representations of relationships in text (and images) ...and some other stuff

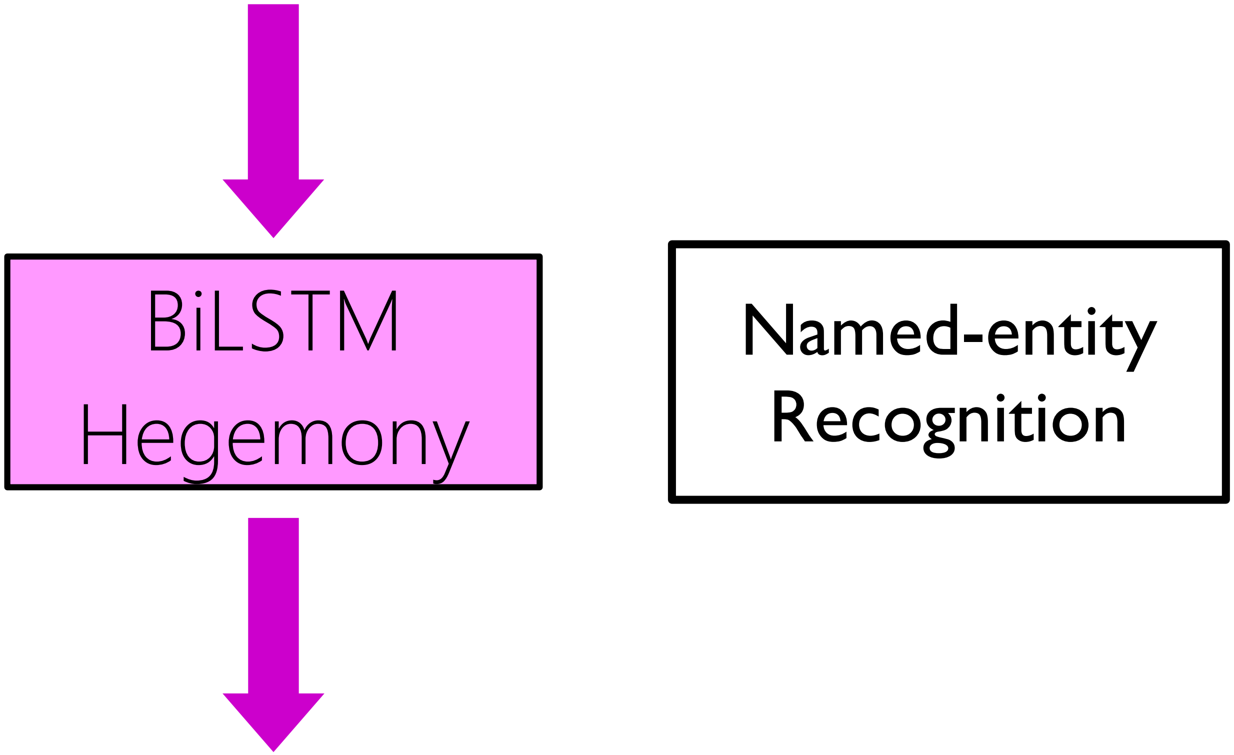
Hal Daumé III | University of Maryland | me@hal3.name | @haldaume3

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?)

NLP via machine learning

Arthur shall strike the blow that sets Lucy free.



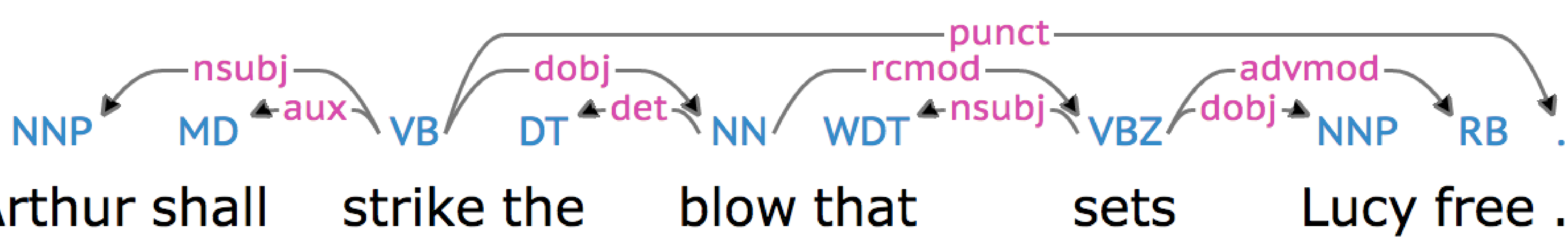
PERSON PERSON
Arthur shall strike the blow that sets Lucy free.

NLP via machine learning

Arthur shall strike the blow that sets Lucy free.

BiLSTM Hegemony

Syntactic Parsing



NLP via machine learning

Arthur shall strike the blow that sets Lucy free.

