LIAAD at SemDeep-5 Challenge

Word-in-Context (WiC)

Daniel Loureiro, Alípio Jorge

SemDeep-5@IJCAI – Macau, 12 August 2019

U.PORTO
FACULDADE DE CIÊNCIAS UNIVERSIDADE DO PORTO
LIAAD
INESCTEC LABORATÓRIO ASSOCIADO

Sense Embeddings

Exploiting the latest Neural Language Models (NLMs) for sense-level representation learning.

Sense Embeddings

Exploiting the latest Neural Language Models (NLMs) for sense-level representation learning.

- Beat SOTA for English Word Sense Disambiguation (WSD).
- Full WordNet in NLM-space (+100K common sense concepts).
- Concept-level analysis of NLMs. [ACL 2019 – LMMS Paper]

Related Work

Related Work

[Iacobacci et al. (2016)]
[Zhong and Ng (2010)]

Bag-of-Features Classifiers

(SVM)

[Luo et al. (2018b)]
[Luo et al. (2018a)]
[Vial et al. (2018)]
[Raganato et al. (2017)]

Deep Sequence Classifiers

(BiLSTM)

[Loureiro and Jorge (2019)]
[Peters et al. (2018)]
[Melamud et al. (2016)]
[Yuan et al. (2016)]

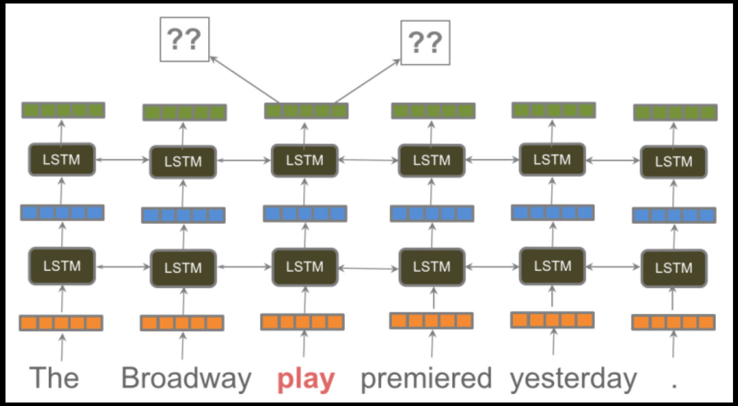
Sense-level Representations

(k-NN)
(over NLM reprs.)

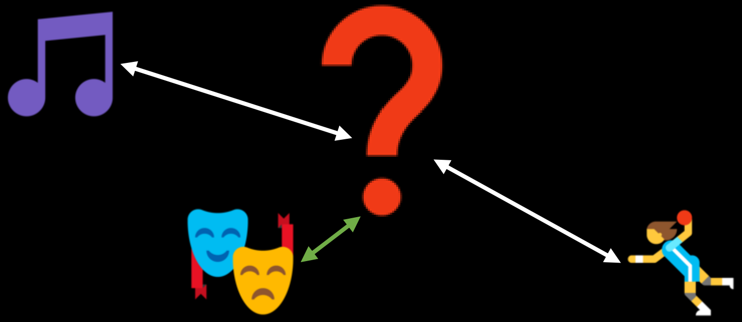
Contextual k-NN

Matching Contextual Word Embeddings:

- Produce Sense Embeddings from NLMs (averaging).
- Sense embs. can be compared with contextual embs.
- Disambiguation = Nearest Neighbour search (1-NN).
- Annotations have limited coverage (16% of WordNet).
- Promising, but early attempts.



[Ruder (2018)]



Our Approach

Our Approach

- Expand the k-NN approach to full-coverage of WordNet.

Our Approach

- Expand the k-NN approach to full-coverage of WordNet.
- Matching senses becomes trivial, no MFS fallbacks needed.

Our Approach

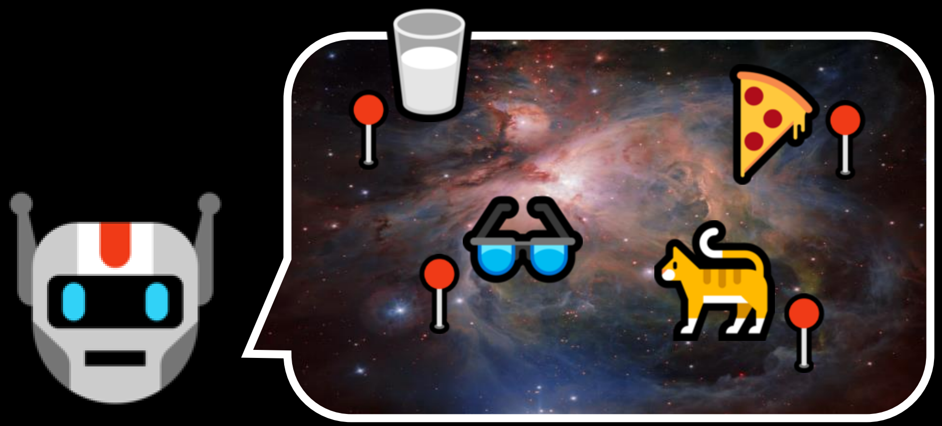
- Expand the k-NN approach to full-coverage of WordNet.
- Matching senses becomes trivial, no MFS fallbacks needed.
- Full-set of sense embeddings in NLM-space is useful beyond WSD.

Our Approach

- Expand the k-NN approach to full-coverage of WordNet.
- Matching senses becomes trivial, no MFS fallbacks needed.
- Full-set of sense embeddings in NLM-space is useful beyond WSD.

Our Approach

- Expand the k-NN approach to full-coverage of WordNet.
- Matching senses becomes trivial, no MFS fallbacks needed.
- Full-set of sense embeddings in NLM-space is useful beyond WSD.



Challenges

Challenges

- Overcome very limited sense annotations (covers 16% senses).
- Infer missing senses correctly so that task performance improves.
- Rely only on sense embeddings, no lemma or POS features.*

Bootstrap →	Propagate →	Enrich →	Reinforce*
Annotated Dataset	WordNet Ontology	WordNet Glosses	Morphological Embeddings

*Covered on our ACL 2019 Paper

Bootstrapping Sense Embeddings

Can your insurance company aid you in reducing administrative costs ?

Would it be feasible to limit the menu in order to reduce feeding costs ?

Bootstrapping Sense Embeddings

insurance_company%1:14:00::
aid%2:41:00::
reduce%2:30:00::
administrative%3:01:00::
cost%1:21:00::

Can your <u>insurance company</u> <u>aid</u> you in <u>reducing</u> <u>administrative</u> <u>costs</u> ?

Would it be <u>feasible</u> to <u>limit</u> the <u>menu</u> in order to <u>reduce</u> <u>feeding</u> <u>costs</u> ?

feasible%5:00:00:possible:00
limit%2:30:00::
menu%1:10:00::
reduce%2:30:00::
feeding%1:04:01::
cost%1:21:00::

Bootstrapping Sense Embeddings

insurance_company%1:14:00::
aid%2:41:00::
reduce%2:30:00::
administrative%3:01:00::
cost%1:21:00::

Can your insurance company aid you in reducing administrative costs ?

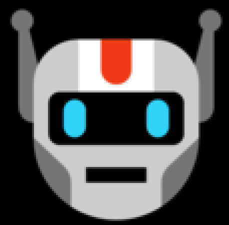
Would it be feasible to limit the menu in order to reduce feeding costs ?

feasible%5:00:00:possible:00
limit%2:30:00::
menu%1:10:00::
reduce%2:30:00::
feeding%1:04:01::
cost%1:21:00::

Bootstrapping Sense Embeddings

insurance_company%1:14:00:: aid%2:41:00:: reduce%2:30:00:: administrative%3:01:00:: cost%1:21:00::

