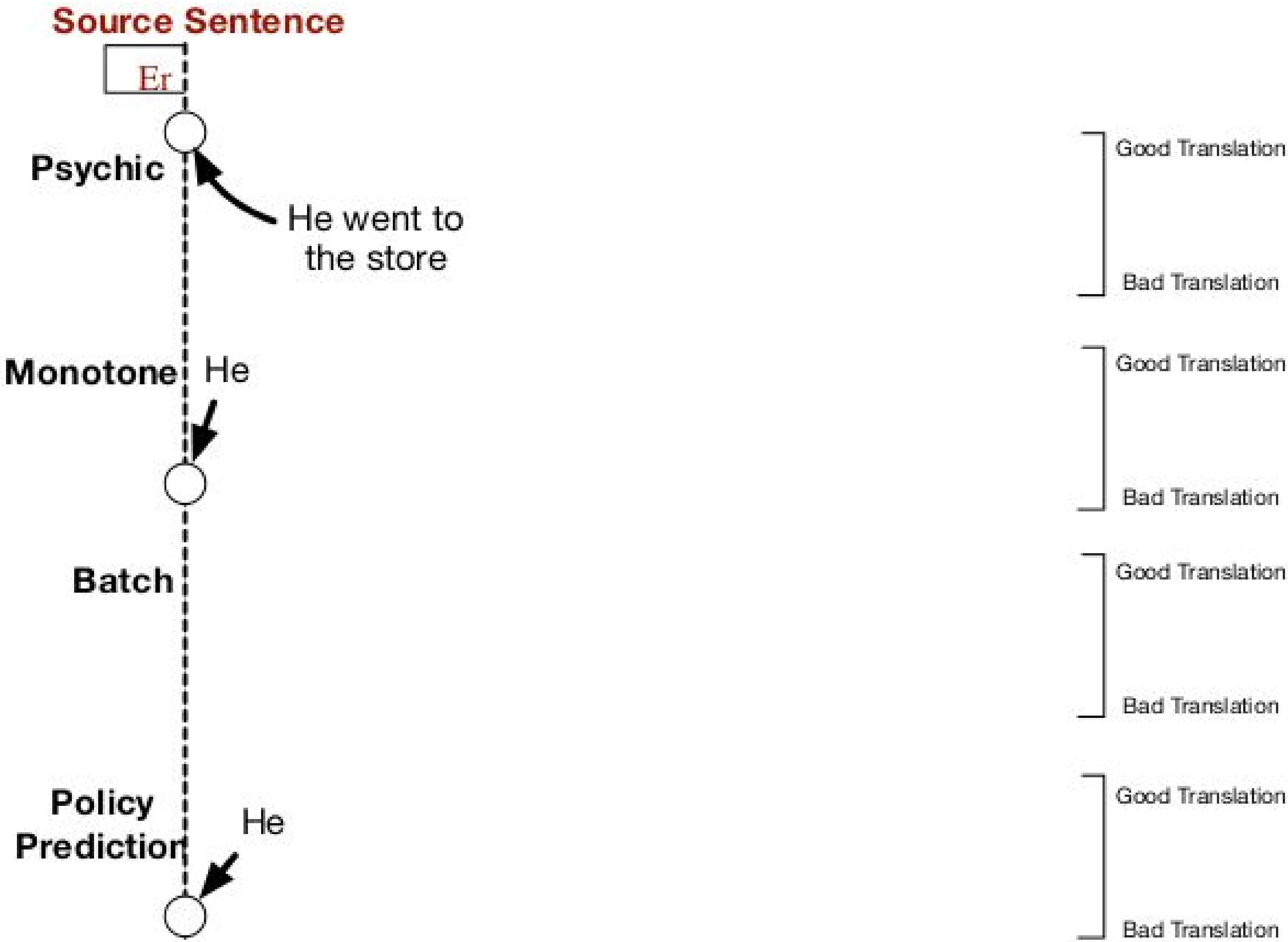
(Ross+Gordon+Bagnell, AIStats 2011)

Evaluating performance and baselines



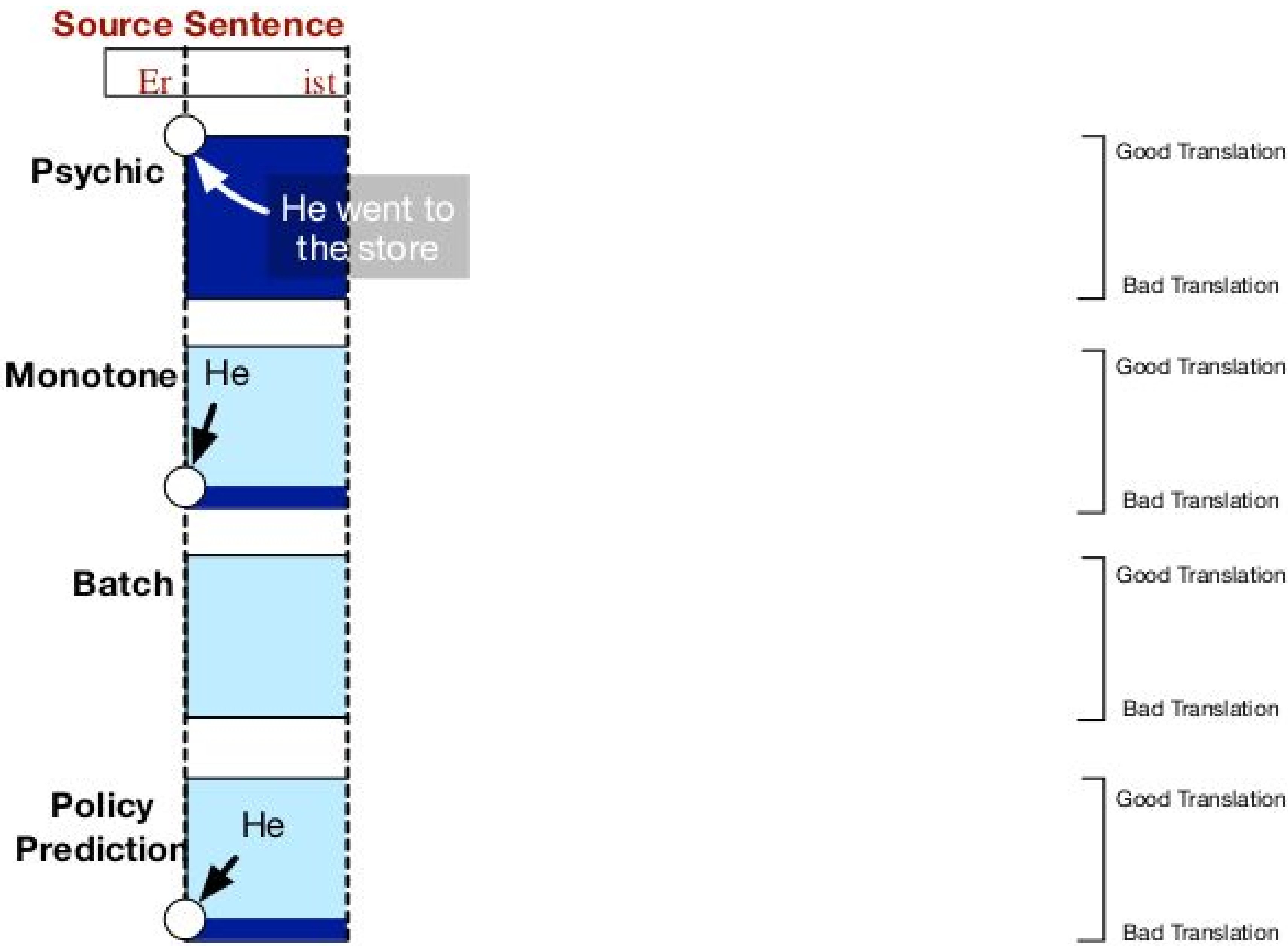
(Grissom II et al., EMNLP 2014)