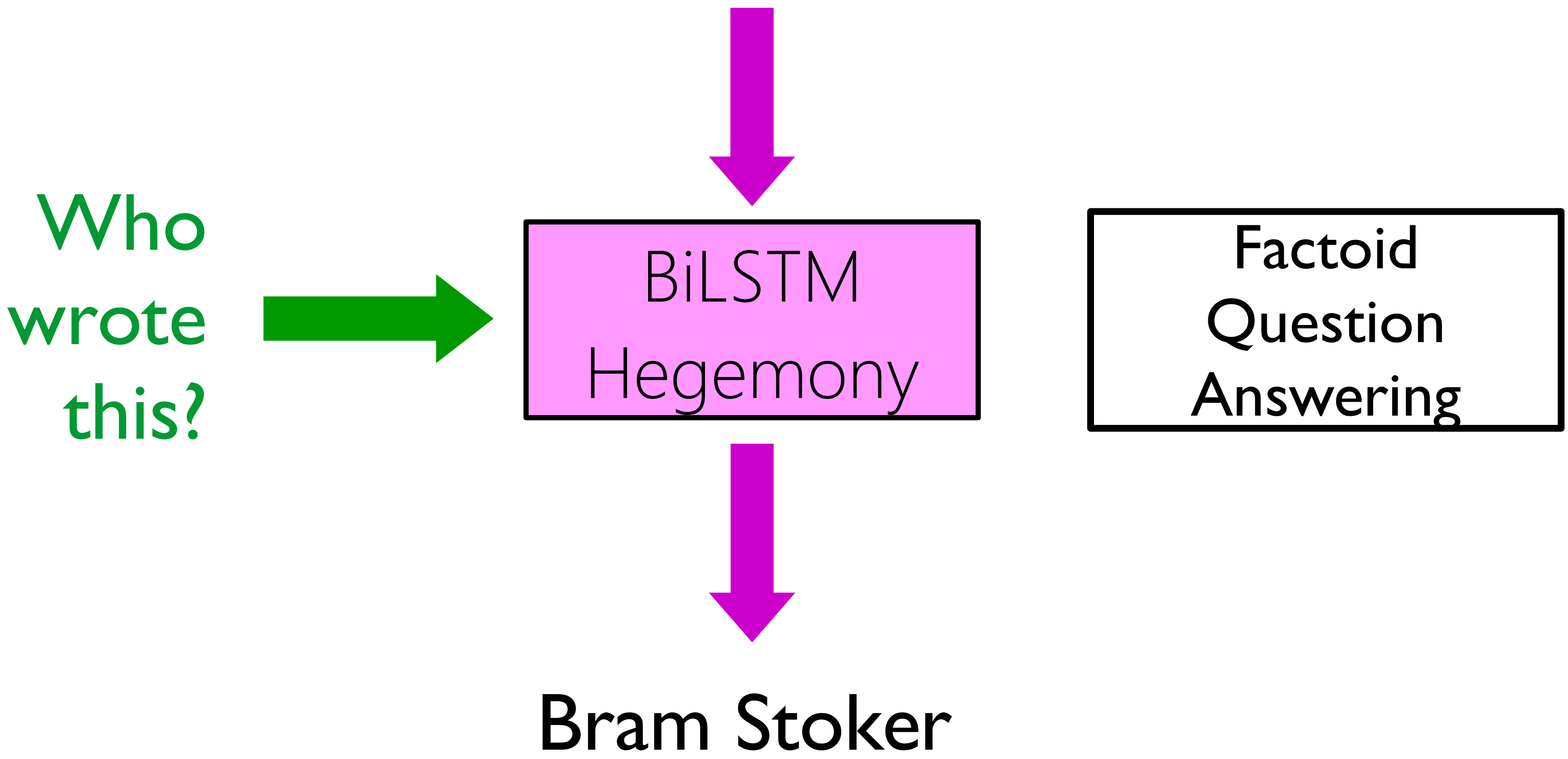
BiLSTM Hegemony

Machine Translation

Es war Arthurs Hand, die Lucy
den Weg zu den Sternen geöffnet hat

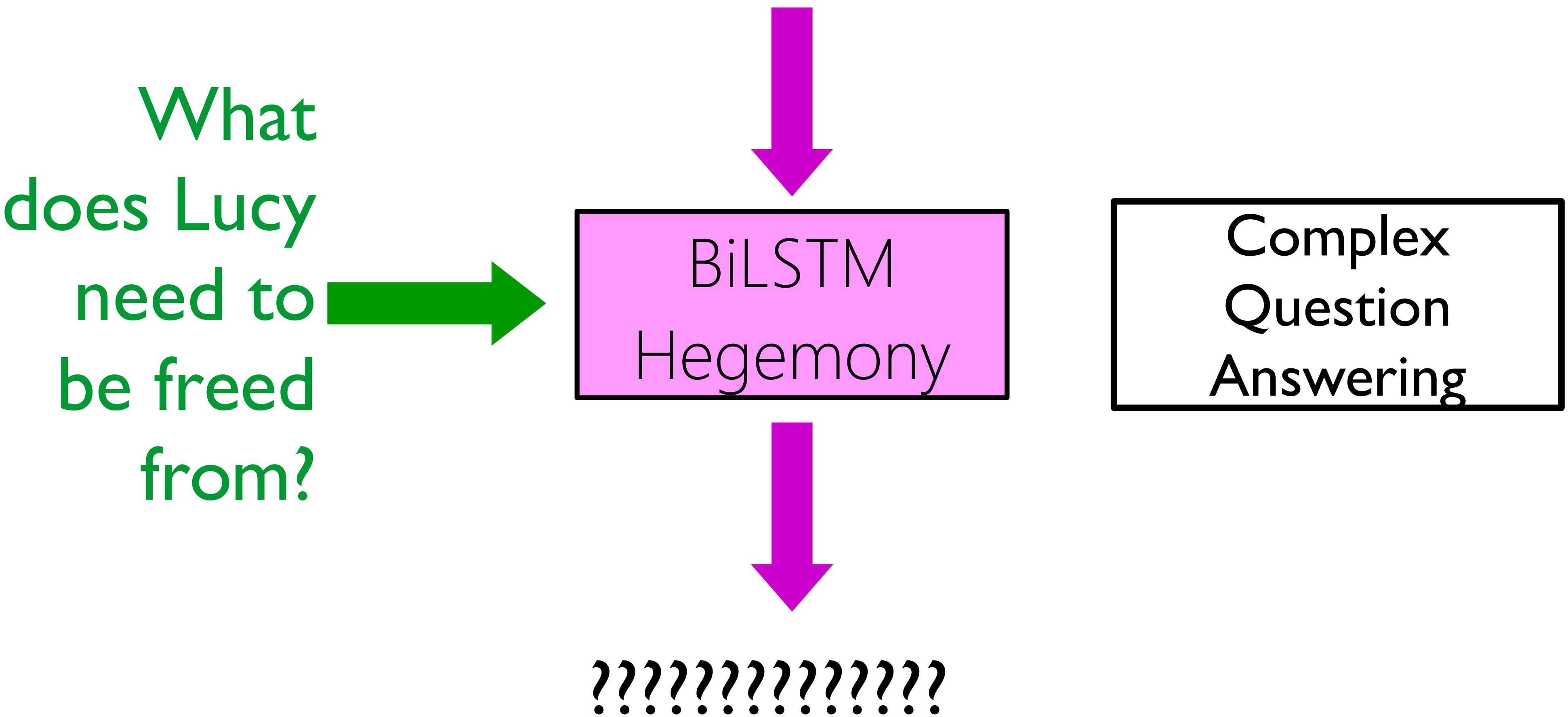
NLP via machine learning

Arthur shall strike the blow that sets Lucy free.



NLP via machine learning

Arthur shall strike the blow that sets Lucy free.



What am I *not* going to talk about?

learning through interaction:

- not so much representation stuff there
- changed my mind
- can see talk linked from my webpage

bag of n-grams + logistic regression:

- my “go-to” representation
- can do better? probably
- do you always? no
- why is averaging enough? regularity?

Narratives

Account of events or experiences

Fictitious or Real

Why Relationships?

Relationships are fundamental for understanding our social behavior

speaking of relationships…

Snigdha Chaturvedi
Postdoc at UIUC

Mohit Iyyer
PhD student at UMD

Existing character-centric approaches model roles

Sometimes narratives are about relationships

* Bamman et. al. ACL 2013, Bamman et. al. ACL 2014, Valls-Vargas et.al. AIIDE 2014

Relationships are not static.

They evolve with the progress of the narrative.

how can we describe a fictional relationship between two characters?

Tom falls in love with **Becky** Thatcher, a new girl in town, and persuades her to get "engaged" to him.

Their romance collapses when she learns that Tom has been "engaged" before—to a girl named Amy Lawrence.

Shortly after being shunned by Becky, Tom ...

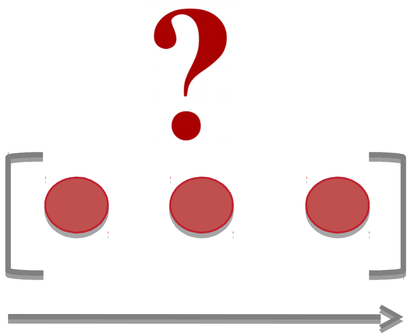
Back in school, Tom gets himself back in Becky's favor after he nobly accepts the blame for a book that she has ripped.

Meanwhile, Tom goes on a picnic to McDougal's Cave with Becky and their classmates.