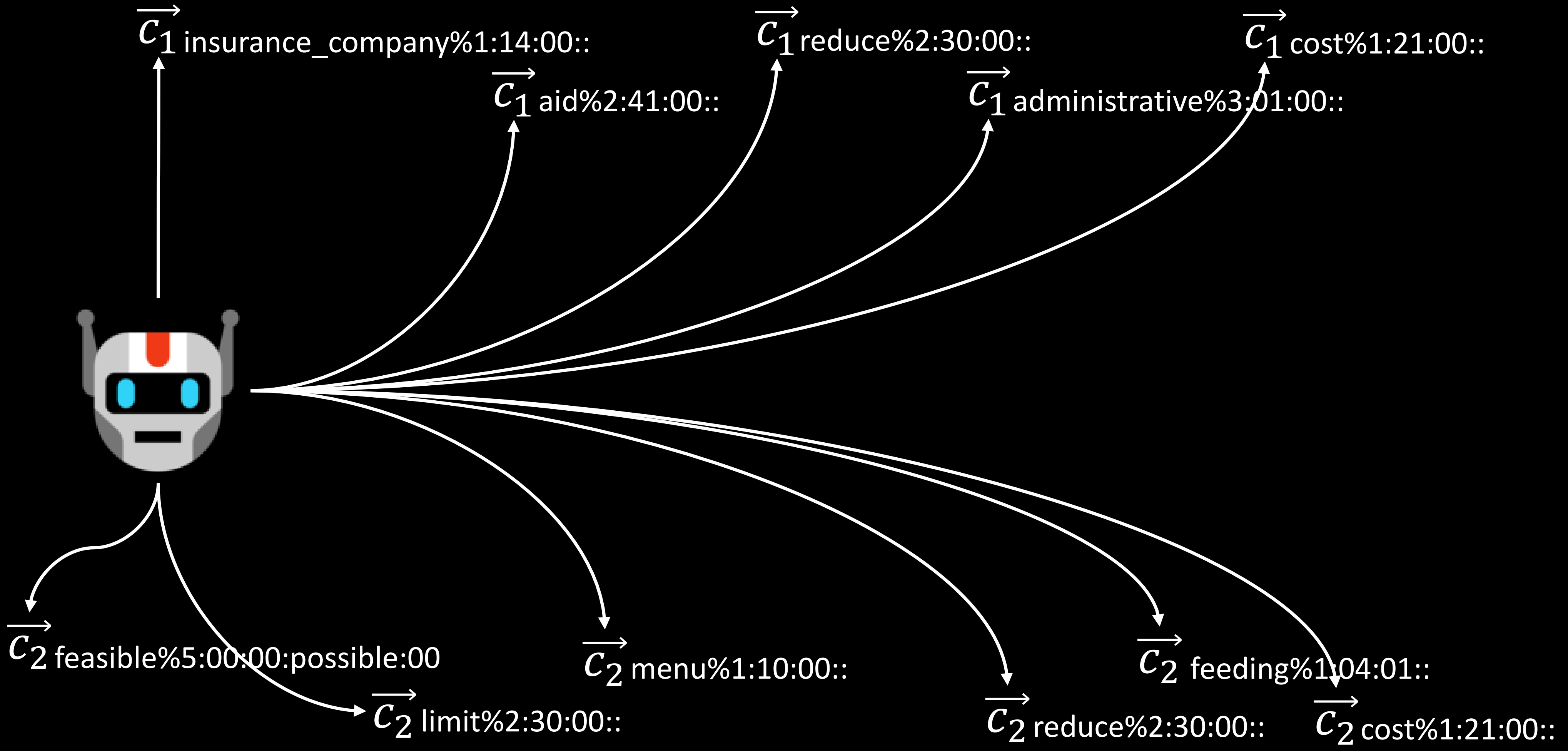
Can your insurance company aid you in reducing administrative costs ?



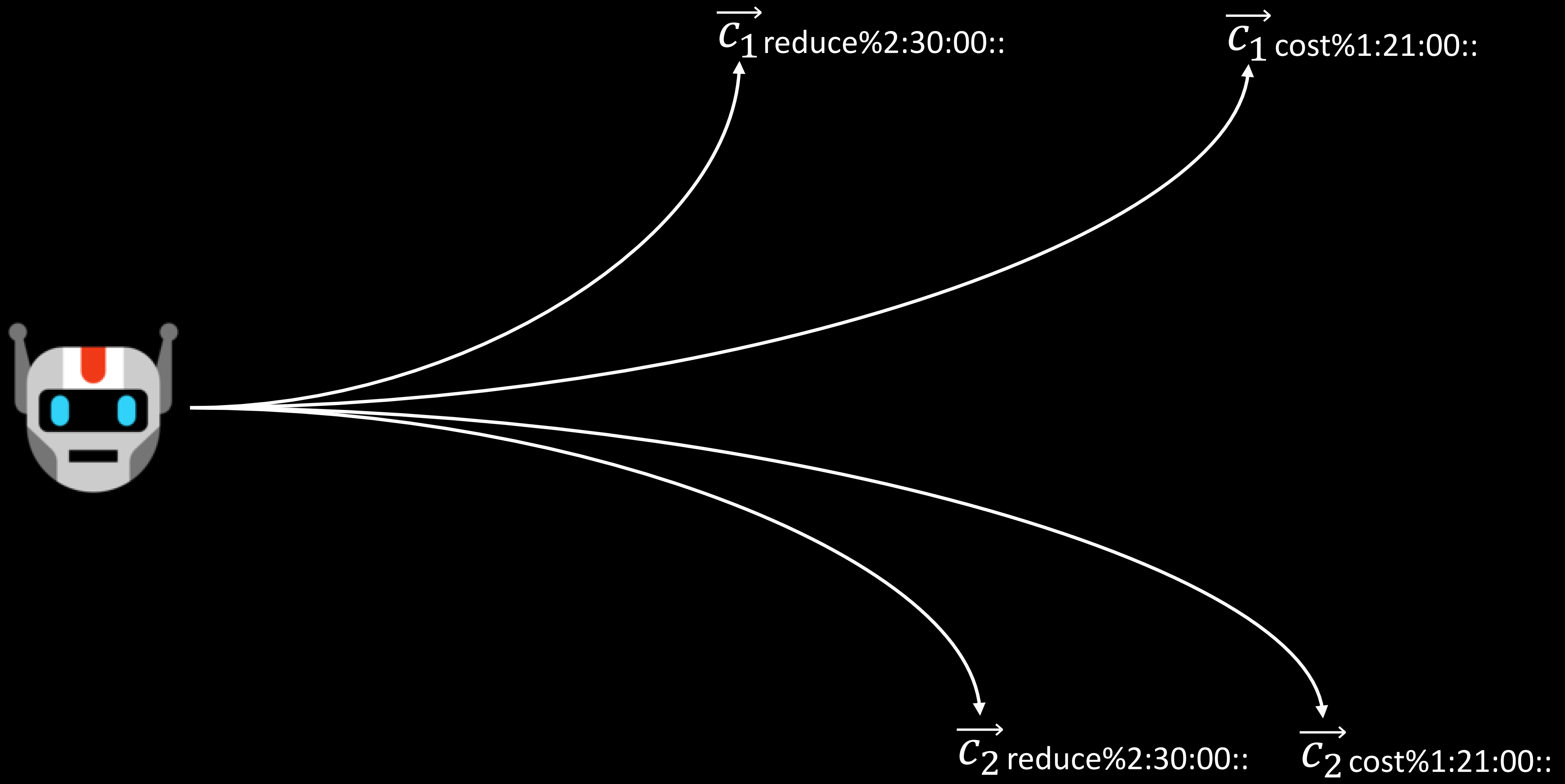
Would it be feasible to limit the menu in order to reduce feeding costs ?

feasible%5:00:00:possible:00 limit%2:30:00:: menu%1:10:00:: reduce%2:30:00:: feeding%1:04:01:: cost%1:21:00::

Bootstrapping Sense Embeddings



Bootstrapping Sense Embeddings



Bootstrapping Sense Embeddings

$$\vec{v}_{\text{reduce\%2:30:00::}} = \frac{\vec{c_1}_{\text{reduce\%2:30:00::}} + \vec{c_2}_{\text{reduce\%2:30:00::}} + \ldots + \vec{c_n}_{\text{reduce\%2:30:00::}}}{n}$$

$$\vec{v}_{\text{cost\%1:21:00::}} = \frac{\vec{c_1}_{\text{cost\%1:21:00::}} + \vec{c_2}_{\text{cost\%1:21:00::}} + \ldots + \vec{c_n}_{\text{cost\%1:21:00::}}}{n}$$

Bootstrapping Sense Embeddings

$$\vec{v}_{\text{reduce\%2:30:00::}} = \frac{\vec{c_1}_{\text{reduce\%2:30:00::}} + \vec{c_2}_{\text{reduce\%2:30:00::}} + \ldots + \vec{c_n}_{\text{reduce\%2:30:00::}}}{n}$$

$$\vec{v}_{\text{cost\%1:21:00::}} = \frac{\vec{c_1}_{\text{cost\%1:21:00::}} + \vec{c_2}_{\text{cost\%1:21:00::}} + \ldots + \vec{c_n}_{\text{cost\%1:21:00::}}}{n}$$