Evaluating performance and baselines



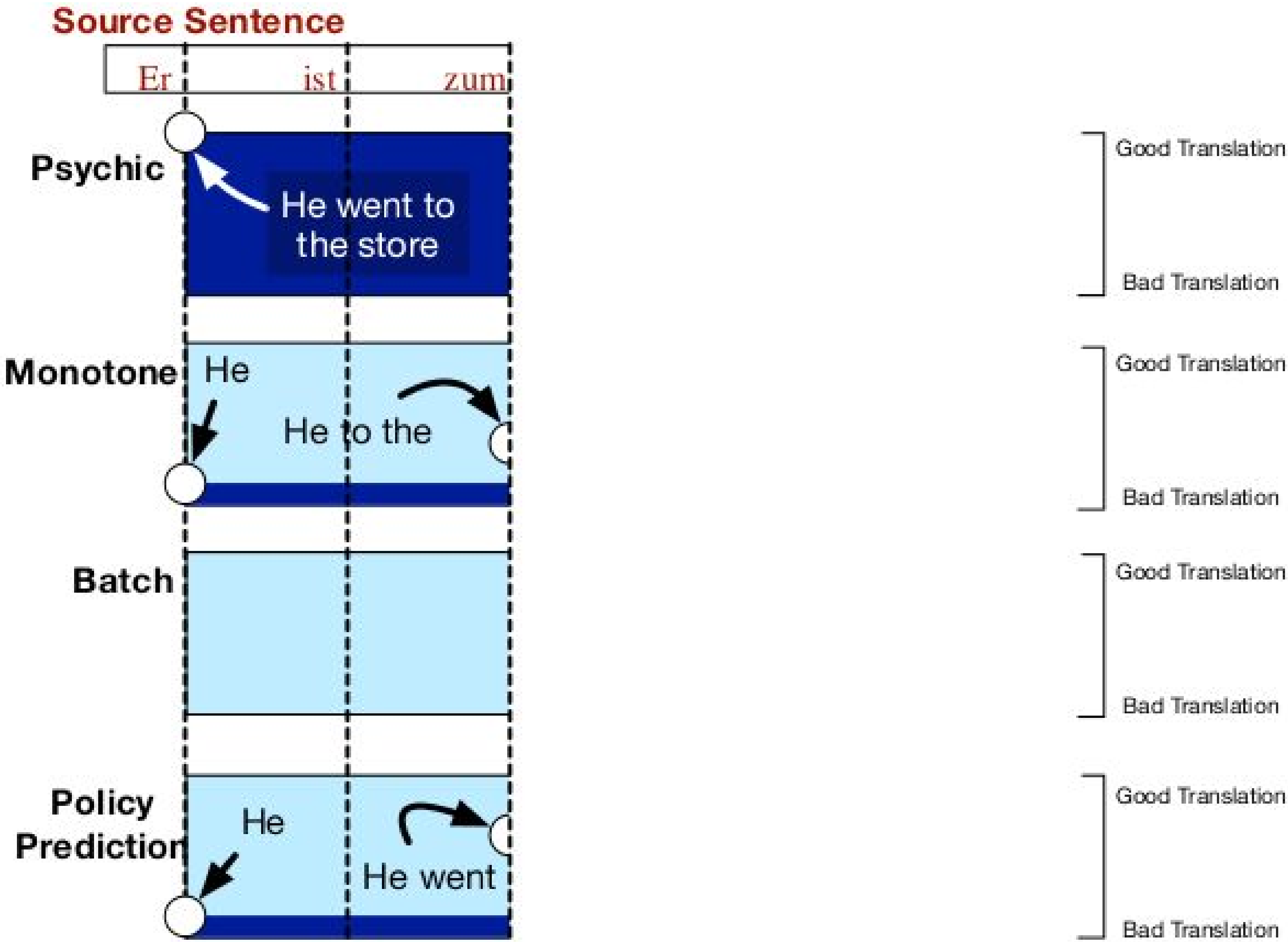
(Grissom II et al., EMNLP 2014)

Evaluating performance and baselines



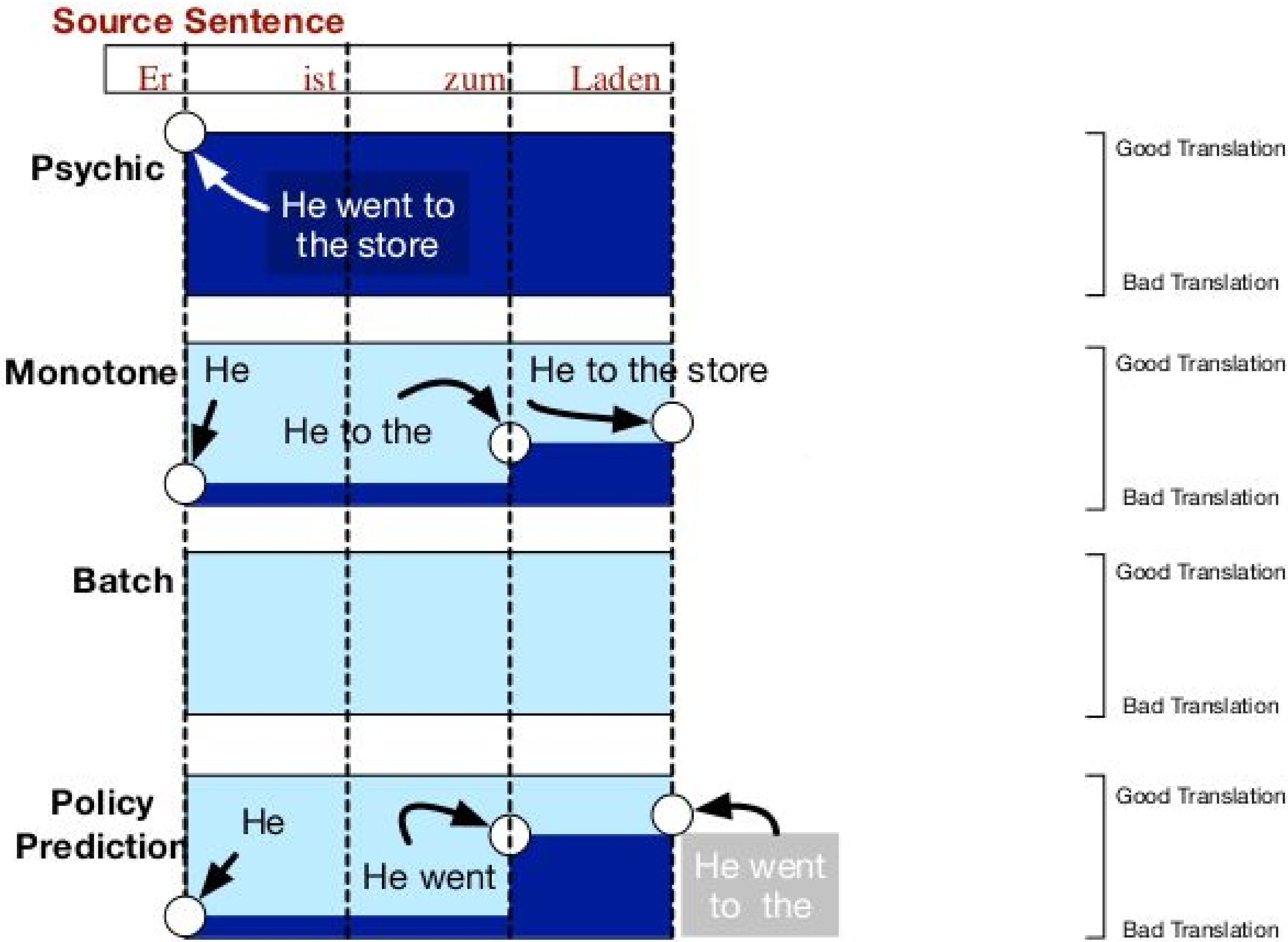
(Grissom II et al., EMNLP 2014)

Evaluating performance and baselines



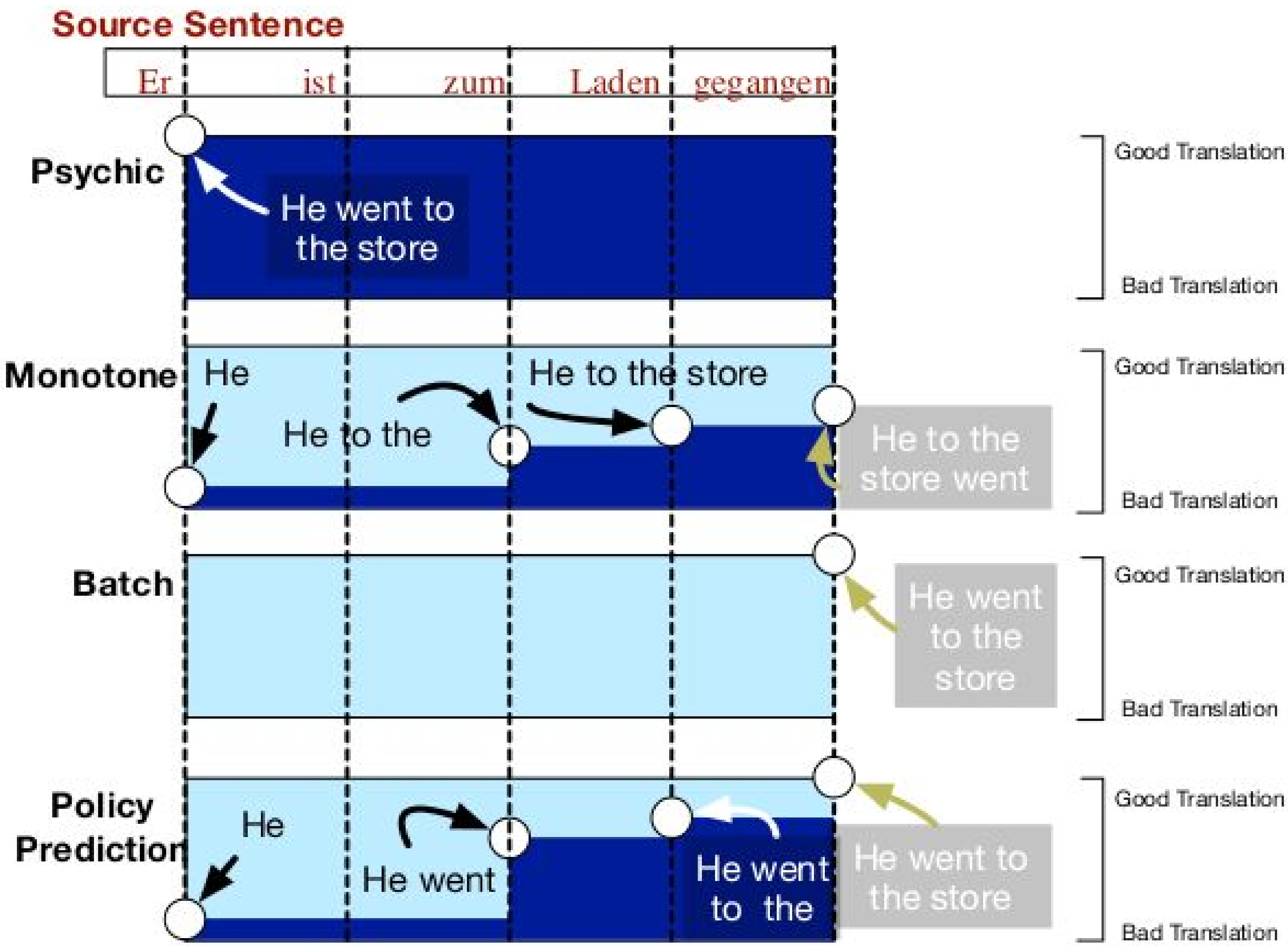
(Grissom II et al., EMNLP 2014)