Problem Statement

Character 1

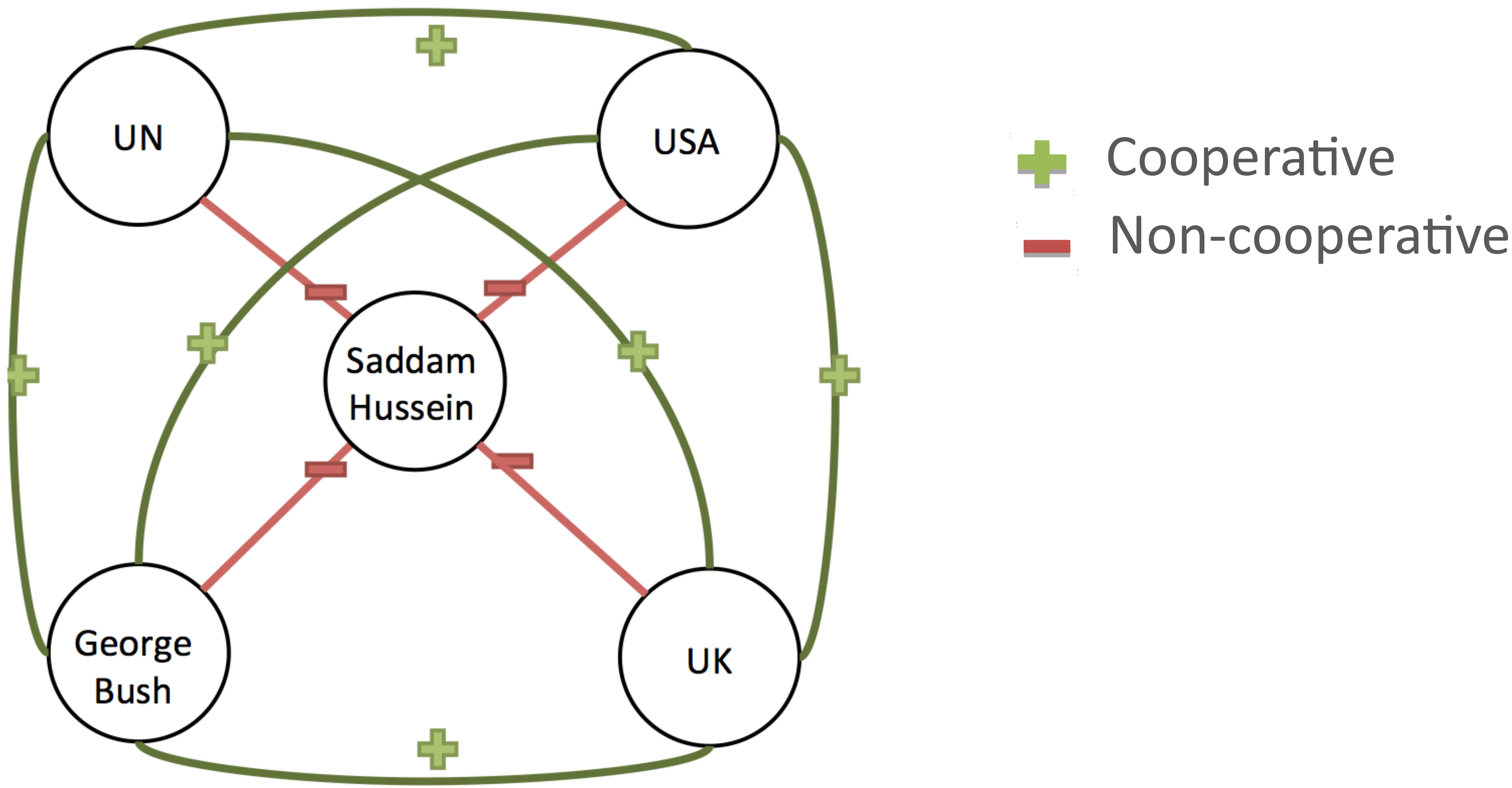
Character 2



Progress of the novel

Sequence of
Relationship states

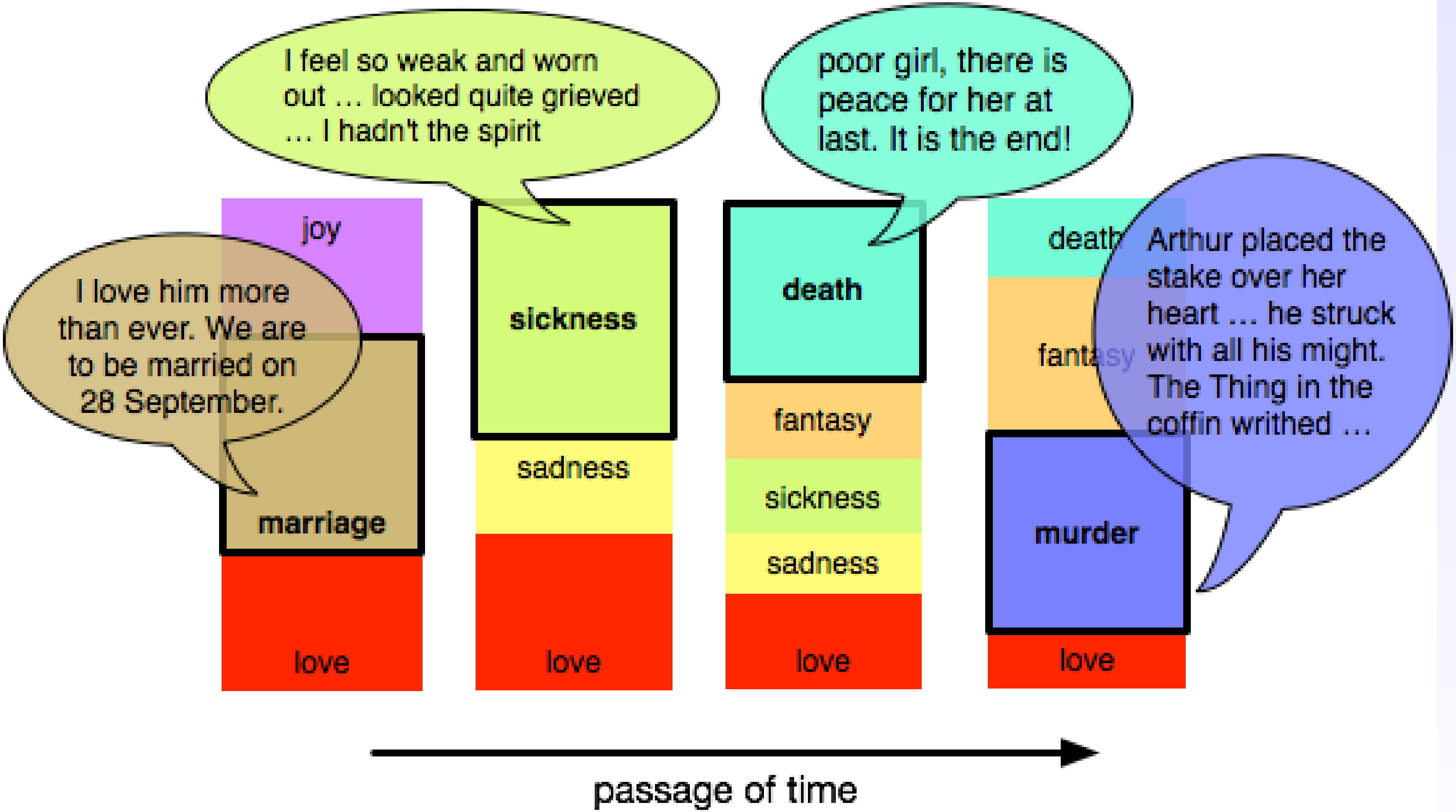
Model not limited to Fictional Narratives



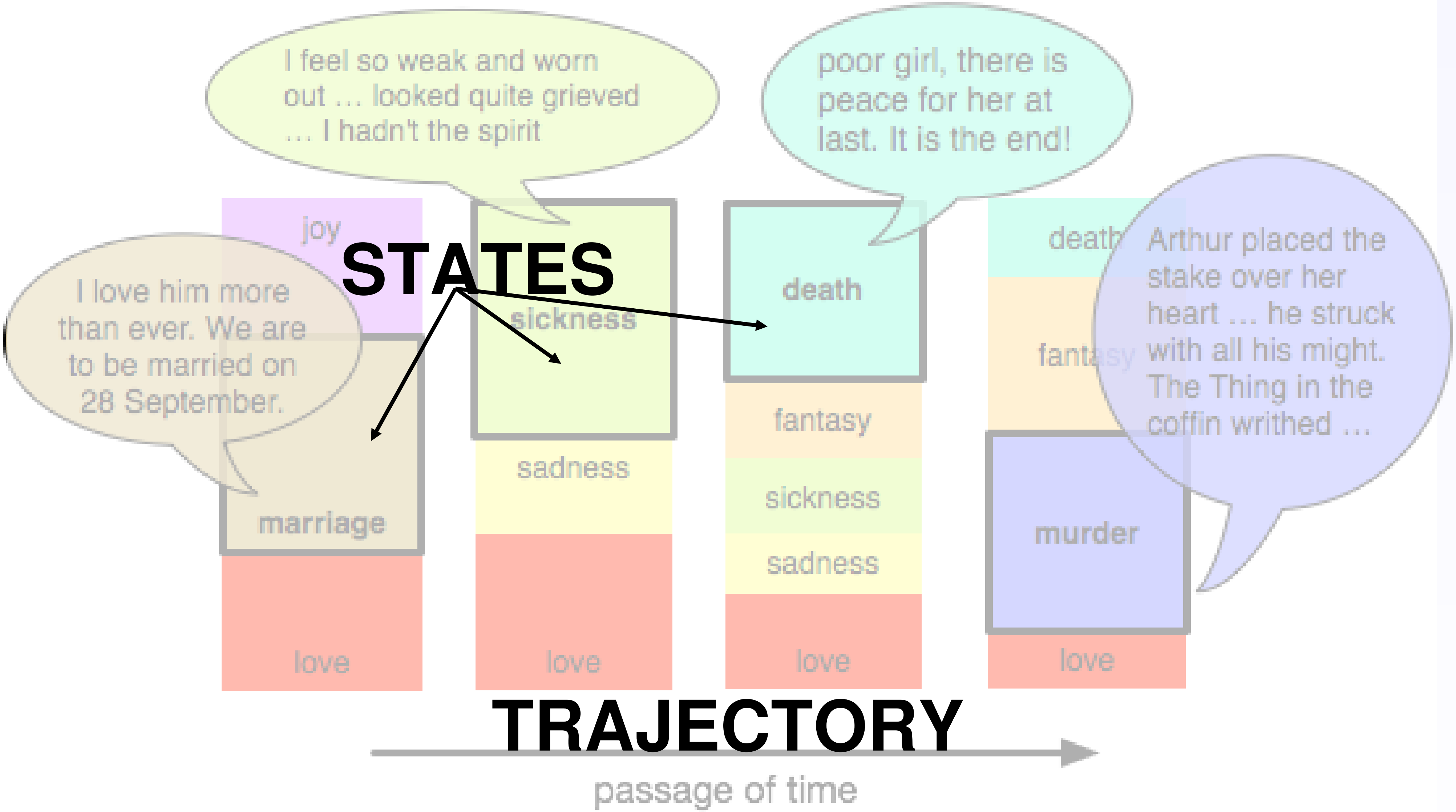
Network Inferred from the Wikipedia article on 2003 Invasion of Iraq

Multiple Facets

Real-world relationships have multiple facets



Arthur and Lucy (*Dracula*)



Arthur and Lucy

(*Dracula*)

why should humanities scholars care?

- "Distant reading" (Moretti, 2005) can help rapidly collect examples of specific relationship types

Do Jane Austen's female and male protagonists have a pattern in their evolving relationships (e.g., mutual disdain followed by romantic love)? (Butler, 1975; Stovel, 1987; Hinant, 2006)
Do certain authors or novels portray relationships of desire more than others? (Polhemus, 1990)
Can we detect positive or negative subtext underlying meals between two characters? (Foster, 2009; Cognard-Black et al., 2014)

a dataset of character interactions

For each pair of characters in a particular book, we extract all **spans** of text that contain mentions to both characters

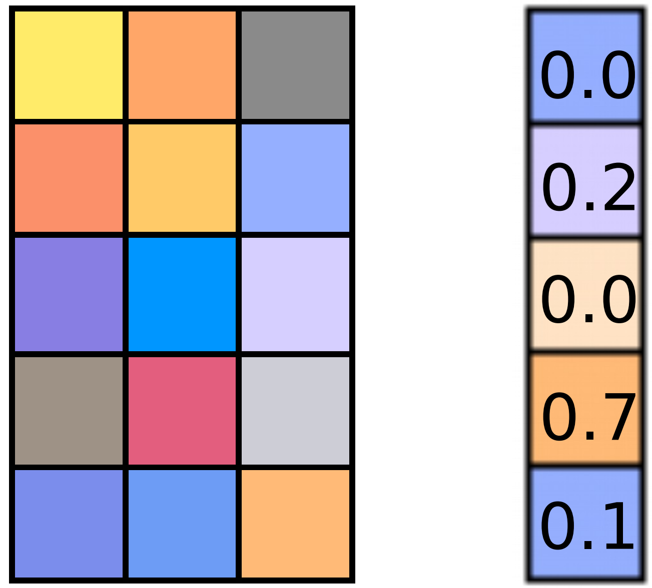
t=0	"If anyone was ever minding his business, it was I," Ignatius breathed. "Please. We must stop. I think I'm going to have a hemorrhage." "Okay." Mrs. Reilly looked at her son's reddening face and realized that he would very happily collapse at her feet just to prove his point."
t=1	"Ignatius belched the gas of a dozen brownies trapped by his valve. "Grant me a little peace…." "You know I appreciate you, babe," Mrs. Reilly sniffed. "Come on and gimme a little goodbye kiss like a good boy."
t=2	Mrs. Reilly looked at her son slyly and asked, "Ignatius, you sure you not a communist?" "Oh, my God!" Ignatius bellowed. "Every day I am subjected to a McCarthyite witchhunt in this crumbling building. No!"