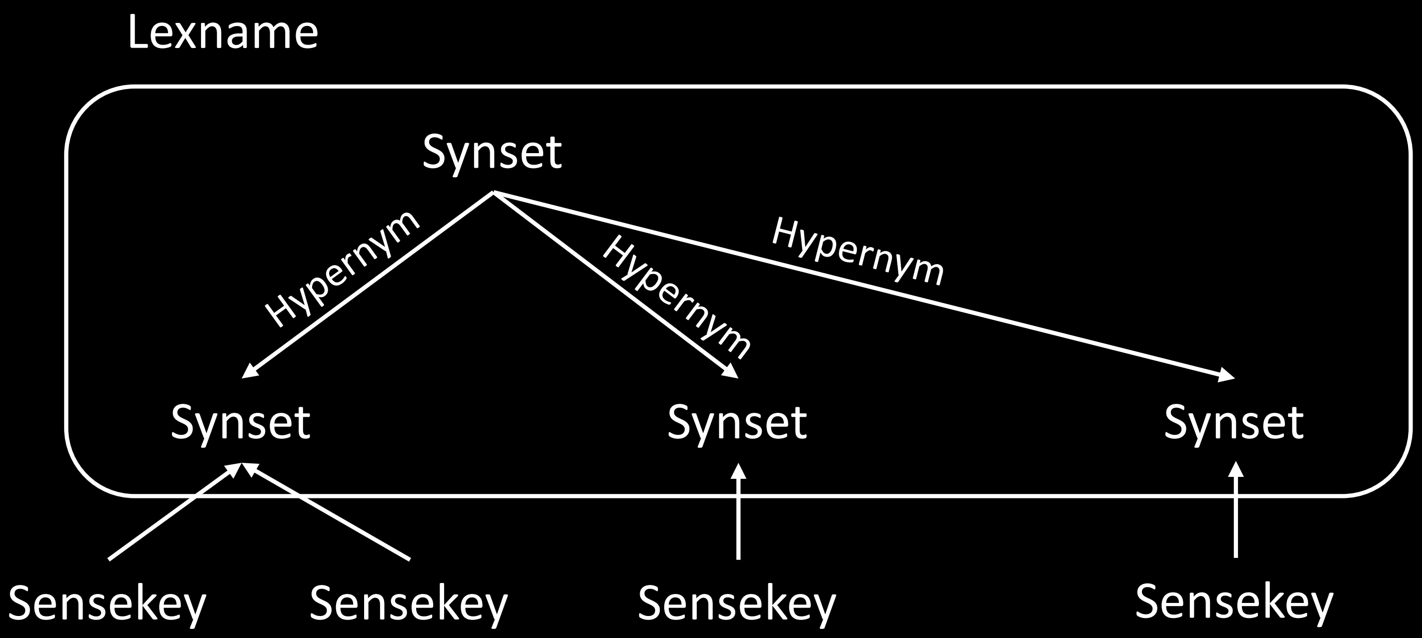
Outcome: 33,360 sense embeddings (16% coverage)

Propagating Sense Embeddings

WordNet’s units, synsets, represent concepts at different levels.

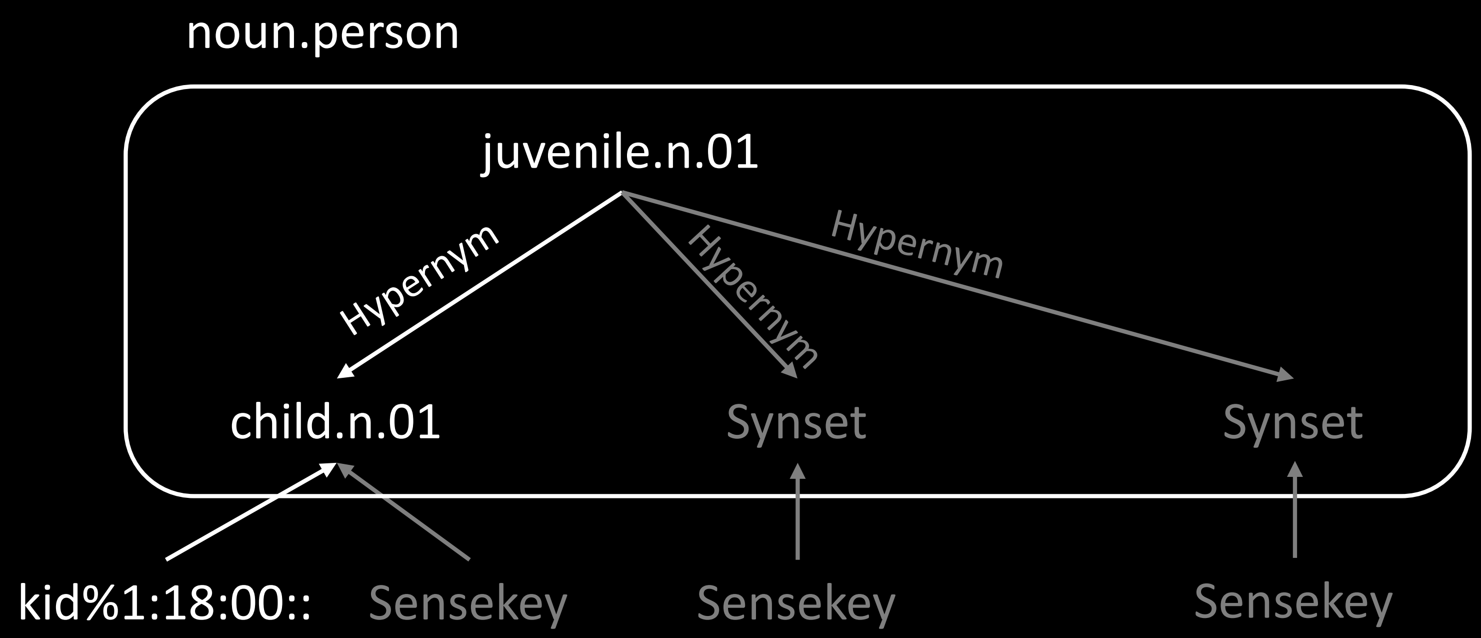
Propagating Sense Embeddings

WordNet's units, synsets, represent concepts at different levels.



Propagating Sense Embeddings

WordNet's units, synsets, represent concepts at different levels.



Propagating Sense Embeddings

burger%1:13:00::

hotdog%1:18:00::

hamburger%1:13:01::

sandwich%1:13:00::

wrap%1:13:00::

potato_chip%1:13:00::

Propagating Sense Embeddings

burger%1:13:00::

hotdog%1:18:00::

hamburger%1:13:01::

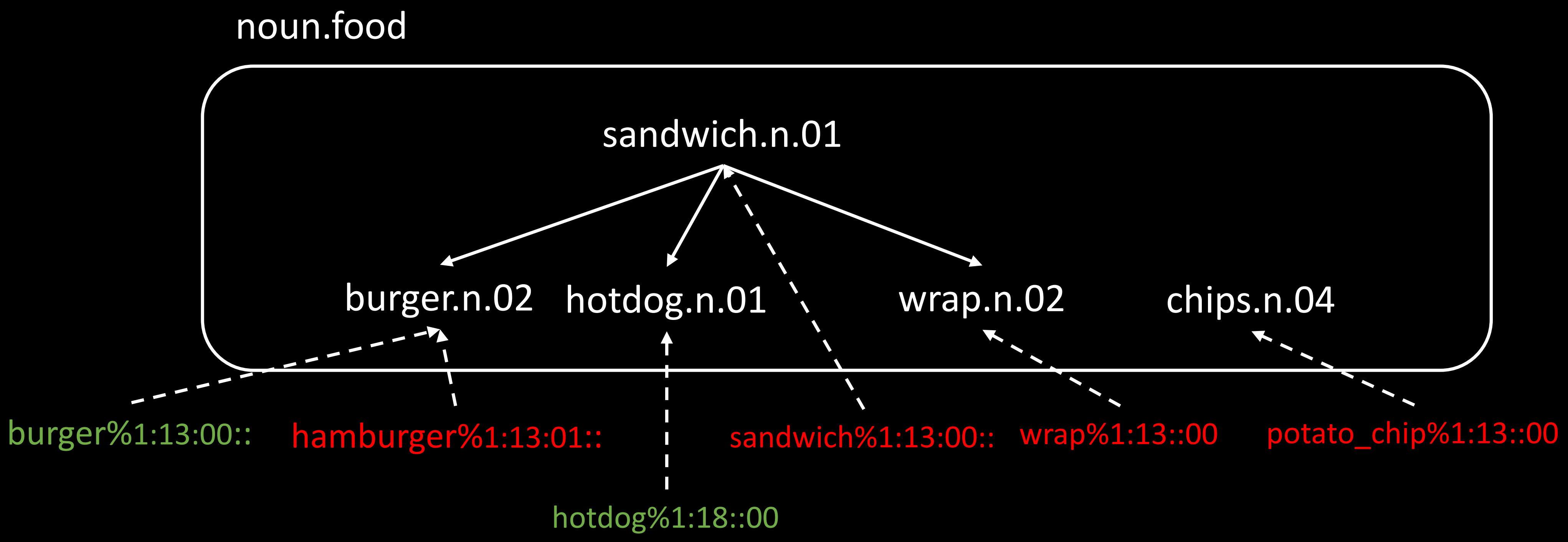
sandwich%1:13:00::

wrap%1:13:00::

potato_chip%1:13:00::