Evaluating performance and baselines



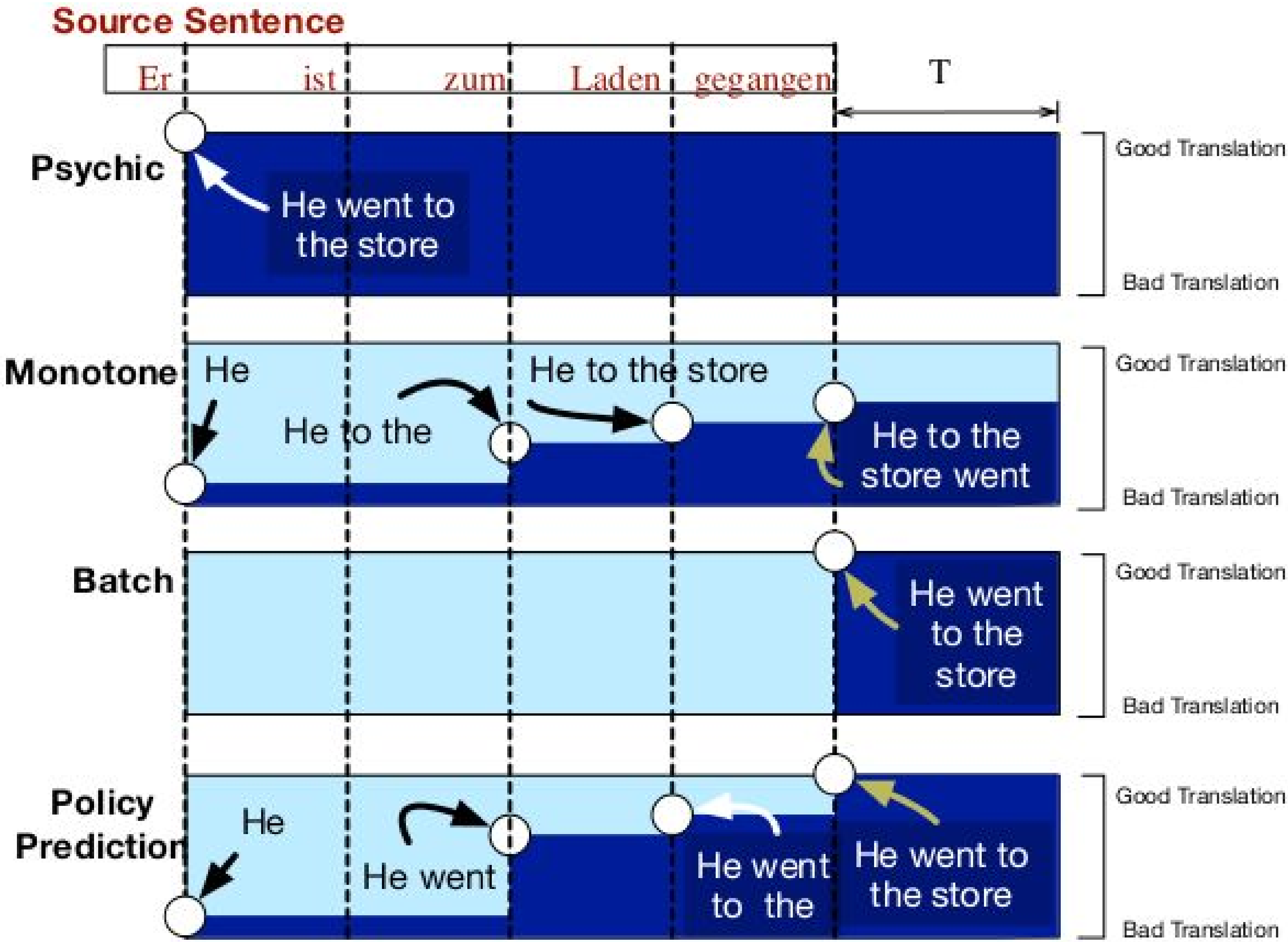
(Grissom II et al., EMNLP 2014)

Evaluating performance and baselines



(Grissom II et al., EMNLP 2014)

Evaluating performance and baselines



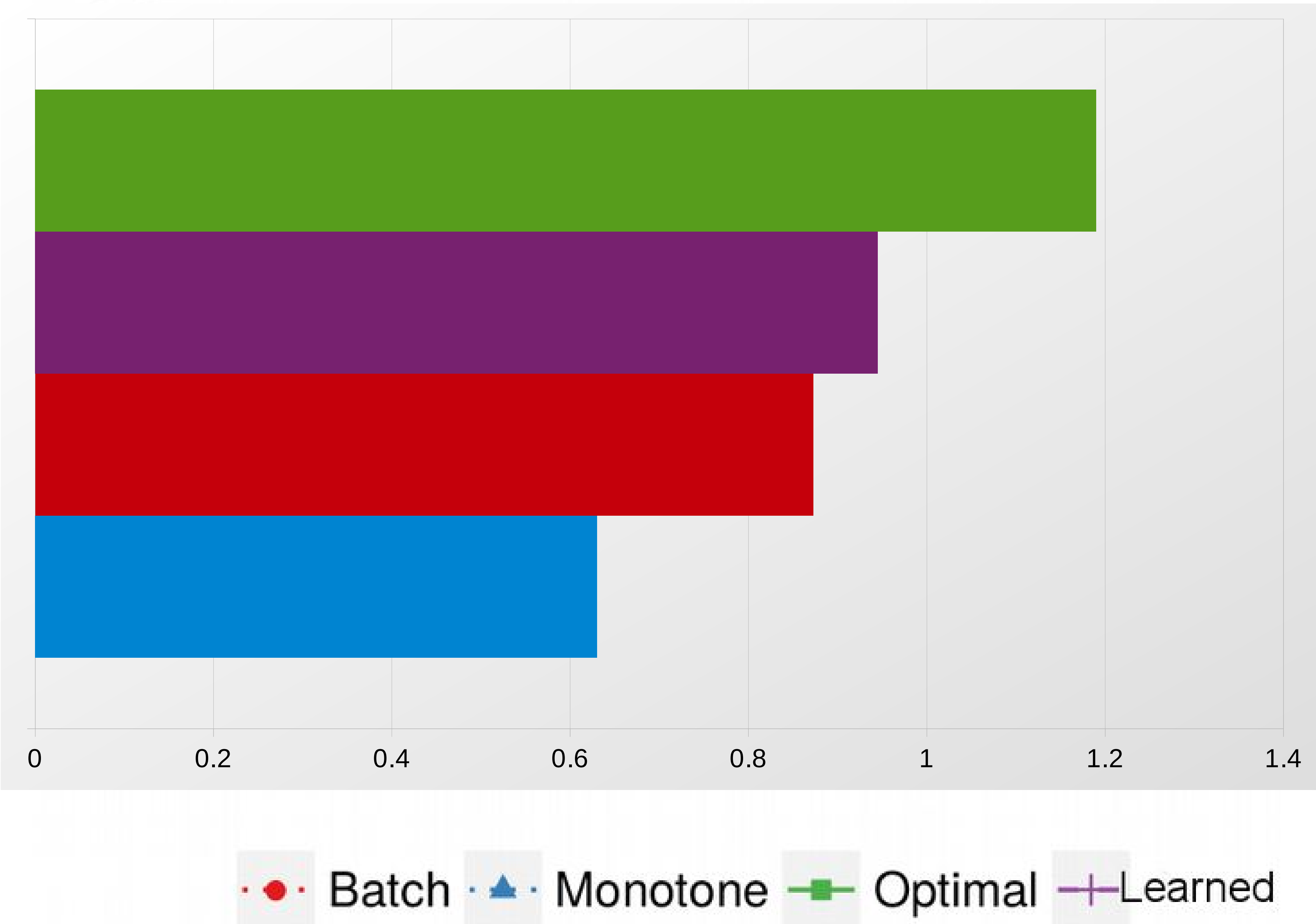
(Grissom II et al., EMNLP 2014)

Training the policy

- Actions:
  - Commit `translate(revealed words)`
  - Predict (verb/next) `translate(revealed + predicted)`
  - Wait `get_next_words()`
- Delayed feedback: latency BLEU
- Features:
  - Output & confidence of predictors
  - Internal translation / language model scores
  - Previous decisions made by policy

(Grissom II et al., EMNLP 2014)

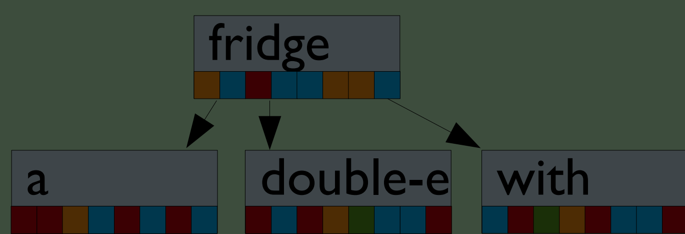
Evaluating performance



(Grissom II et al., EMNLP 2014)

Outline

Quizbowl (Incremental Question Answering)