relationship modeling network (RMN)

recurrent autoencoder with dictionary learning

dictionary matrix; each row is a state

violence
sadness
politics
love
suffering

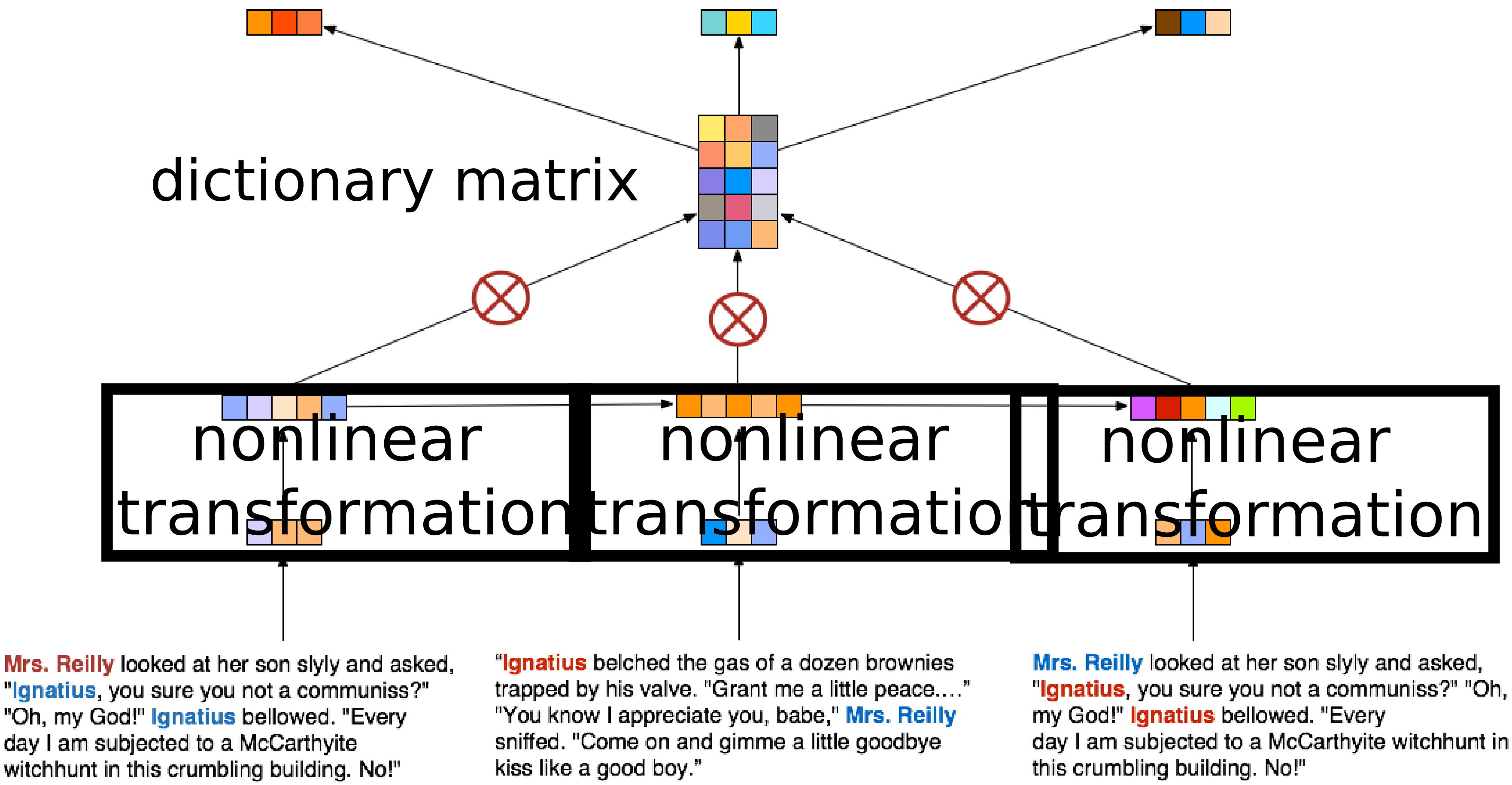


probability distribution over states

“Ignatius belched the gas of a dozen brownies trapped by his valve. "Grant me a little peace….” "You know I appreciate you, babe," Mrs. Reilly sniffed. "Come on and gimme a little goodbye kiss like a good boy.”

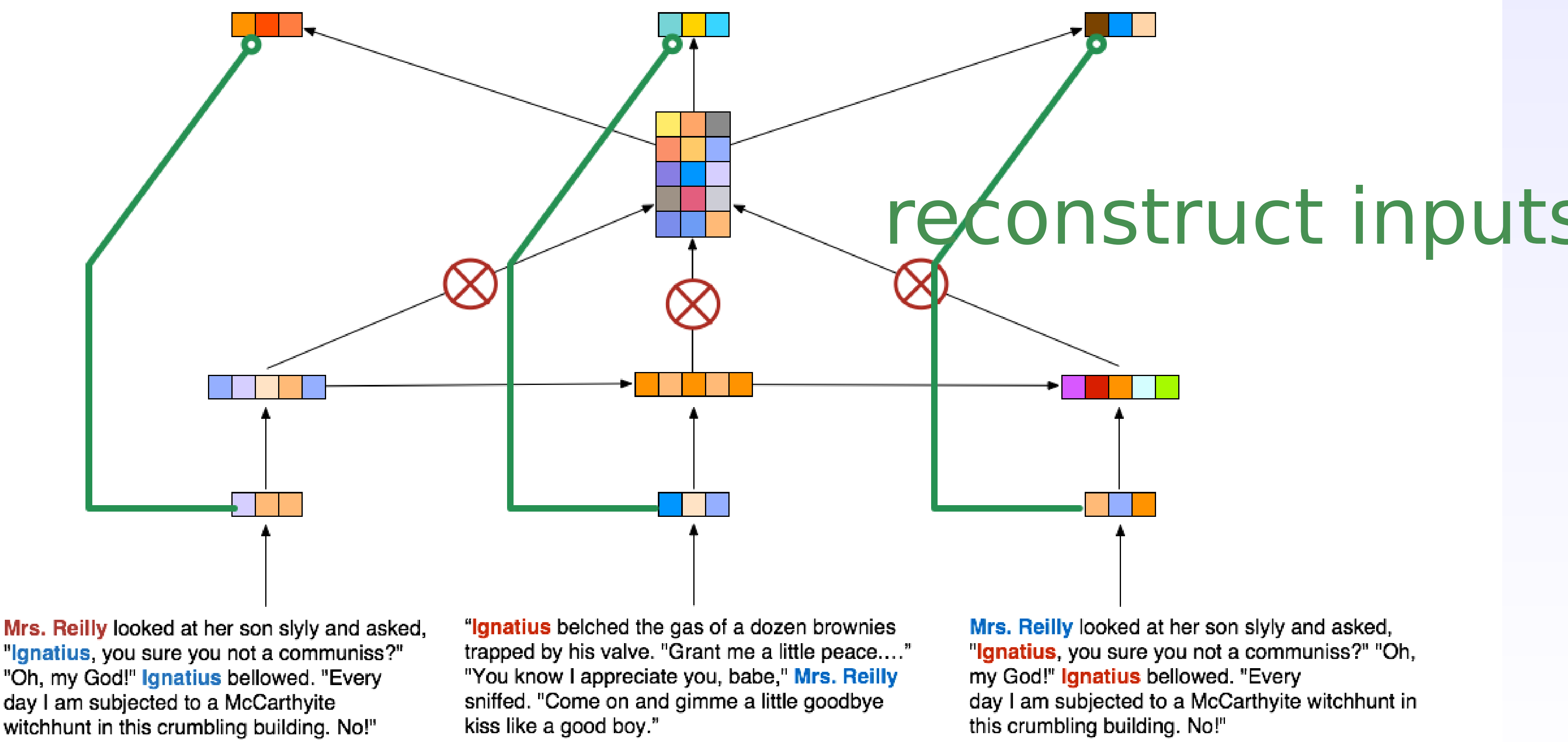
relationship modeling network (RMN)

recurrent autoencoder with dictionary learning

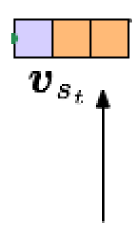


relationship modeling network (RMN)

recurrent autoencoder with dictionary learning



1. word embedding average



Mrs. Reilly looked at her son slyly and asked, "Ignatius, you sure you not a communiss?" "Oh, my God!" Ignatius bellowed. "Every day I am subjected to a McCarthyite witchhunt in this crumbling building. No!"

most coherent relationship states

<u>relationship modeling network</u>

outdoors: outdoors trail trails hillside grassy slopes
sadness: regretful rueful pity pained despondent
education: teaching graduate year teacher attended
love: love delightful happiness enjoyed enjoyable
murder: autopsy arrested homicide murdered