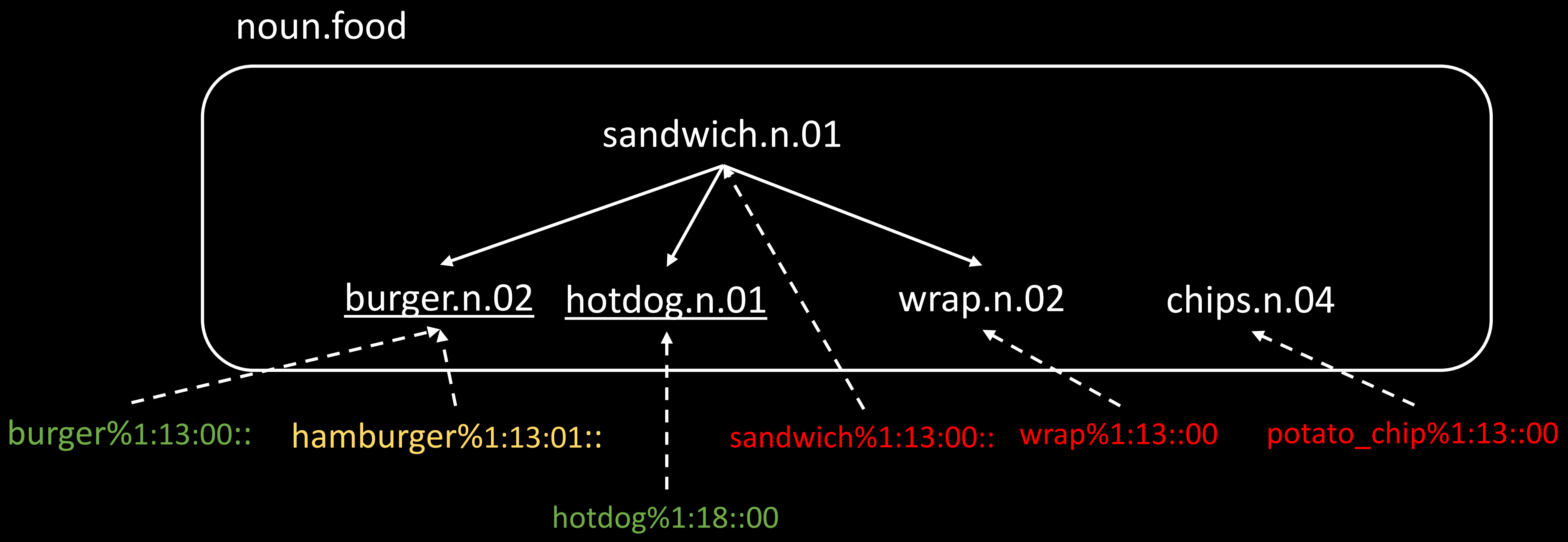
Propagating Sense Embeddings

Retrieve Synsets, Relations and Categories



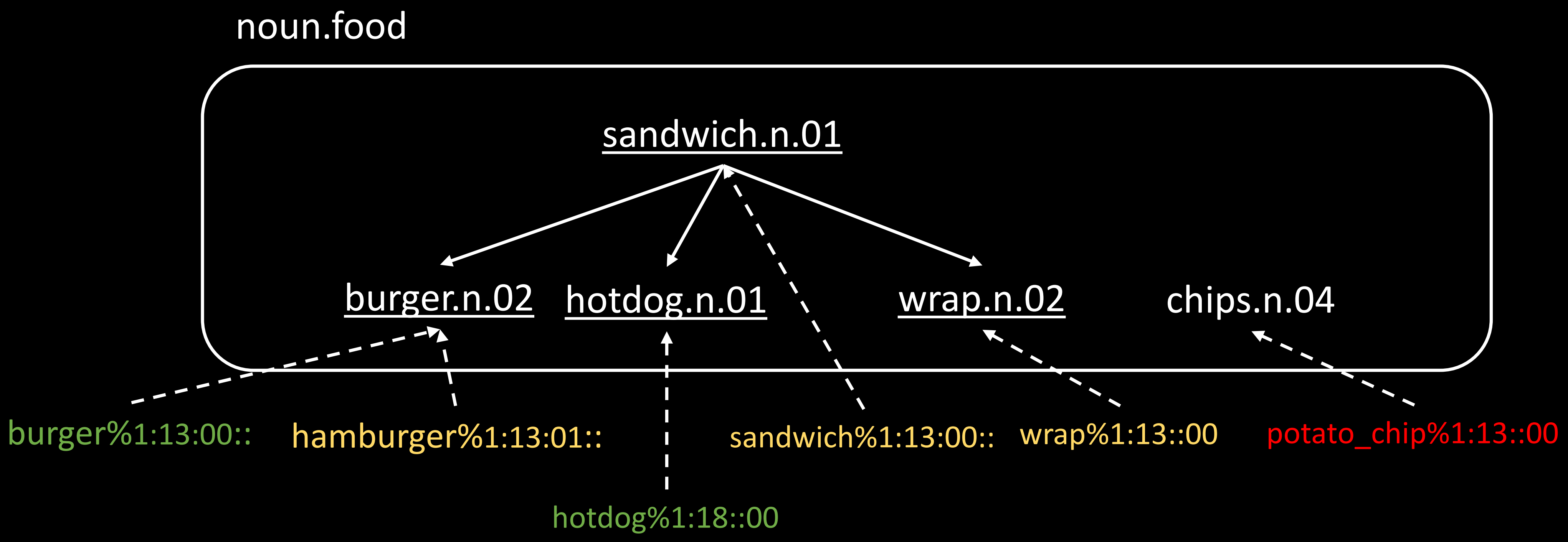
Propagating Sense Embeddings

1st stage: Synset Embeddings



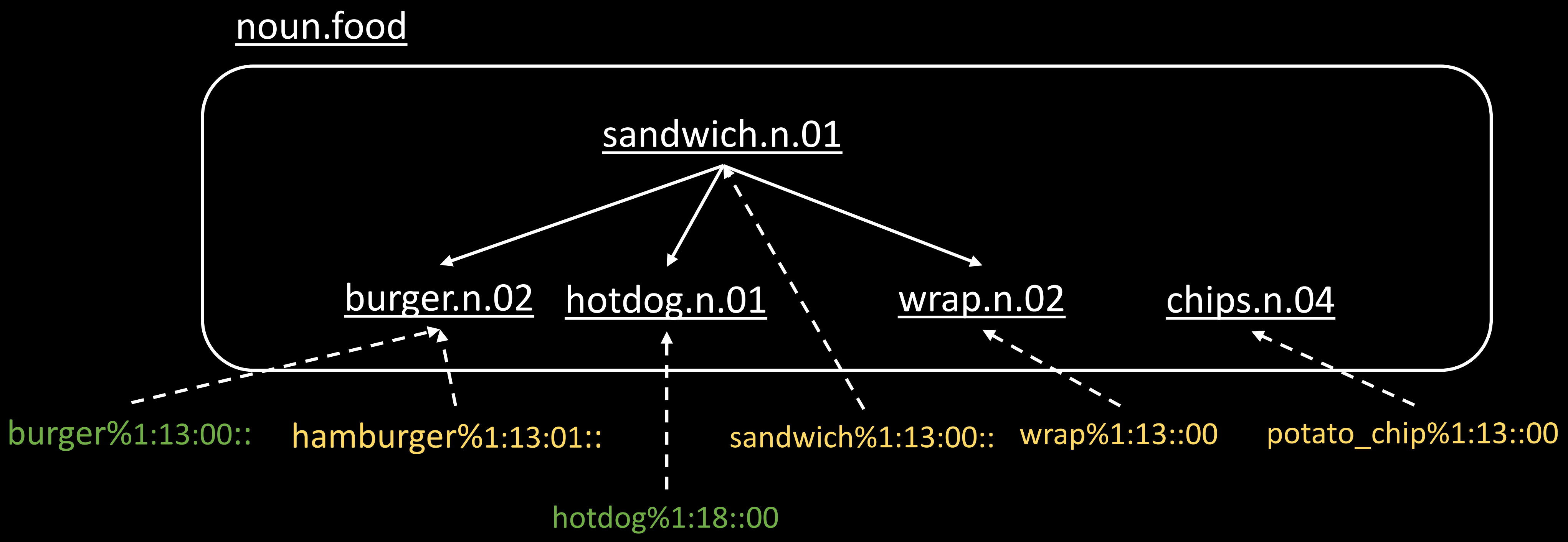
Propagating Sense Embeddings

2^{nd} Stage: Hypernym Embeddings (ind. Synsets)