Mohit Iyyer

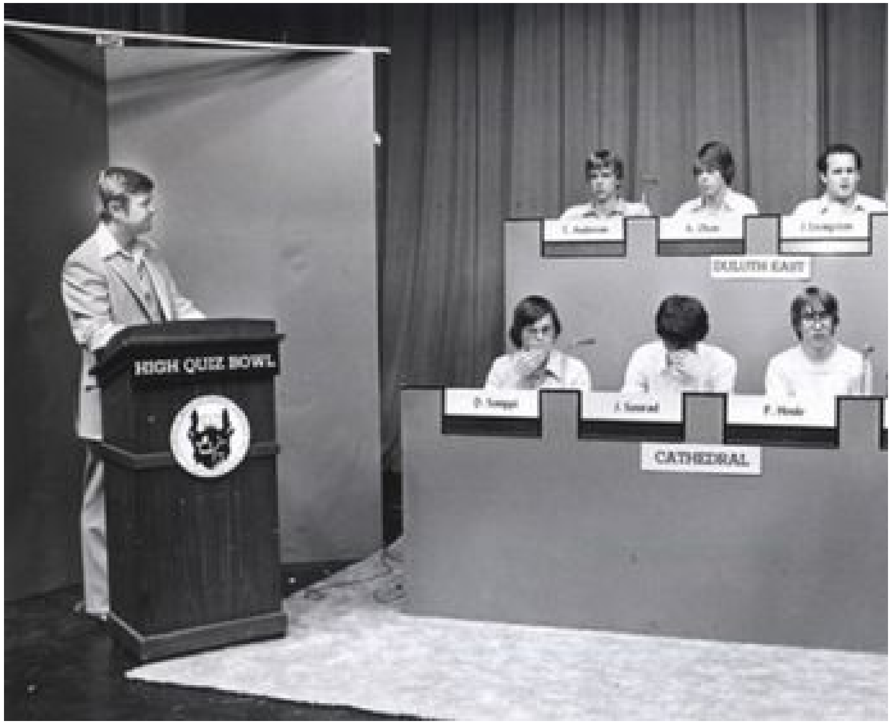
He He

train to Ulm

taneous
Machine
retation

Humans doing incremental prediction

- Game called "quiz bowl"
- Two teams play each other
  - Moderator reads a question
  - When a team knows the answer, they buzz in
  - **If right**, they get points; **otherwise**, rest of the question is read to the other team
- Hundreds of teams in the US alone
- Example …



(Iyyer et al., ACL 2014)

Quizbowl example

With Leo Szilard, he invented a doubly-eponymous

(Iyyer et al., ACL 2014)

Quizbowl example

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

(Iyyer et al., ACL 2014)

Quizbowl example

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

(Iyyer et al., ACL 2014)

Quizbowl example

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

(Iyyer et al., ACL 2014)

Quizbowl example

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

(Iyyer et al., ACL 2014)

Quizbowl example

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

(Iyyer et al., ACL 2014)

Quizbowl example

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?

(Iyyer et al., ACL 2014)

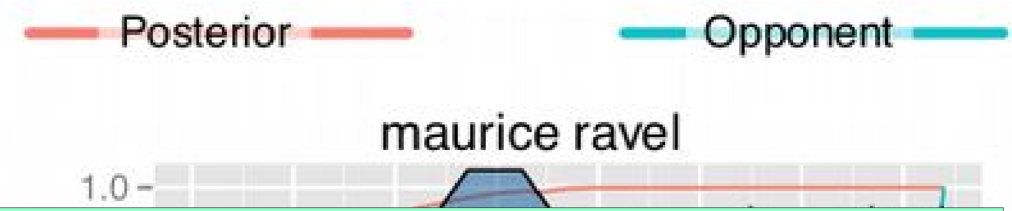
Solving incrementally

- Action: buzz now or wait
  - Content Model is constantly generating guesses
  - Oracle provides examples where it is correct
  - The Policy generalizes to test data
  - Features represent our state

(Iyyer et al., ACL 2014)

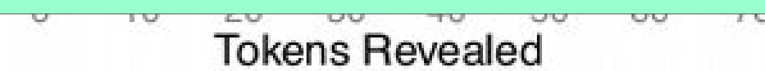
Evaluation methodology

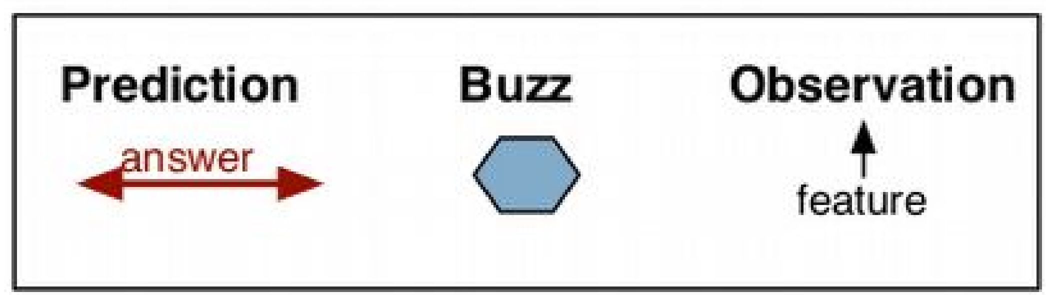
- Mechanical Turk to collect human data
- 7000 questions were
-
- Total of 461 unique users
- Leaderboard to encourage users



Big problem:

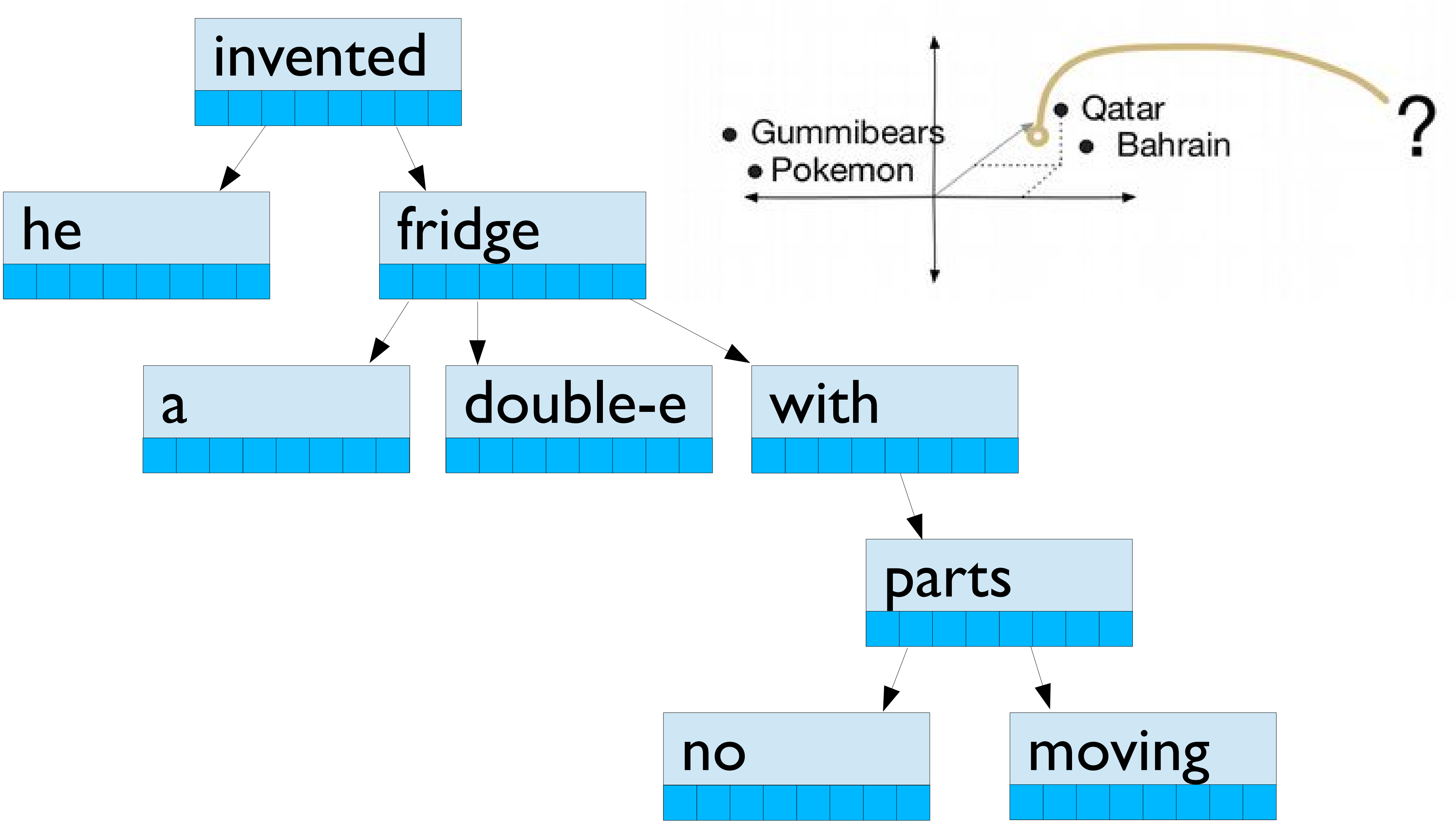
"this man shot at Aaron Burr"
is very different from
"Aaron Burr shot at this man"