<u>baseline: hidden topic markov model (HTMM)</u>

crime: blood knife pain legs steal
food: kitchen mouth glass food bread
violence: sword shot blood shouted swung
boats: ship boat captain deck crew
outdoors: stone rock path darkness desert

most coherent relationship states

relationship modeling network

outdoors: outdoors trail trails hillside grassy slopes

sadness: regretful rueful pity pained despondent

education: teaching graduate year teacher attended

love: love delightful happiness enjoyed enjoyable

murder: autopsy arrested homicide murdered

baseline: hidden topic markov model (HTMM)

crime: blood knife pain legs steal

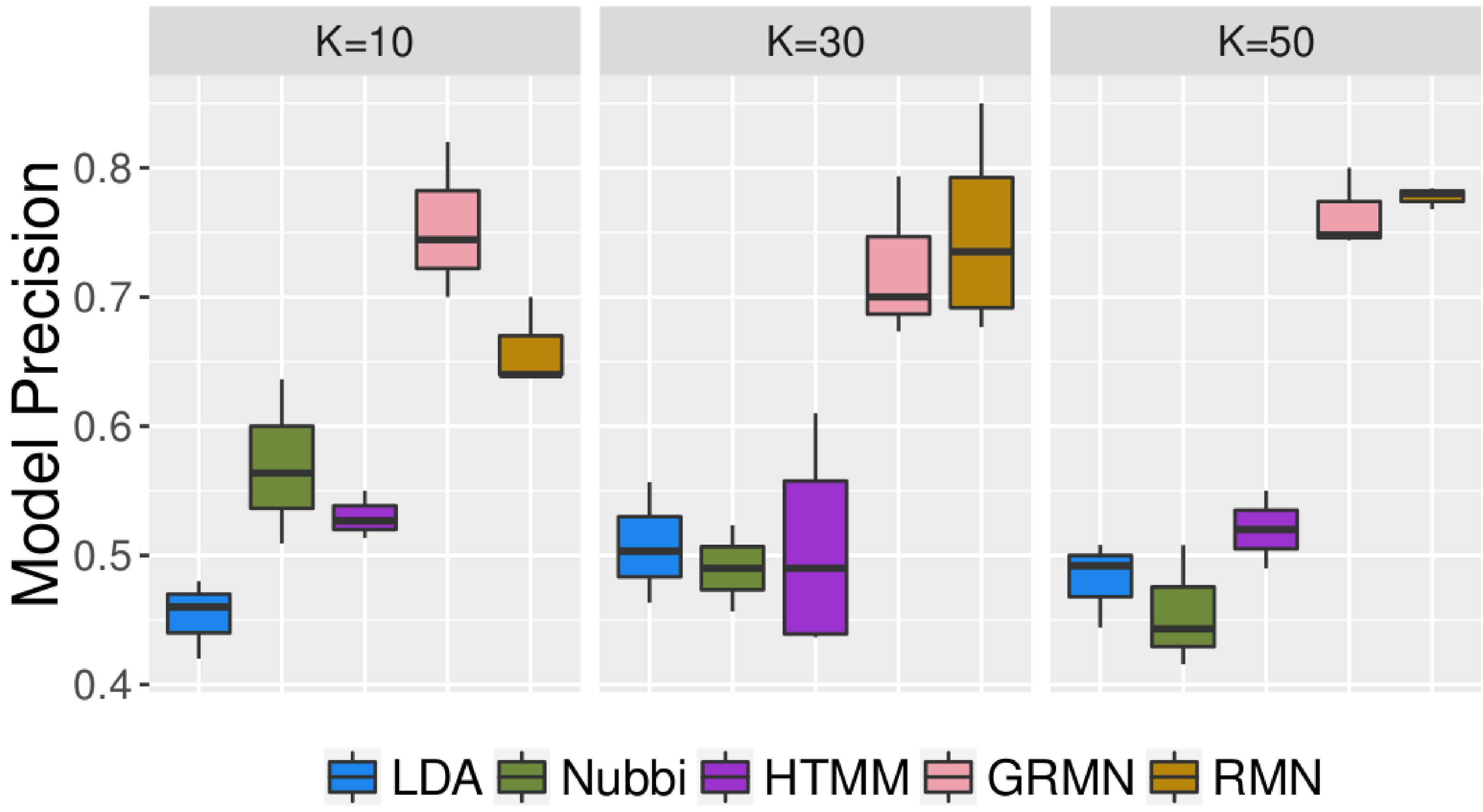
food: kitchen mouth glass food bread

violence: sword shot blood shouted swung

boats: ship boat captain deck crew

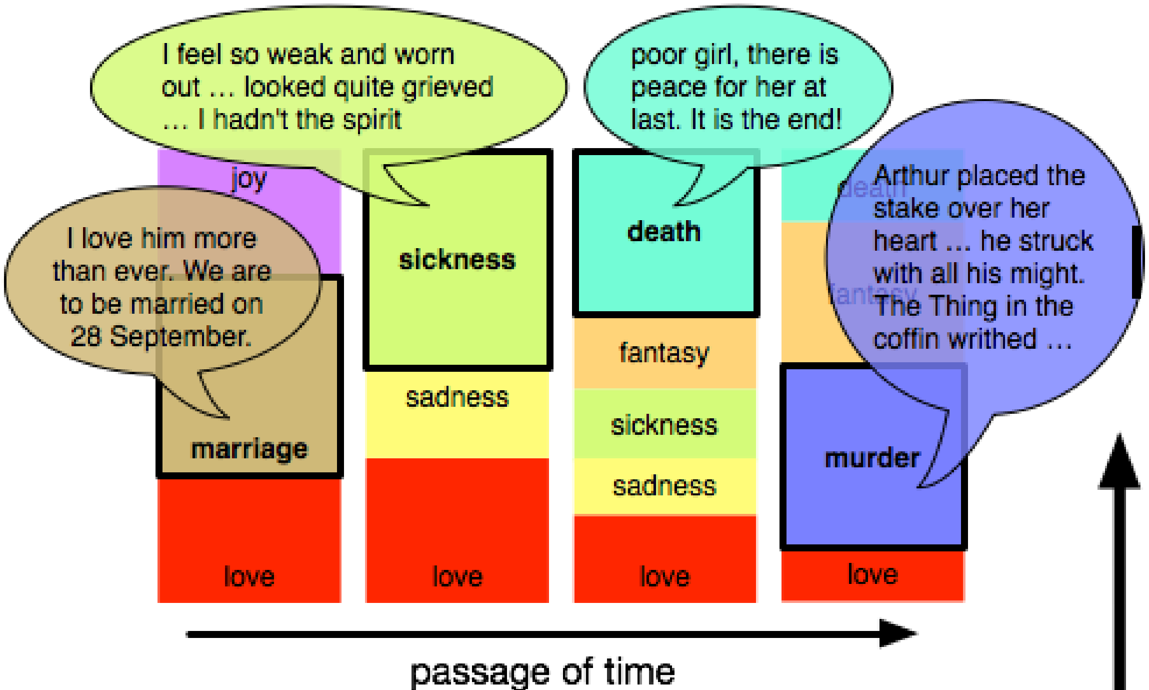
outdoors: stone rock path darkness desert

word intrusion task

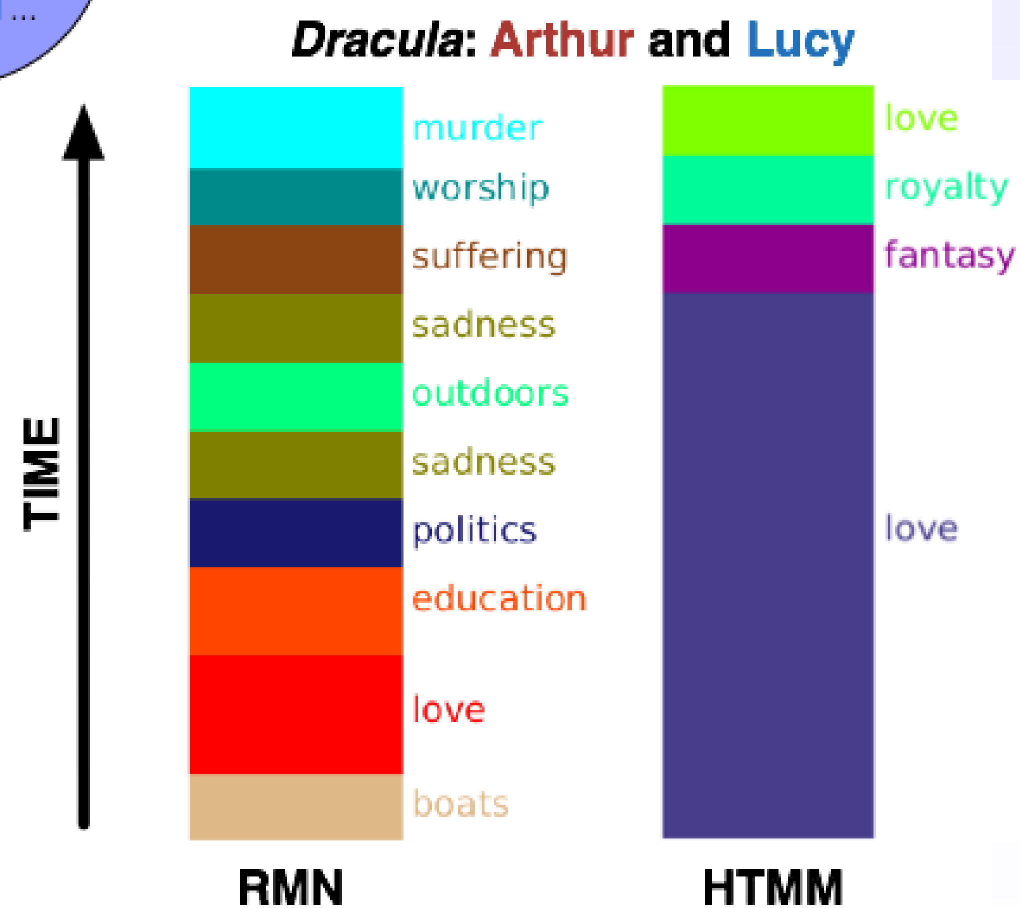


Arthur and Lucy "ground-truth":

marriage -> sickness -> death -> murder



learned trajectories:



RMN contributions

humanities:

- learns global relationship states from corpus of novels
- interpretable visualizations of relationship trajectories

machine learning:

- novel combination of dictionary learning and neural networks
- unsupervised training promotes model interpretability

commonsense inferences

Lex Luthor's limousine
came late today.