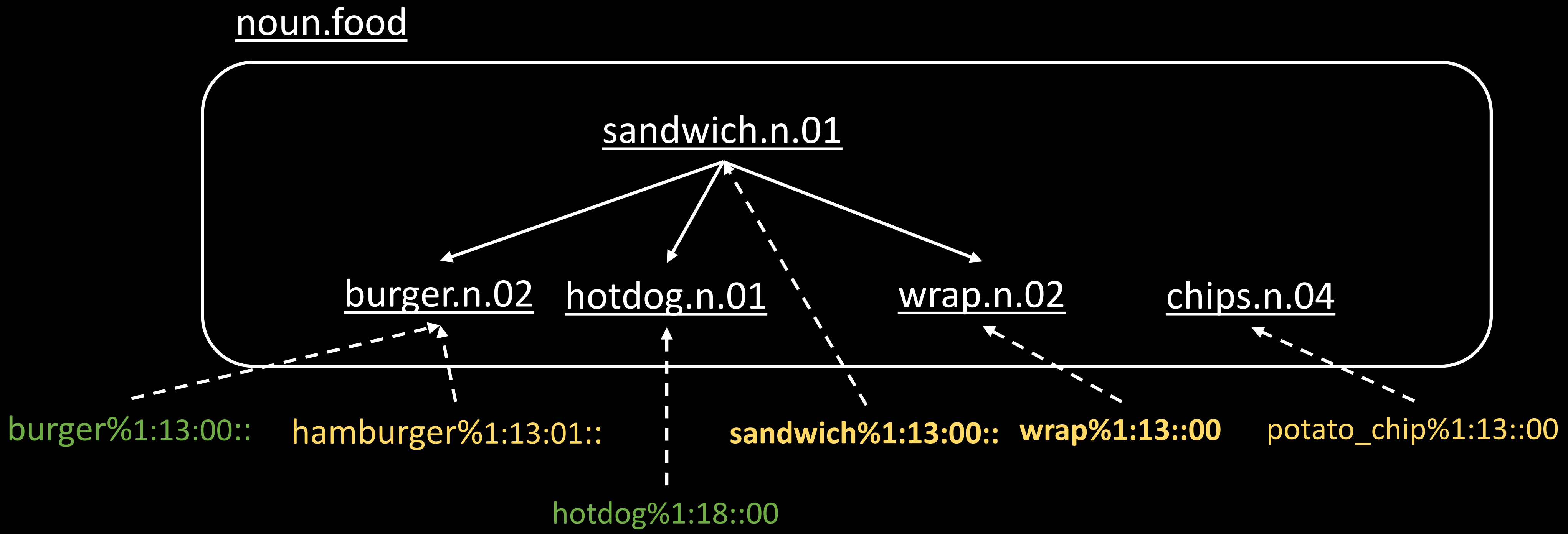
Propagating Sense Embeddings

3rd Stage: Lexname Embeddings



Propagating Sense Embeddings

But  !=  …



Enriching Sense Embeddings

Leverage Synset Definitions and Lemmas for Differentiation

Enriching Sense Embeddings

Leverage Synset Definitions and Lemmas for Differentiation

sandwich:%1:13:00:: (sandwich.n.01)

Definition: two (or more) slices of bread with a filling between them

Lemmas: sandwich

wrap:%1:13:00:: (wrap.n.02)

Definition: a sandwich in which the filling is rolled up in a soft tortilla

Lemmas: wrap, tortilla

Enriching Sense Embeddings

Compose a new context

sandwich:%1:13:00:: (sandwich.n.01)

sandwich - two (or more) slices of bread with a filling between them

wrap:%1:13:00:: (wrap.n.02)

wrap, tortilla - a sandwich in which the filling is rolled up in a soft tortilla

Enriching Sense Embeddings

Make the context specific to sensekey (repeat lemma)

sandwich:%1:13:00::

sandwich - sandwich - two (or more) slices of bread with a filling between them

wrap%1:13:00::

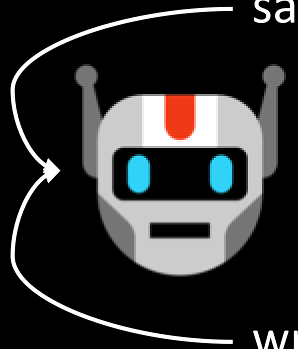
wrap - wrap, tortilla - a sandwich in which the filling is rolled up in a soft tortilla

Enriching Sense Embeddings

Make the context specific to sensekey (repeat lemma)

sandwich:%1:13:00::

sandwich - sandwich - two (or more) slices of bread with a filling between them



wrap%1:13:00::

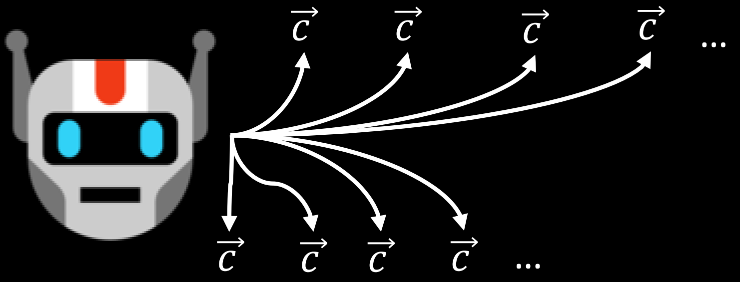
wrap - wrap, tortilla - a sandwich in which the filling is rolled up in a soft tortilla

Enriching Sense Embeddings

Obtain contextual embeddings for every token

sandwich:%1:13:00::

sandwich - sandwich - two (or more) slices of bread with a filling between them

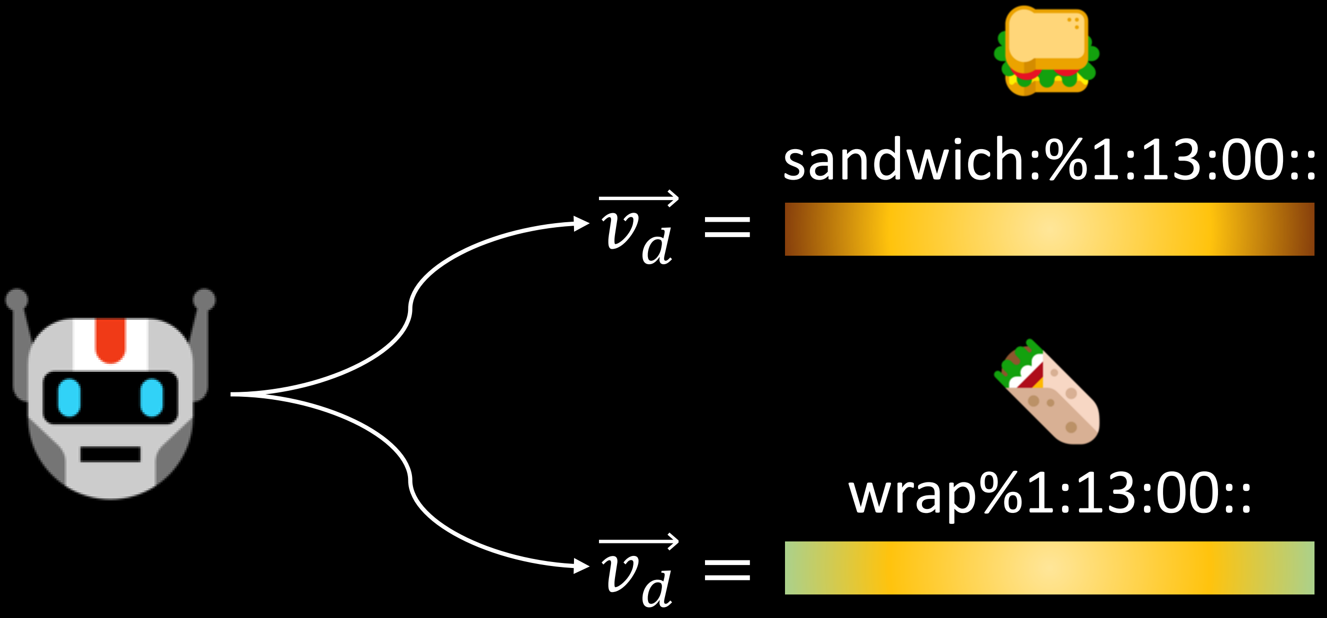


wrap%1:13:00::

wrap – wrap, tortilla - a sandwich in which the filling is rolled up in a soft tortilla