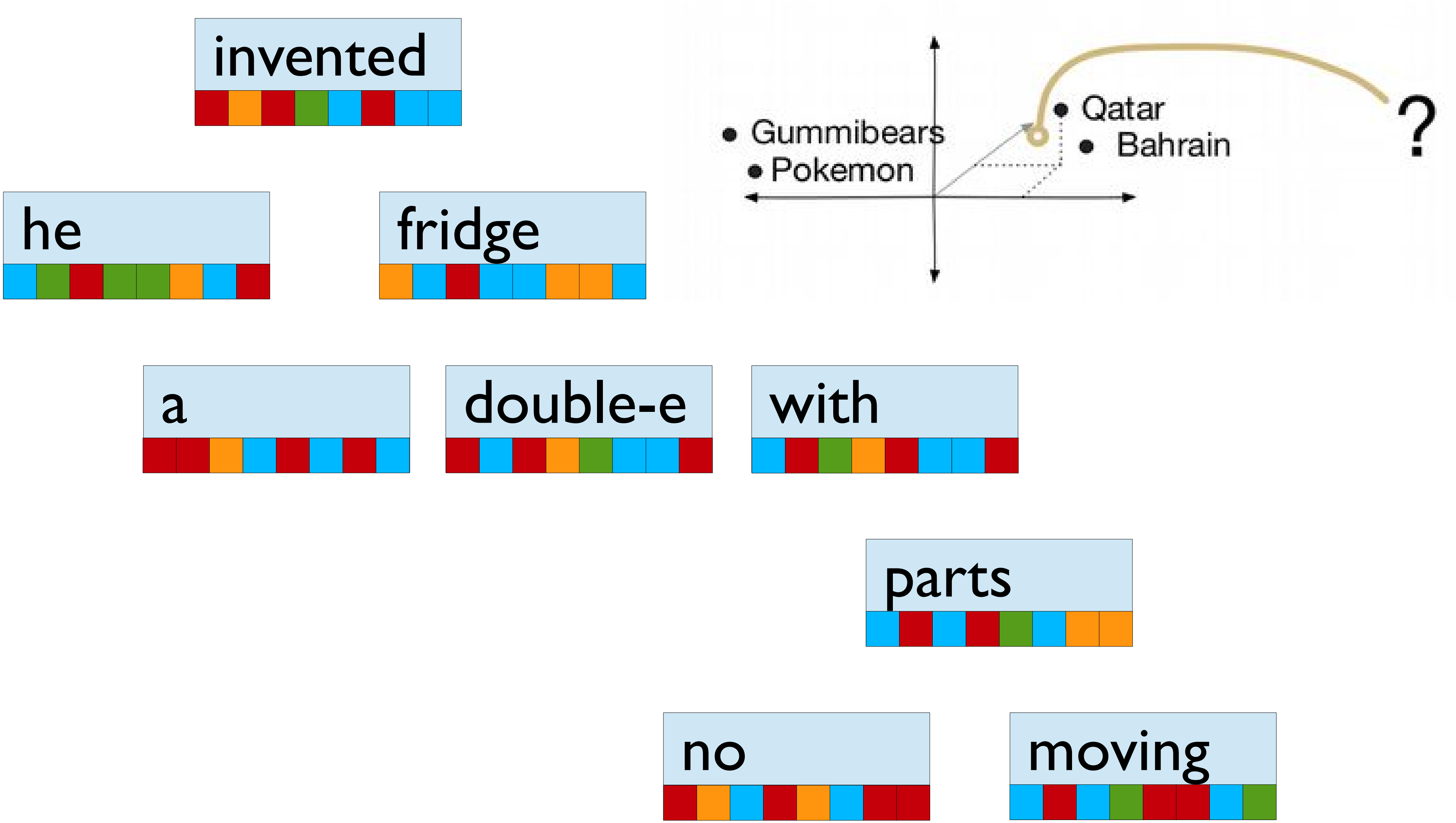
(Iyyer et al., ACL 2014)

Challenge: modeling compositionality



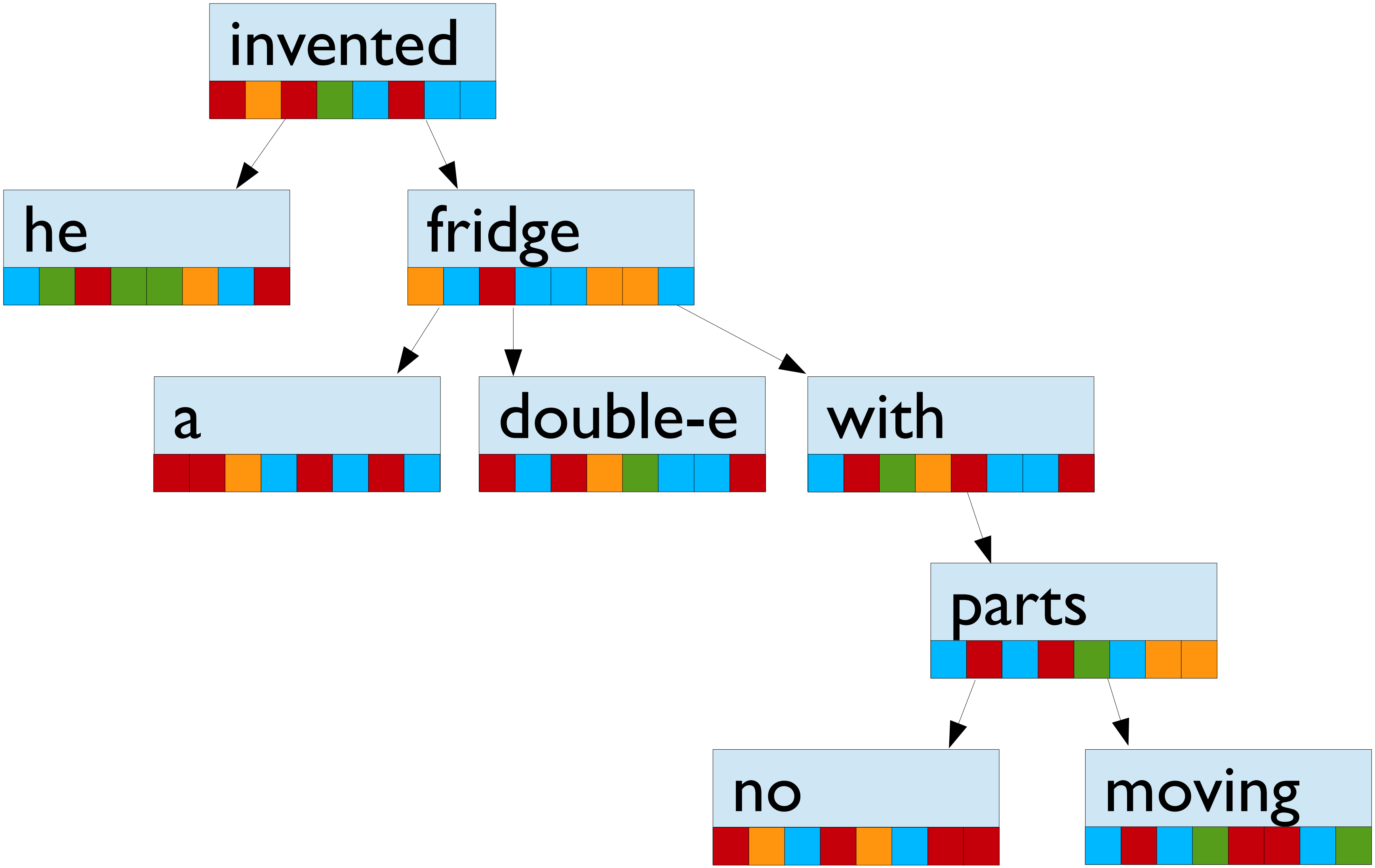
(Iyyer et al., ACL 2014)

Challenge: modeling compositionality



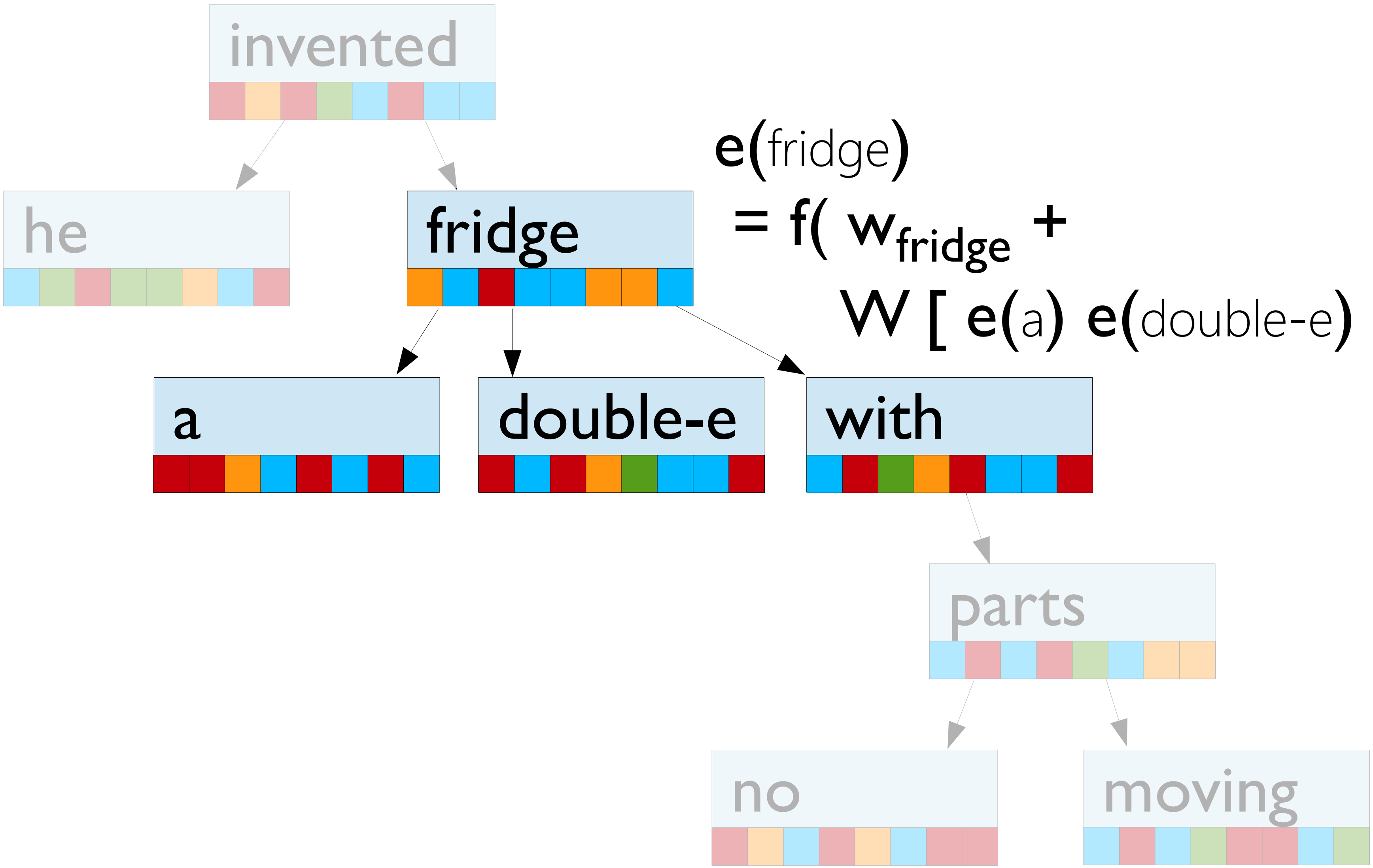
(Iyer et al., ACL 2014)