The driver had overslept.

closure: the process by which we connect panels together

what is exploding, and why?

a dataset of COMICS to study closure

- 4,000 comic books from the Digital Comics Museum www.digitalcomicsmuseum.com
  - books are from the "Golden Age of Comics" (1938-1954), in the public domain due to copyright expiration

1.2 million panels with 2.5 million textboxes

cloze tasks for testing closure

The Simons workshop has been interesting so far.

I hope the rest of the week is just as __________

A. boring
B. awful
C. compelling

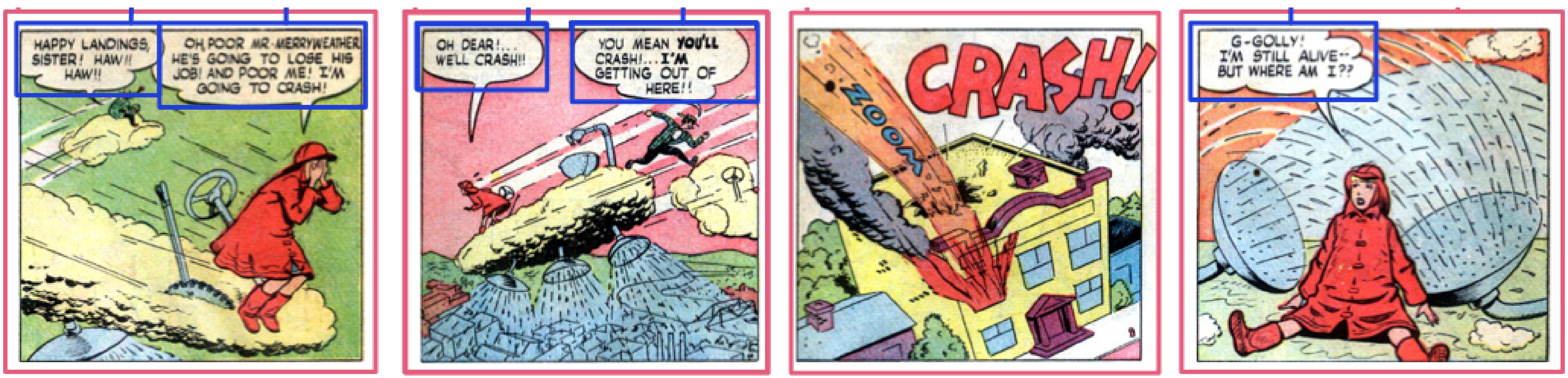
task: predict dialogue in a panel given previous panels as context

task: predict dialogue in a panel given previous panels as context

A. Alice! I've been looking all over for you!

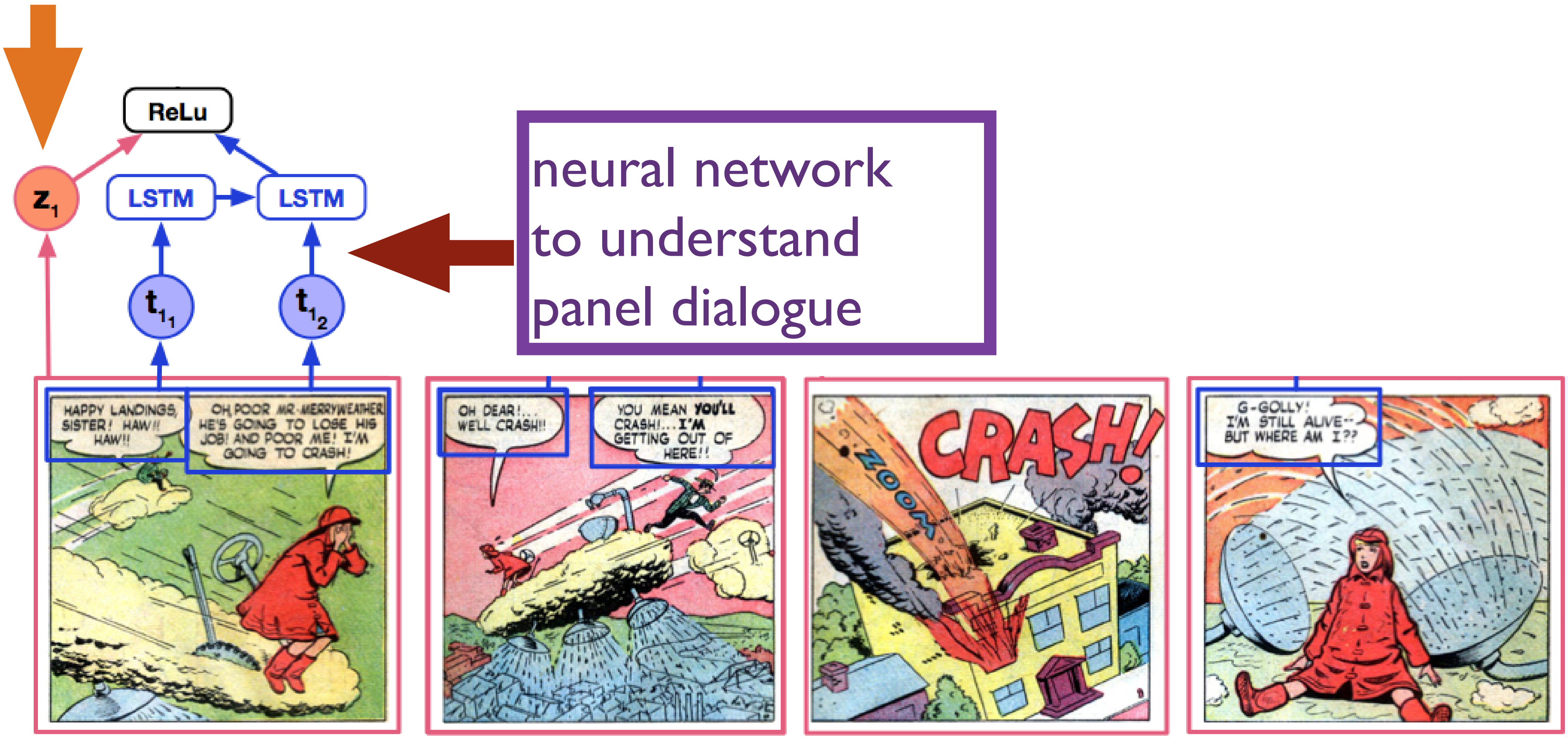
B. Hiya kid! All alone?