Enriching Sense Embeddings

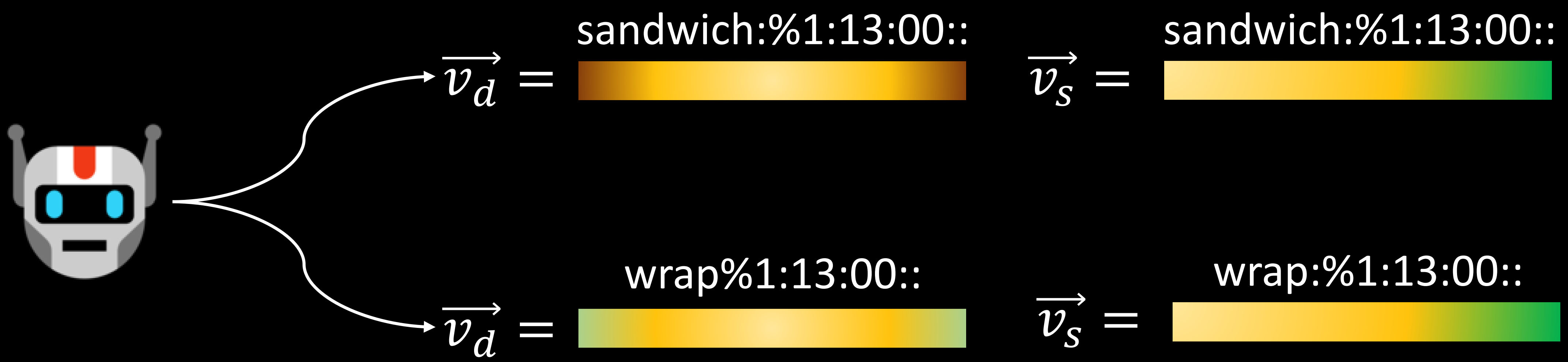
Sentence Embedding from avg. of Contextual Embeddings



$d = 1024$

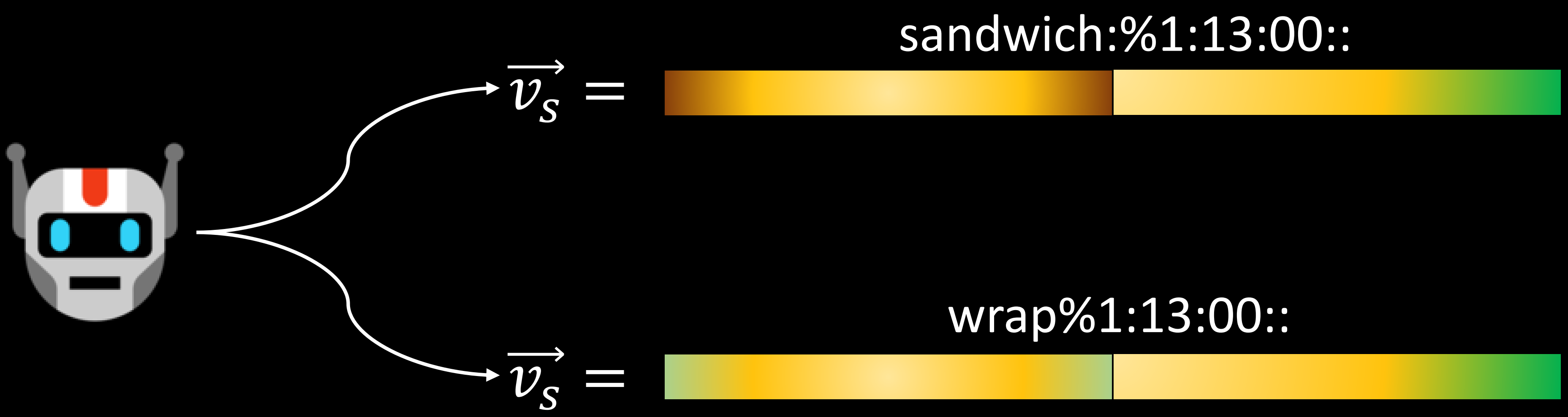
Enriching Sense Embeddings

Merge Sentence Embedding with previous Sense Embedding



Enriching Sense Embeddings

Merge Sentence Embedding with previous Sense Embedding



$d = 2048$

Matching Sense Embeddings

The glasses are in the cupboard.

Matching Sense Embeddings

 $\longrightarrow \vec{c}$

The <u>glasses</u> are in the cupboard.

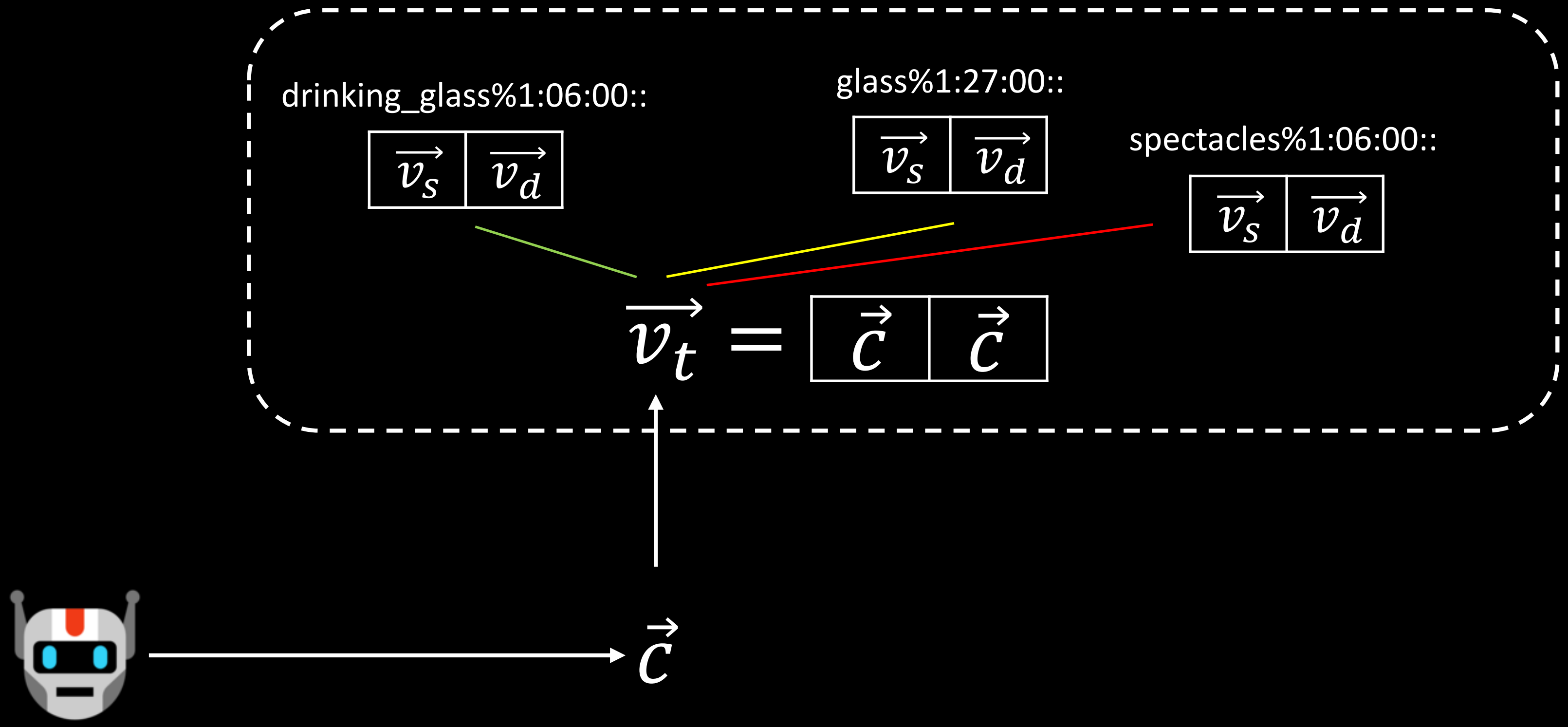
Matching Sense Embeddings

$$\overrightarrow{v_t} = \boxed{\vec{c}}\boxed{\vec{c}}$$

$\vec{c}$

The <u>glasses</u> are in the cupboard.