Challenge: modeling compositionality



(Iyyer et al., ACL 2014)

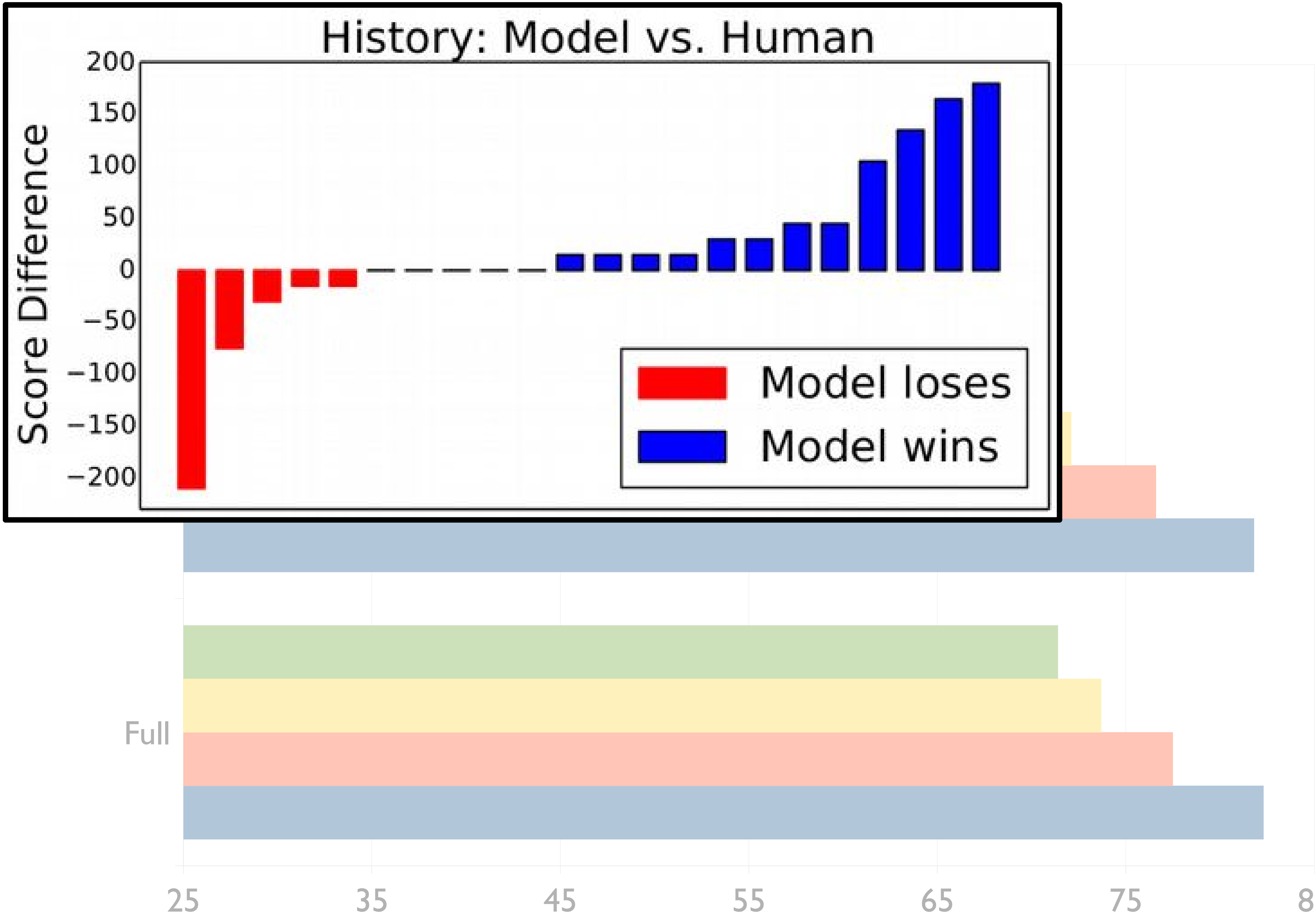
Challenge: modeling compositionality



$$e(\text{fridge}) = f(\, w_{\text{fridge}} + W\,[\, e(\text{a})\ e(\text{double-e})\ e(\text{with})\,]\,)$$

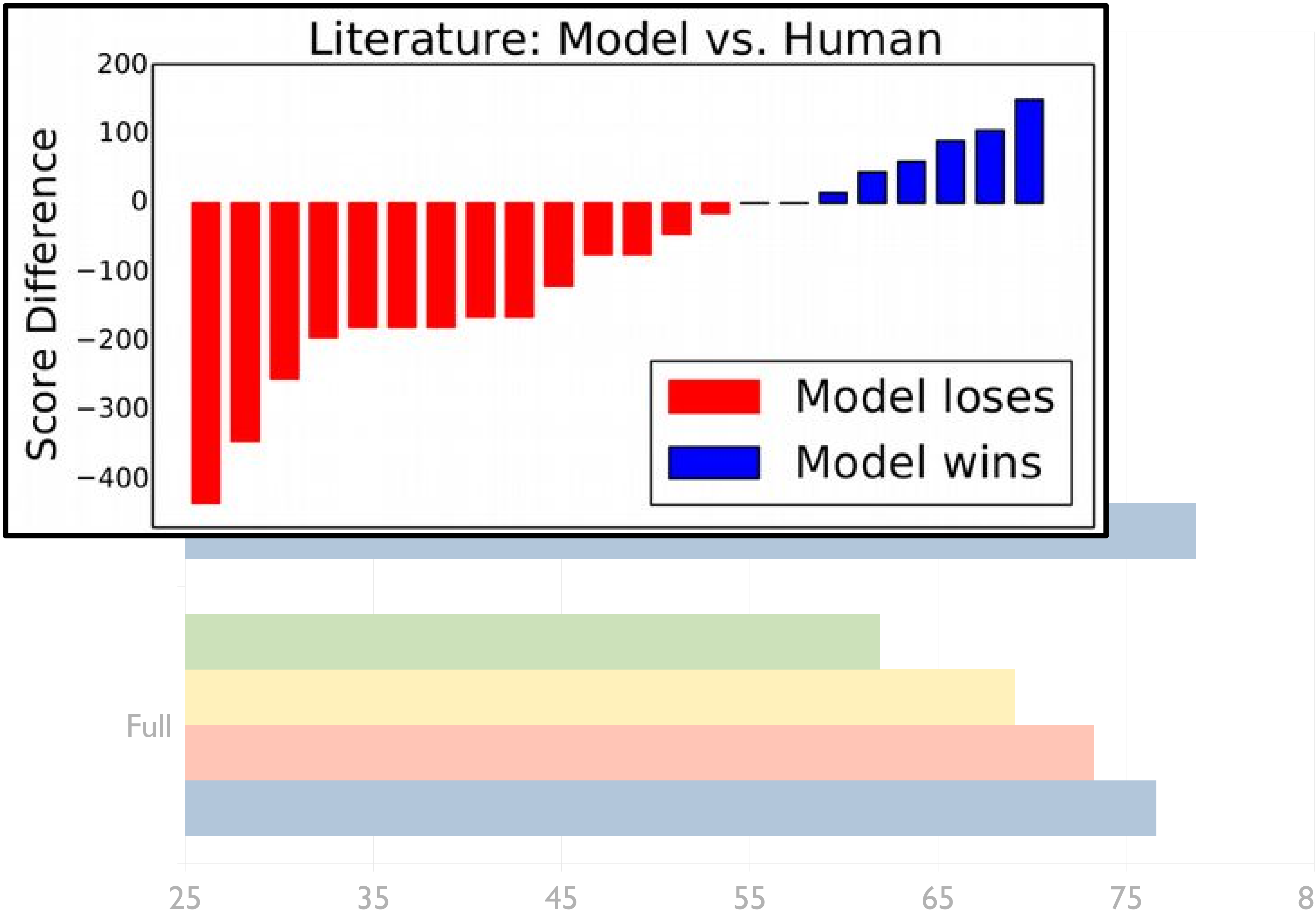
(Iyyer et al., ACL 2014)

Results on question-answering task



(Iyyer et al., ACL 2014)

Results on question-answering task



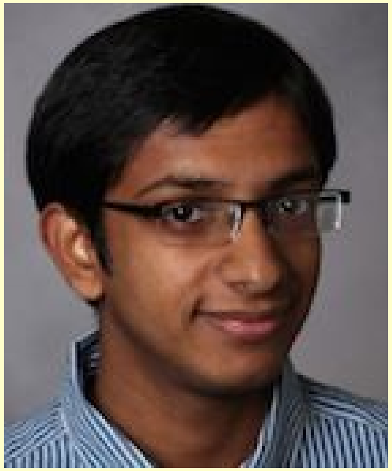
(Iyyer et al., ACL 2014)

But the true test…

But the true test… RESULTS!