C. G-Golly! I'm still alive… but where am I??

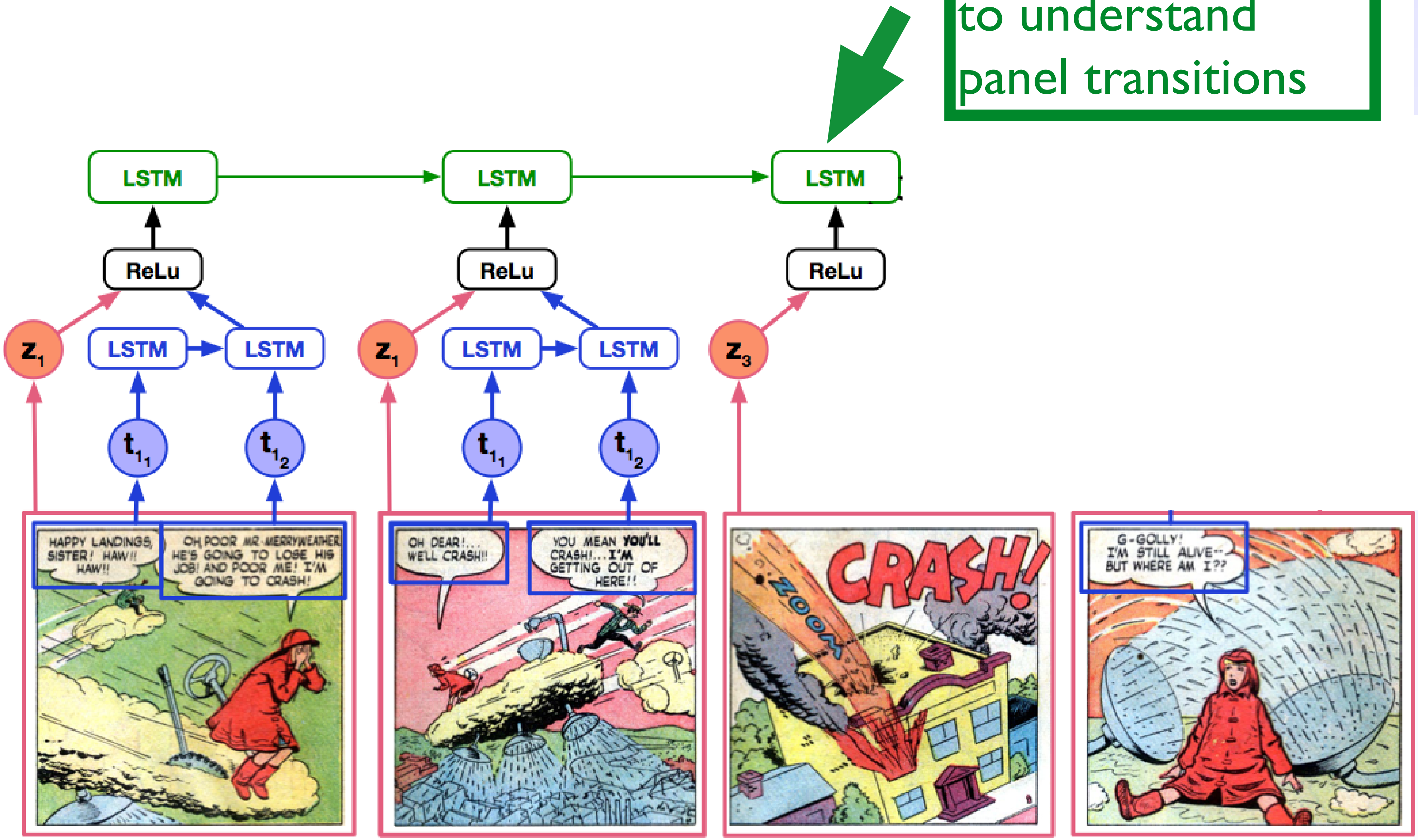

HAPPY LANDINGS, SISTER! HAW!! HAW!!
OH, POOR MR. MERRYWEATHER! HE'S GOING TO LOSE HIS JOB! AND POOR ME! I'M GOING TO CRASH!
OH DEAR!... WELL CRASH!!
YOU MEAN YOU'LL CRASH!... I'M GETTING OUT OF HERE!!
ZOOM
CRASH!
G-GOLLY! I'M STILL ALIVE-- BUT WHERE AM I??

neural network
to understand
panel artwork

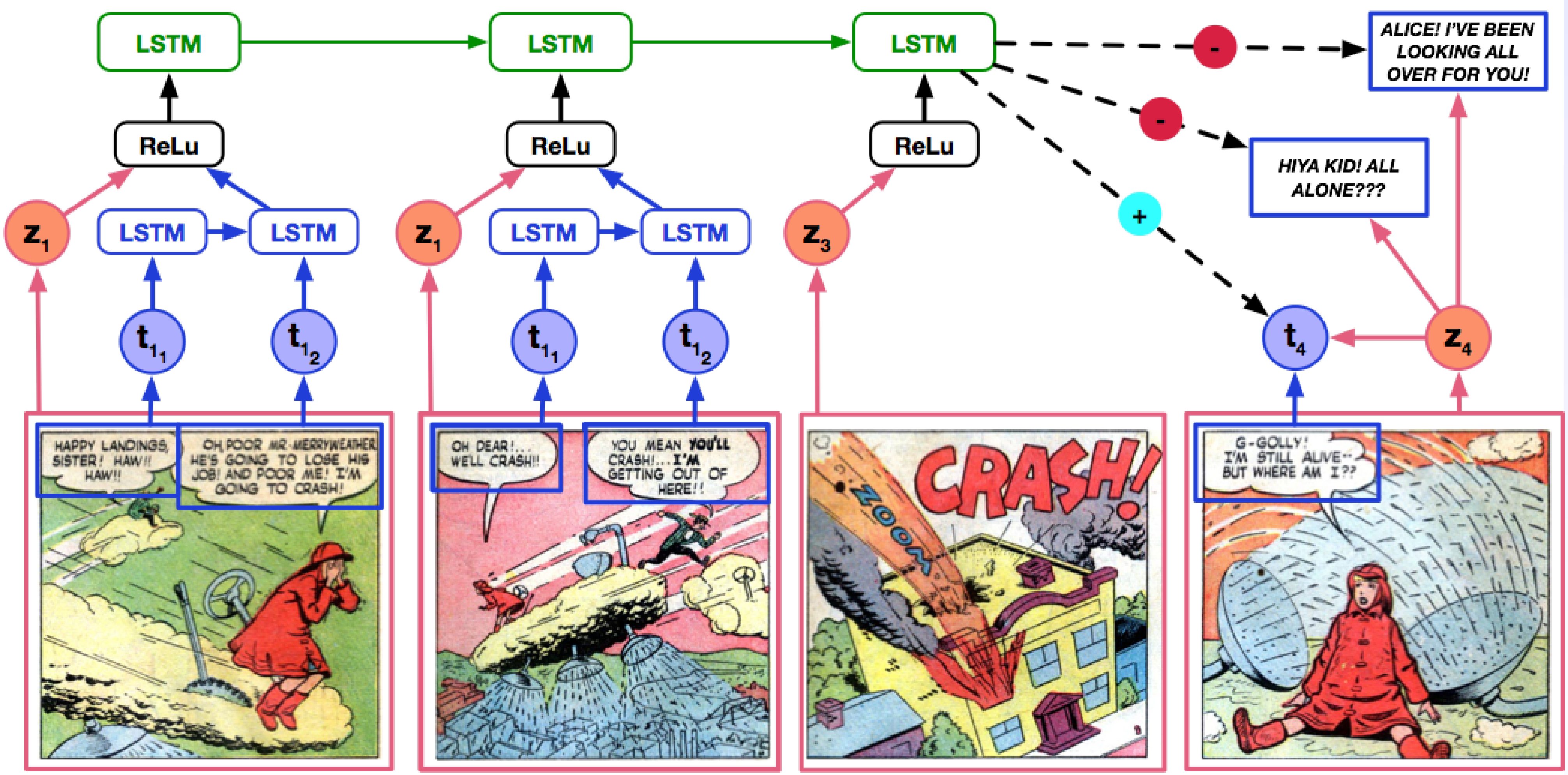
neural network
to understand
panel dialogue



neural network
to understand
panel transitions



dialogue candidates scored against network output



closure is hard!

Model	Cloze accuracy
Random	33
Image-only	48.7
Text-only	51.9
Image-text •No context	59.6
Image-text •Full context	63.4
Human	**84**

COMICS recap

- new dataset & tasks that test a computer's ability to make commonsense inferences
- deep learning can learn some degree of closure
- however, our best models lag behind humans despite huge amount of data

inductive bias in networks or data?

Simultaneous machine interpretation

He He 何河

Alvin Grissom II

Why simultaneous interpretation is hard

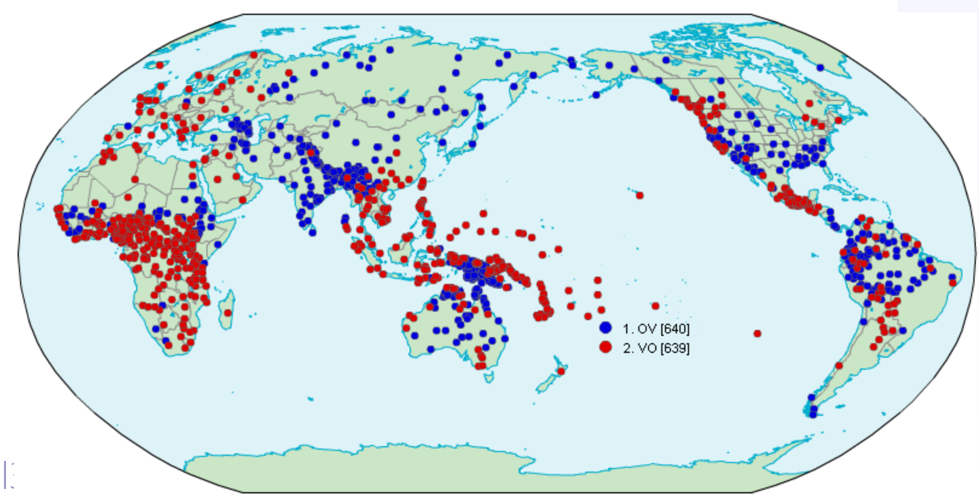
- Human languages have vastly different word orders
  - About half are OV, the other half are VO
  - This comes with a lot more baggage than just verb-final

man-TOP store-LOC go-PAST

the man went to the store

food-OBJ buy-DESIRE man-TOP store-LOC go-PAST

the man who wanted to buy food went to the store



General diffs of Interp vs Batch

- Inversion
  - Segmentation into multiple sentences
  - Passivization of single sentence
- Word generalization
  - (lower retrieval time)
- Summarization and omission
  - (to catch up)

Example (gen + segment)

(S) この日本語の待遇表現の特徴ですが英語から日本語へ直訳しただけでは表現できないといった特徴があります

(T) One of the characteristics of **honorific** Japanese is that it can not be **adequately** expressed when using a direct translation from English to Japanese.

(I) Now let me talk about the characteristic of the Japanese **polite** expressions. <segment/> And such expressions can not be expressed **enough** just by translating directly.

Example (gen + passivize)

(S) 以上のお話をまとめますと自然な発話というものを扱うことができる音声対話の方法ということを考案しました.

(T) In summary, we have **devised** a way for voice interaction systems to handle natural speech.

(I) And this is the summary of what I have so far stated. The spontaneous speech can be dealt with by the speech dialog method **<segment/>** and that method was **proposed**.

Example (gen + summarize)

(S) で三番目の特徴としてはですねえ出来る限り自然な日本語の話言葉とてその出力をするといったような特徴があります .

(T) Its third **characteristic** is that its output is, as much as possible, in the natural language of spoken ((Japanese)).

(I) And the third **feature** is that the translation could be produced in a very natural spoken language.