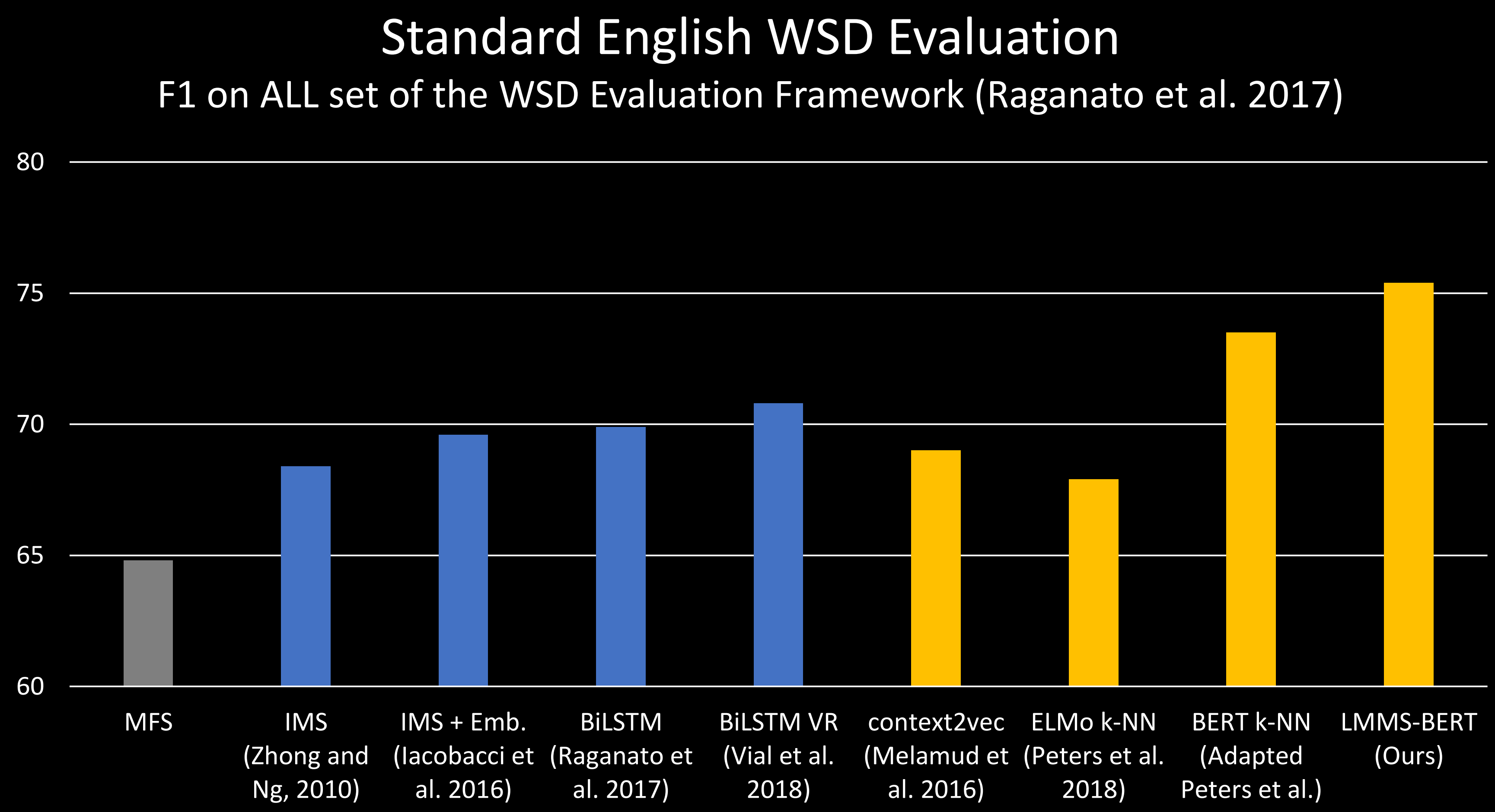
Matching Sense Embeddings



The <u>glasses</u> are in the cupboard.

WSD Results

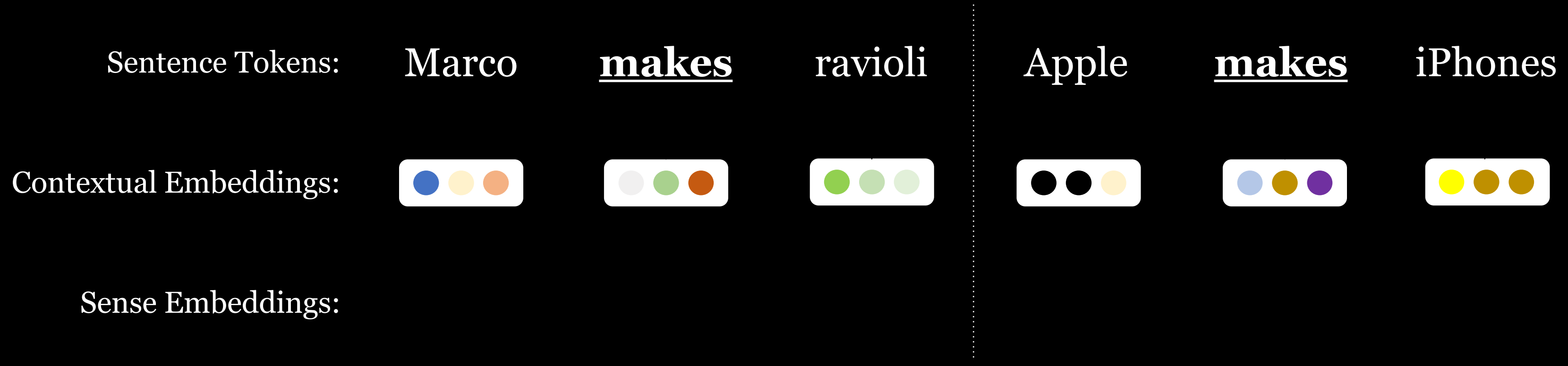
WSD Performance



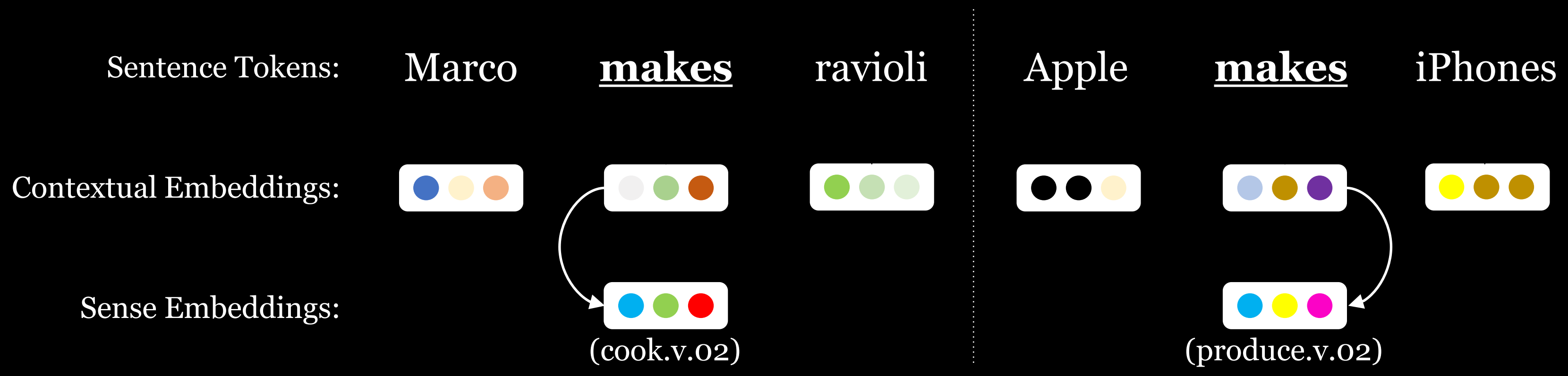
Classifying Embedding Similarities

Sentence Tokens:	Marco	**makes**	ravioli	Apple	**makes**	iPhones
Contextual Embeddings:						
Sense Embeddings:						