SUPER thanks to Ken Jennings for being a great sport!

Outline



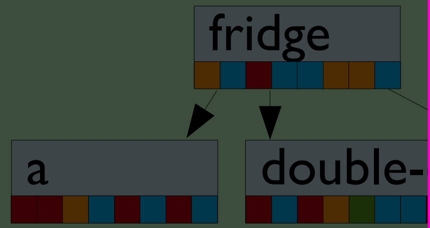
Alekh Agarwal

Kai-Wei Chang

Akshay Krishnamurthy

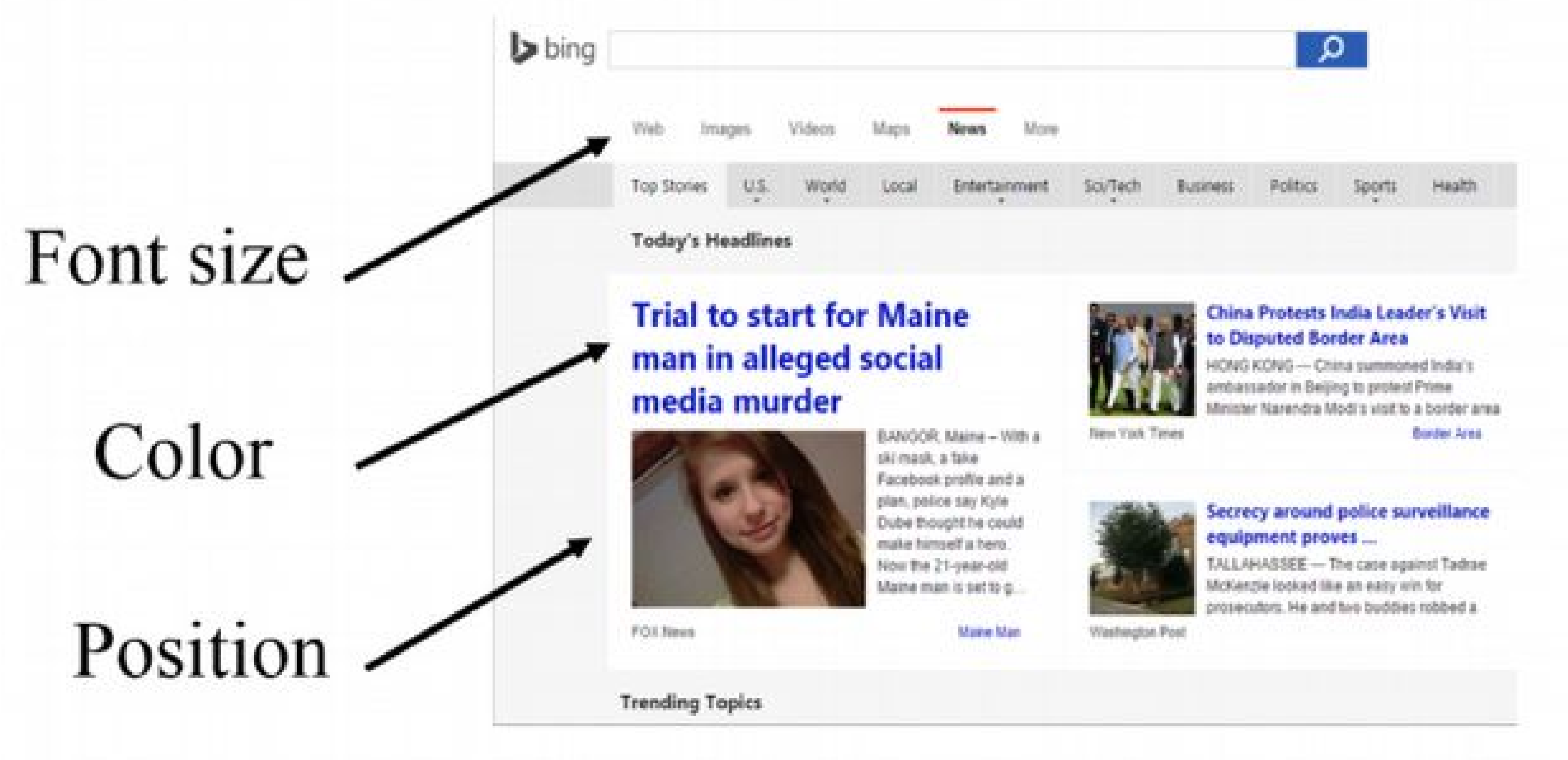
John Langford

Question
Answering)



BONUS PRIZE!

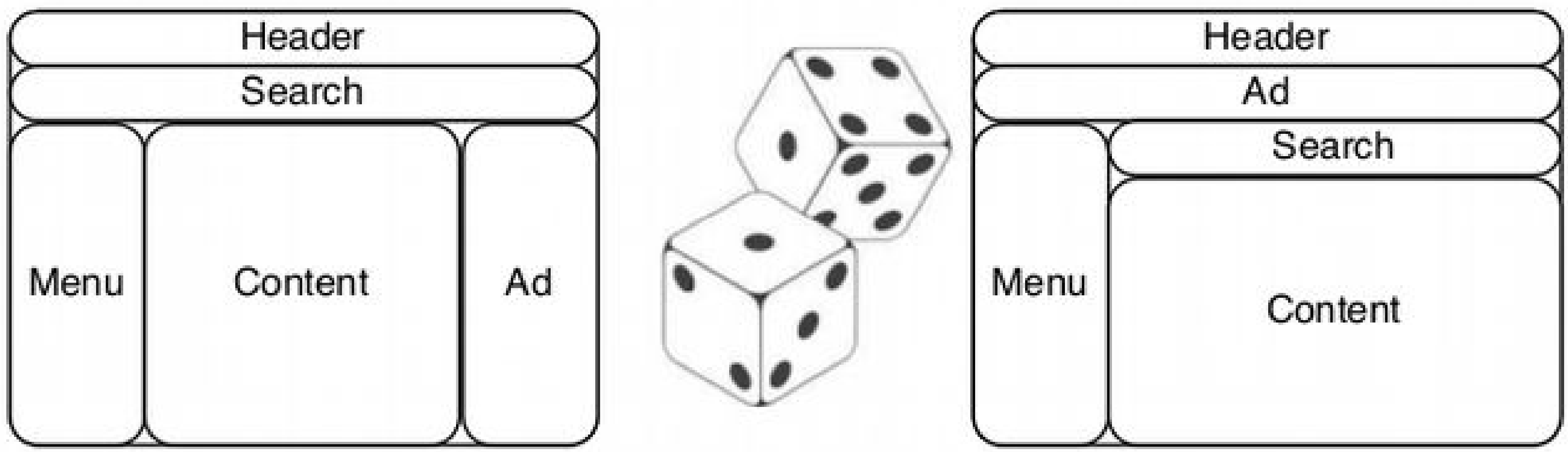
Structured learning with partial feedback



- Loss of a **single structured label** can be observed
- Labels are *never* observed

Solution strategy

- Use randomization to estimate losses
- Apply "standard" learning-to-search to losses



Learning to search

- Convert structured prediction into a search problem
  - search space and actions
- Define structured features over each state
- Construct a reference policy (Ref)
  - Ref usually defined using true label
- Learn a policy that imitates Ref
  - Implement with a cost-sensitive classifier