How can we integrate this information into a neural machine translation system?

rewriting the *training data*

我々は 政府の 構造 や 組織を 変更 すべき だ
We-TOP government-GEN structure and composition-ACC change should COP

我々は || 政府の || 構造や組織を || 変更すべきだ
the government's structure and composition should be changed by us

We || should change the structure and composition of the government

Batch ⟶ Monotone

… of the government *the government's …*

change … *… be changed*

Requirement of missing constituents is postponed

Transformation rules

- Passivization

  We should change the structure and composition of the government

  The structure and composition of the governmentshould be changed by us

- Genitive reordering

  the structure and composition of the government

  the government 's the structure and composition

- Quotative verbs

  They announced that the president will restructure the division

  The president will restructure the division, they announced

  - *know, realize, observe, doubt, deny*

- *that* clause

  It is important to remain watchful

  To remain watchful is important

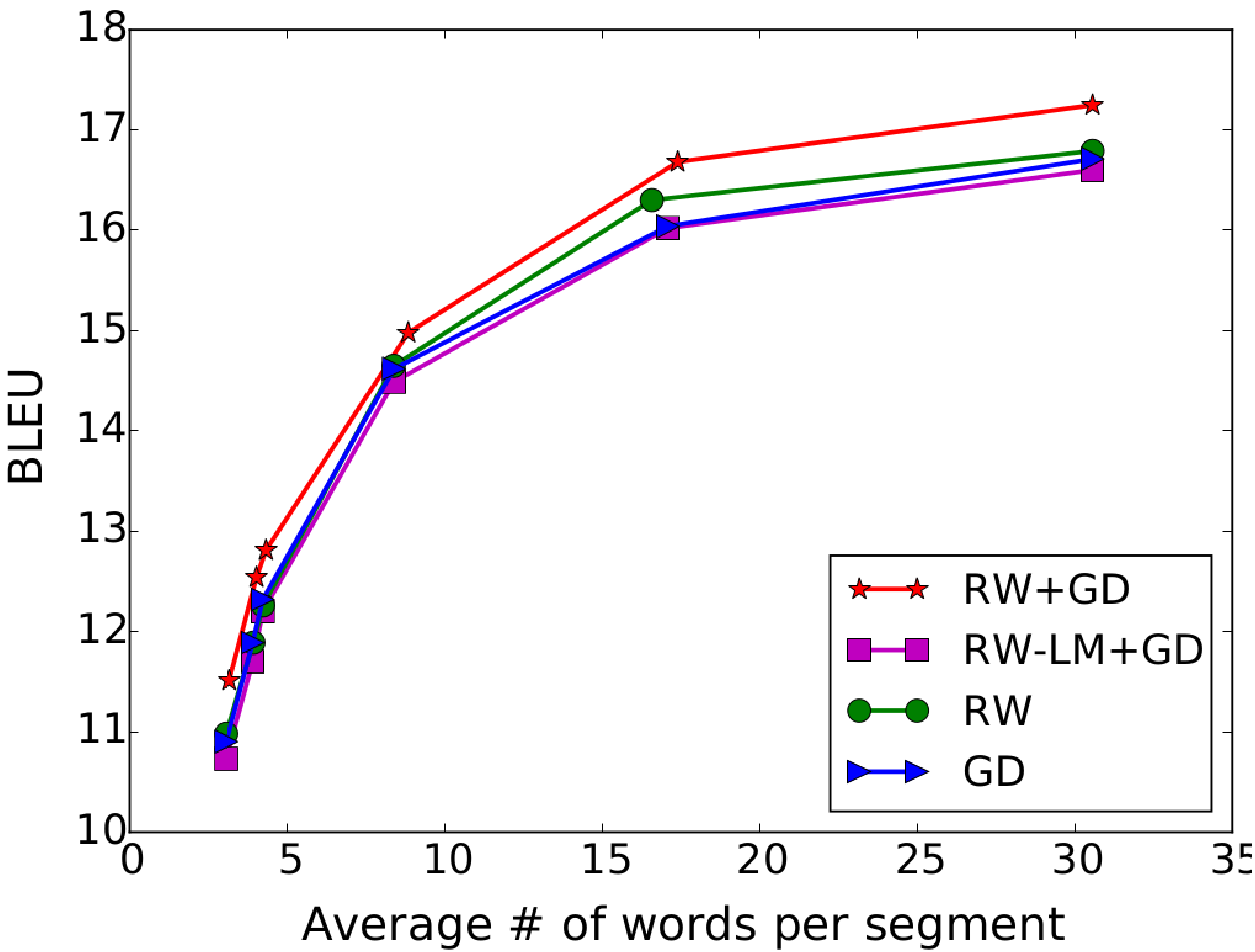
- Conjunction clause

  We should march because winter is coming

  Winter is coming ; because of this , we should march

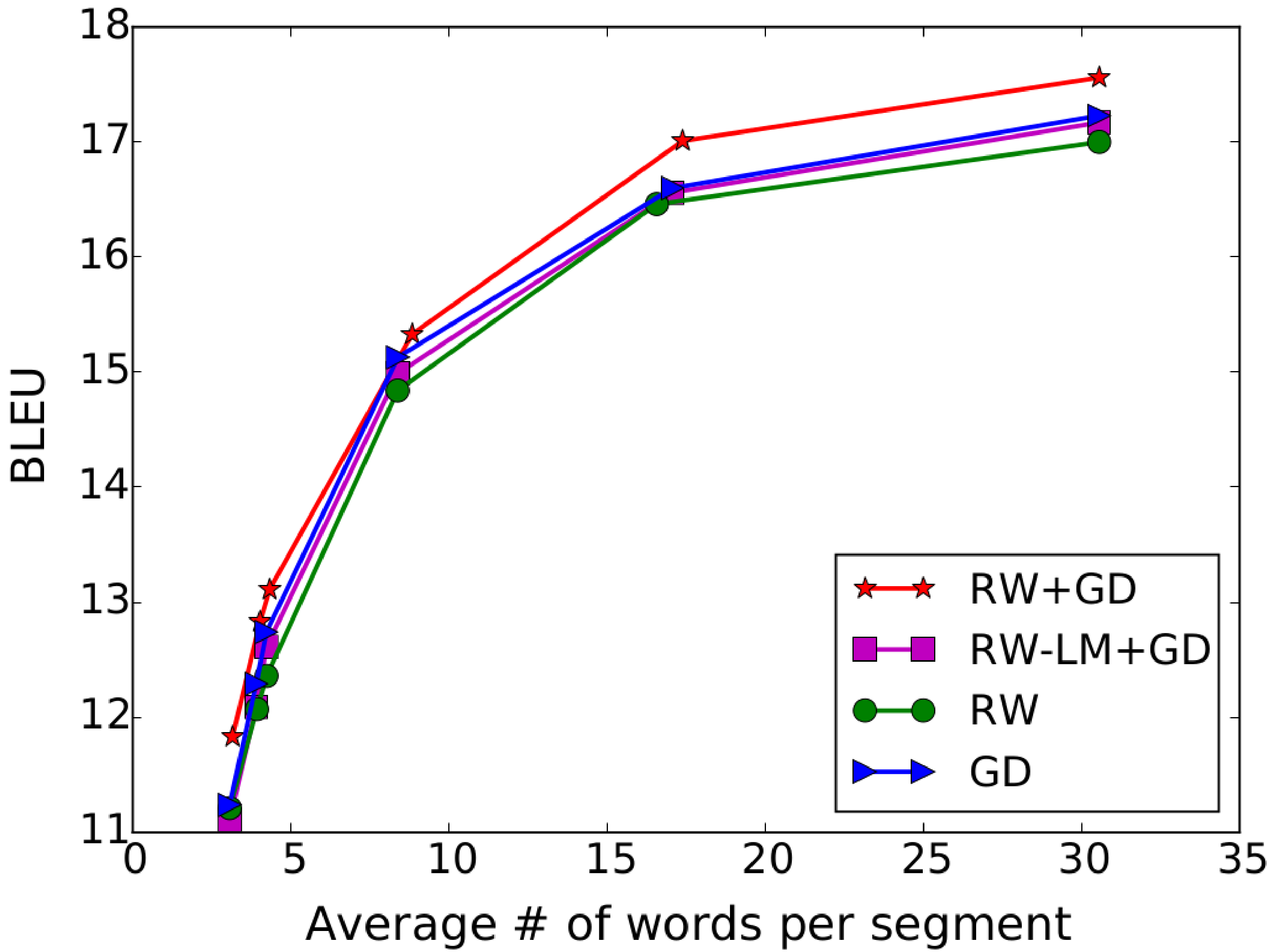
  - *despite, even though, although*

Accuracy at different delay levels



(c) BLEU w.r.t. rewritten ref

Accuracy at different delay levels



(a) BLEU w.r.t. gold ref

learning with invariants

- MNIST folks have been doing this for ages
- Often much easier to talk about data invariants than trying to construct "better" hypothesis classes
- Added benefit: can use highly optimized code
- Can try to understand though lens of covariate shift:
  - More data → good
  - Not-quite-right-data → bad
- Are there other ways to understand this?

Discussion

- Dictionary learning embedded in neural network → interpretability
- Moving toward evaluating unsupervised learning via actual use cases
- Not quite there on complex common-sense reasoning tasks

- Invariants on *outputs* as a mechanism for additional supervision
- Not sure how to analyze

Snigdha Chaturvedi

Mohit Iyyer

He He

Alvin Grissom II