Classifying Embedding Similarities



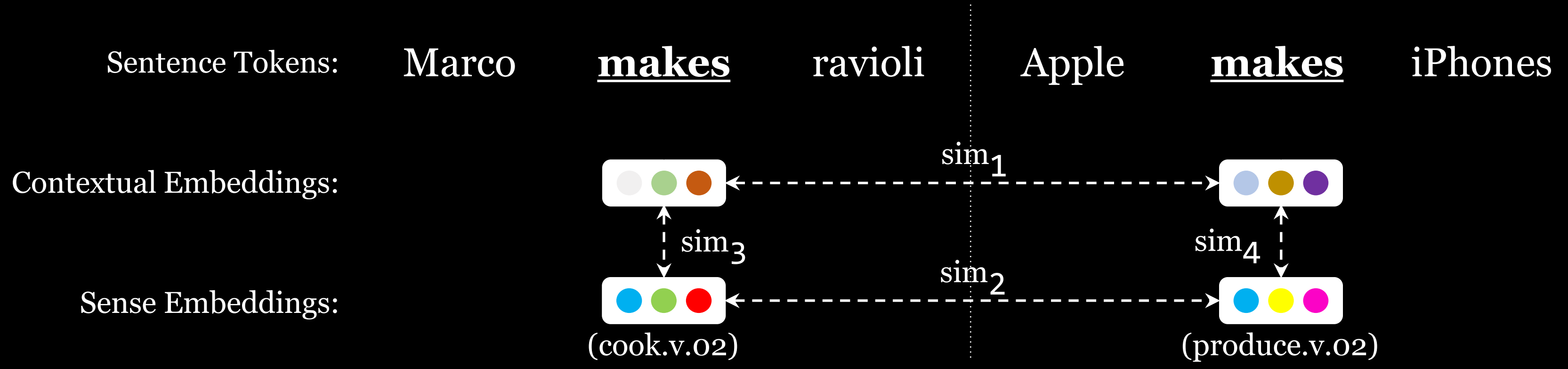
Classifying Embedding Similarities



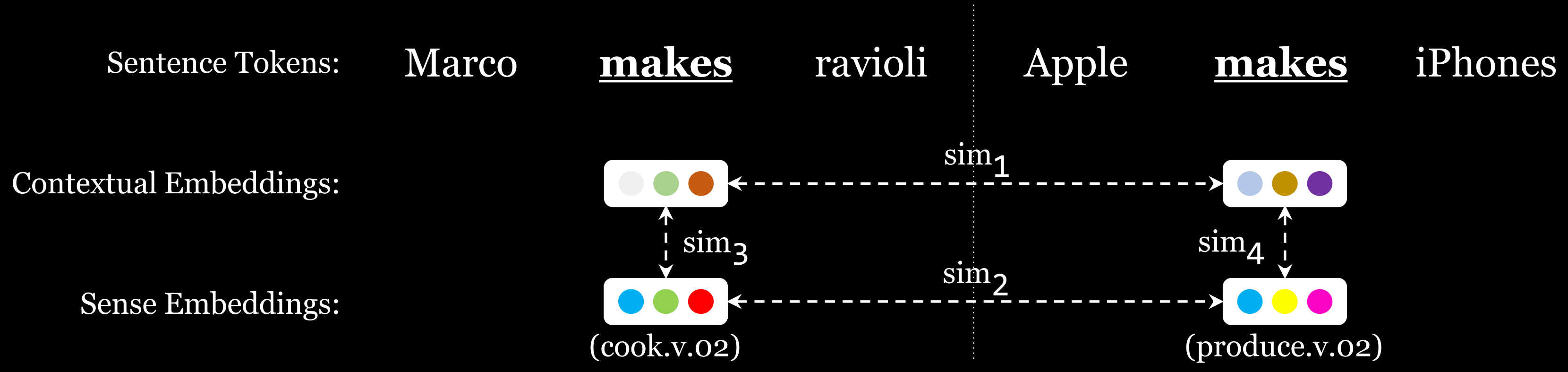
Classifying Embedding Similarities

Sentence Tokens:	Marco	**makes**	ravioli	Apple	**makes**	iPhones
Contextual Embeddings:						
Sense Embeddings:		(cook.v.02)			(produce.v.02)	

Classifying Embedding Similarities

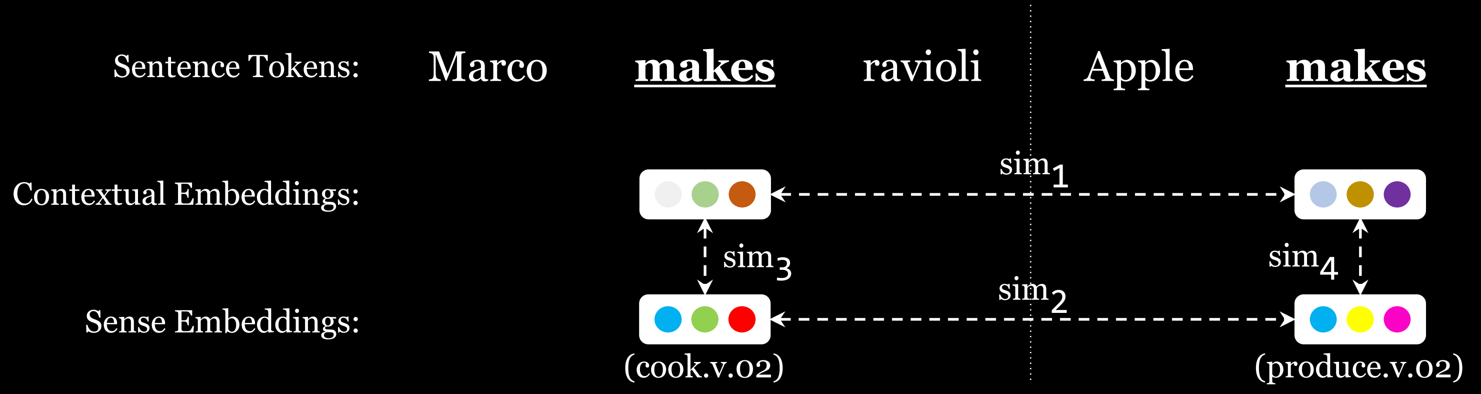


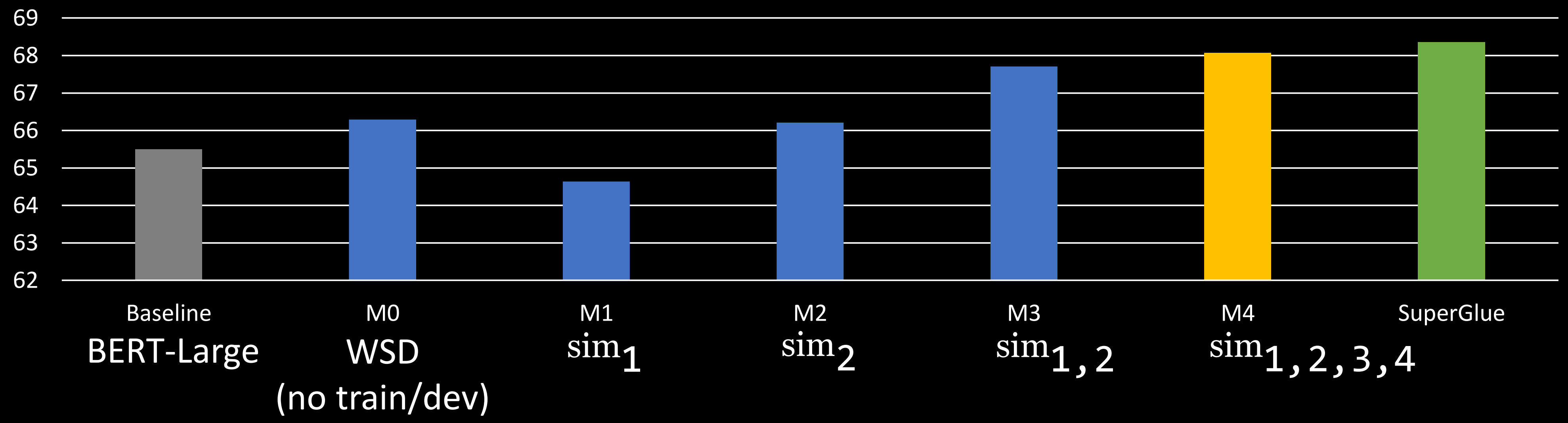
Classifying Embedding Similarities



Now, we classify different similarity combinations using Binary Logistic Regression

Classifying Embedding Similarities





Conclusions

- Systems designed for WSD, without being trained for the WiC task, can perform competitively.
- Sense Embeddings can still benefit from information captured by contextual embeddings, as shown by similarities classifier.
- In future work, progress on the WiC task could lead to better semi-supervised annotations for WSD.

Thanks

Code and Sense Embeddings:
github.com/danlou/LMMS

dloureiro@fc.up.pt

@danielbloureiro