Structured contextual bandit challenge

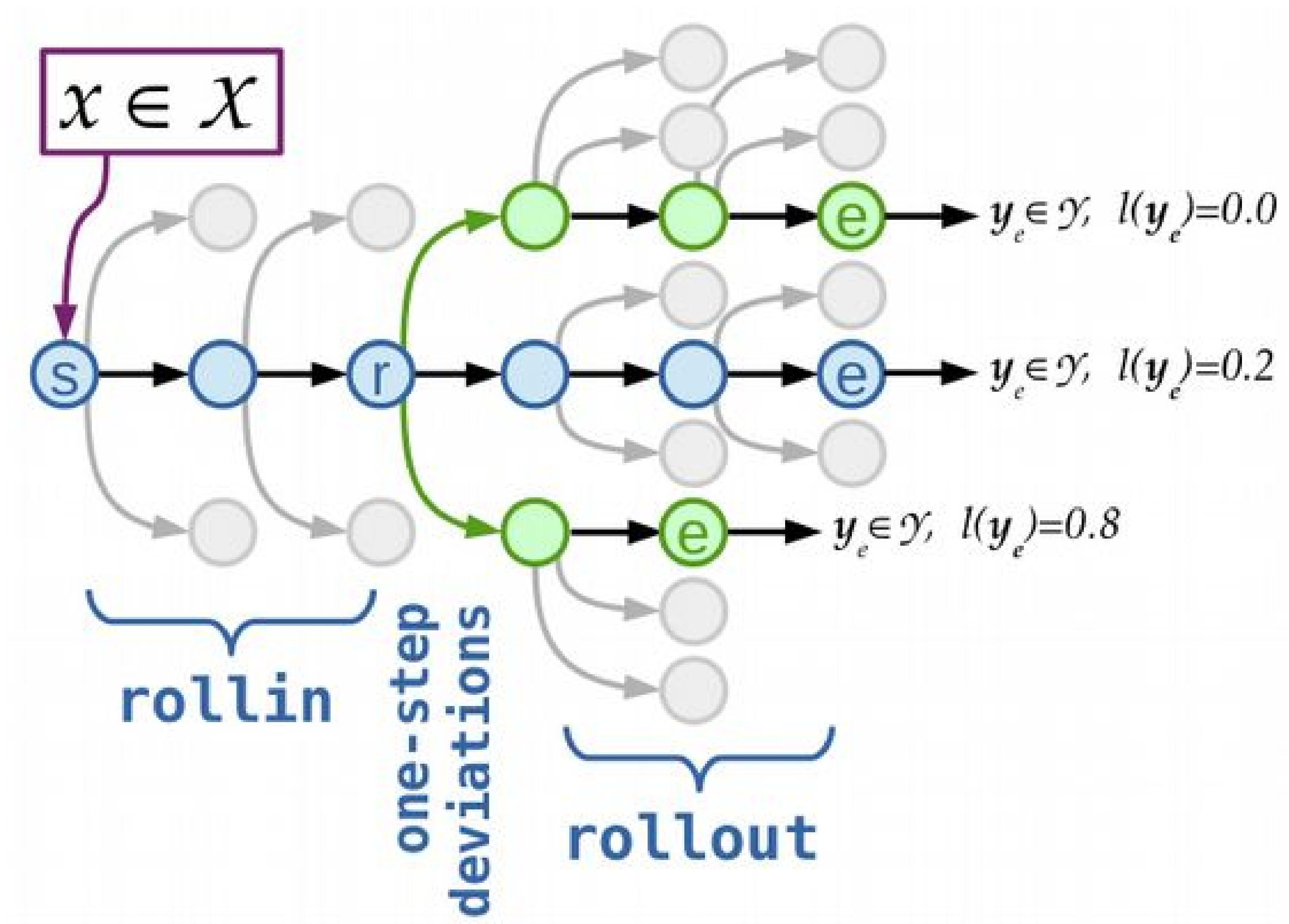
- True label is not available => Hard to define good Ref
  - Existing L2S algorithms give:

$$R(\pi) \leq R(\pi^{ref}) + o(1)$$

- Can use status quo system as Ref,
  - But competing with this Ref is not useful!

- **Main goal: Learning to search with:**
  - **A suboptimal reference => *improve* on Ref**
  - **Partial feedback**

Learning to search "schematic"



- Desiridata:
  - Compete with Ref (global opt if Ref is optimal and realizable)
  - Local optimality

Effect of Roll-{in,out} policies

roll-out → ↓ roll-in	Reference	Mixture	Learned
Reference	Inconsistent		
Learned	No local opt	Good	RL

Mixture: w.p. β use Ref, else use Learned

Theorem: LOLS minimizes a combination of regret to Ref and regret to its own one-step deviations

Theorem: Can take $\Omega(2^T)$ steps to reach *local* optimality

Does it work in practice?

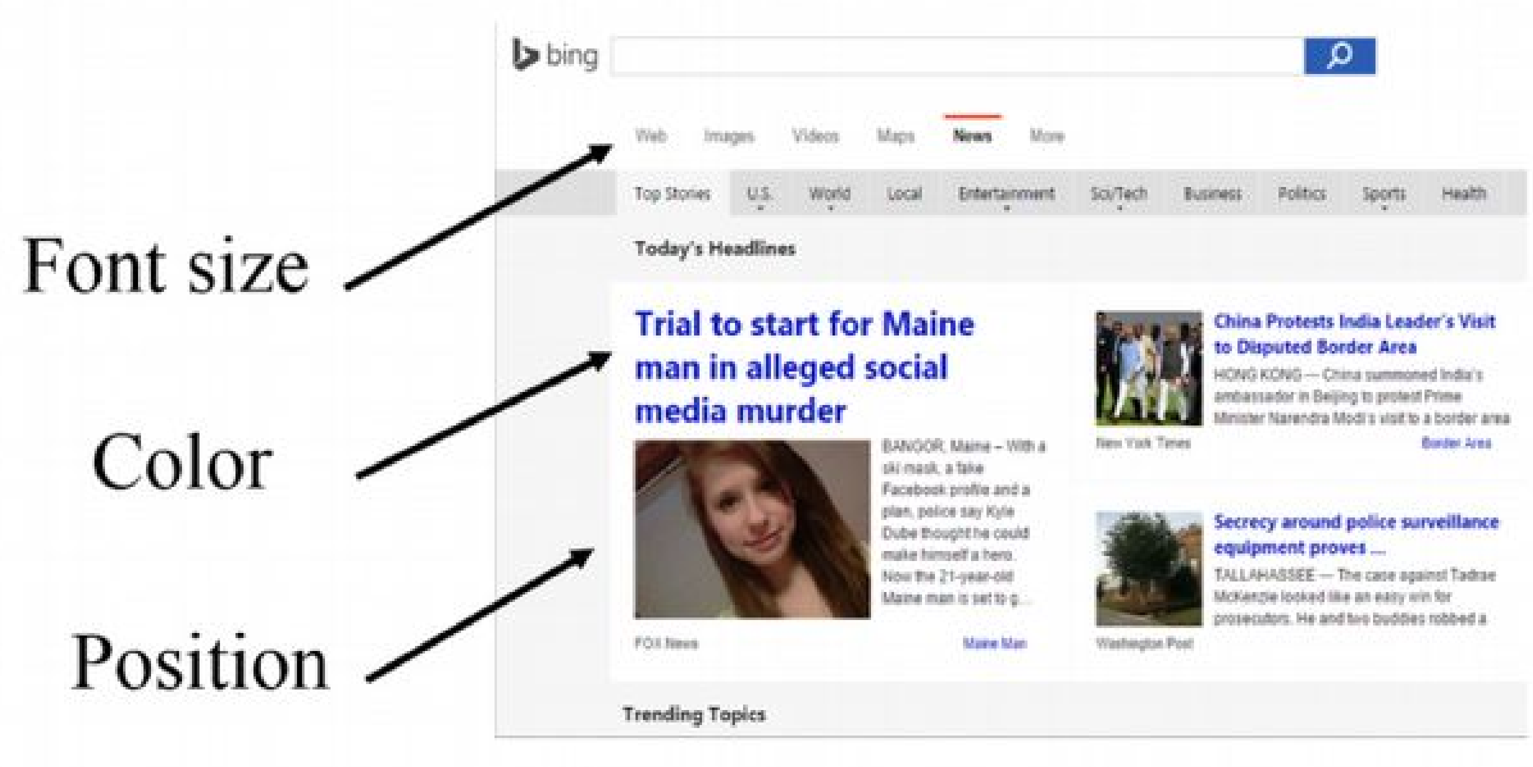
- Experiments on Dependency Parsing

roll-out → ↓ roll-in	**Reference**	**Mixture**	**Learned**
	Reference is optimal		
Reference	87.2	89.7	88.2
Learned	**90.7**	90.5	86.9
	Reference is suboptimal		
Reference	83.3	87.2	81.6
Learned	87.1	**90.2**	86.8
	Reference is bad		
Reference	68.7	65.4	66.7
Learned	75.8	**89.4**	87.5

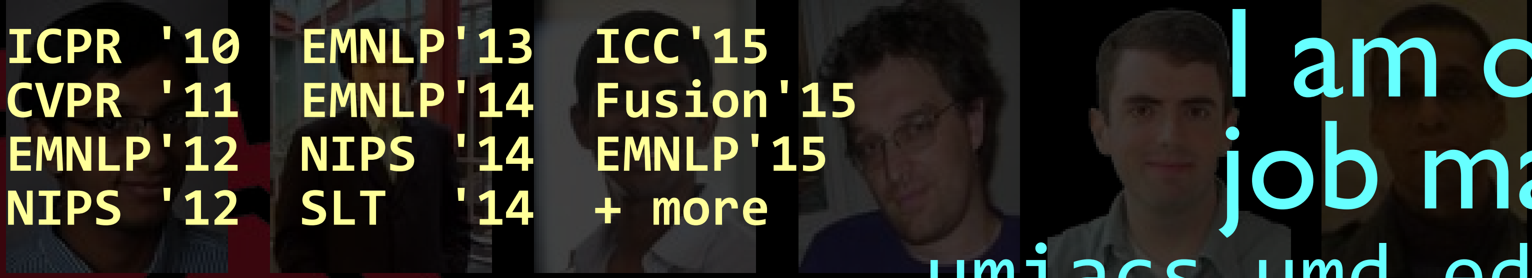
- LOLS always good, even with Ref is bad

Learning with partial feedback

- Loss of a *single* structured label can be observed
- Reference policy is *not optimal*



- Apply an ∈-greedy strategy
- Regret to ref and one-step deviations still bounded



ICPR '10 EMNLP'13 ICC'15
CVPR '11 EMNLP'14 Fusion'15
EMNLP'12 NIPS '14 EMNLP'15
NIPS '12 SLT '14 + more

Alekh Agarwal K-W Chang Akshay K. John Langford Jordan B-G Alvin Grissom II Mohit Iyyer

I am on the job market!

umiacs.umd.edu/~hhe

He He

- Reasoning with incomplete information is useful for *speed* and *modeling*
- *Imitation learning* can help us build such systems
  - Even when you can't construct a perfect oracle & have incomplete information!
- Wide range of new, interesting problems to work on!
  - *Improve* upon human interpreters?
  - Compete against specific opponents?
  - Distance supervision via structured bandits

Thanks! Questions?