

Cross-Lingual Pronoun Prediction

Again

Participation in DiscoMT Shared Task 2015

Jimmy Callin
May 2015

Everybody talks about happiness these days .

Everybody talks about happiness these days .

Aujourd' hui tout le monde parle du bonheur .

Everybody talks about happiness these days .

Aujourd' hui tout le monde parle du bonheur .

it expresses our view of how we should try to relate to the future

??? exprime notre manière d' aborder le futur

it expresses our view of how we should try to relate to the future

ce exprime notre manière d' aborder le futur
cela
elle
elles
il
ils

it expresses our view of how we should try to relate to the future

ce exprime notre manière d' aborder le futur
cela
elle
elles
il
ils

it expresses our view of how we should try to relate to the future

ce exprime notre manière d' aborder le futur

cela
elle
elles
il
ils

ils	14898
ce	14555
il	8467
elle	2999
elles	2888
cela	2256

it expresses our view of how we should try to relate to the future

ce exprime notre manière d' aborder le futur

~~cela~~
elle
elles
il
ils

ils	14898
ce	14555
il	8467
elle	2999
elles	2888
cela	2256

it expresses our view of how we should try to relate to the future

ce
cela
elle
elles

il exprime notre manière d' aborder le futur
ils

ils	14898
ce	14555
il	8467
elle	2999
elles	2888
cela	2256

it expresses our view of how we should try to relate to the future

ce

cela

elle

elles

il exprime notre manière d' aborder le futur

ils

We have this banner in our offices in Palo Alto

it expresses our view of how we should try to relate to the future

ce

cela

elle

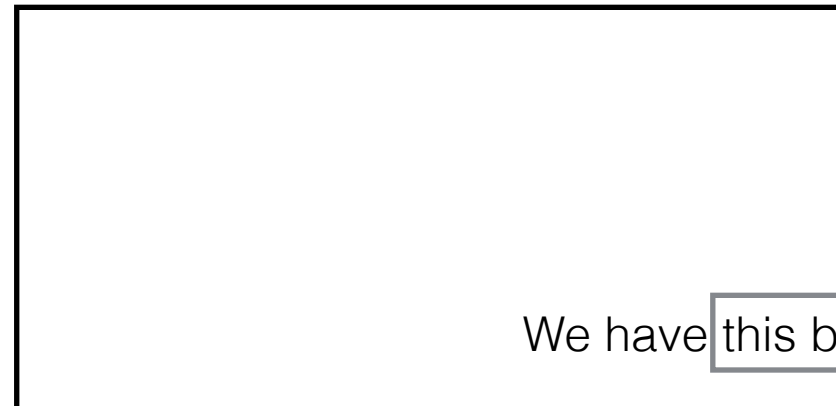
elles

Nous avons cette bannière dans nos bureaux à Palo Alto

il

exprime notre manière d' aborder le futur

ils



We have **this banner** in our offices in Palo Alto

it expresses our view of how we should try to relate to the future

ce

cela

elle

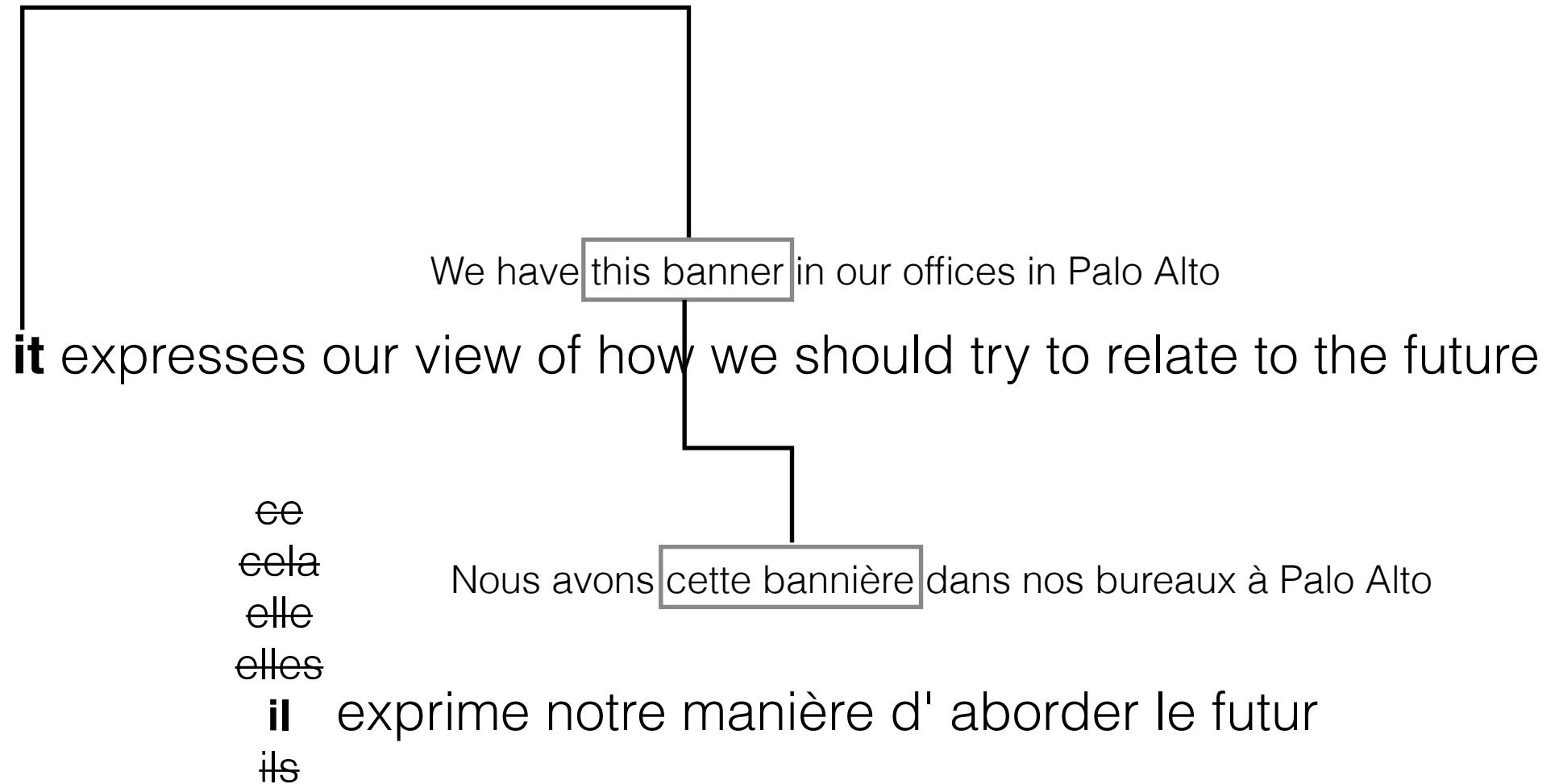
elles

il

ils

Nous avons cette bannière dans nos bureaux à Palo Alto

exprime notre manière d' aborder le futur



We have this banner in our offices in Palo Alto

it expresses our view of how we should try to relate to the future

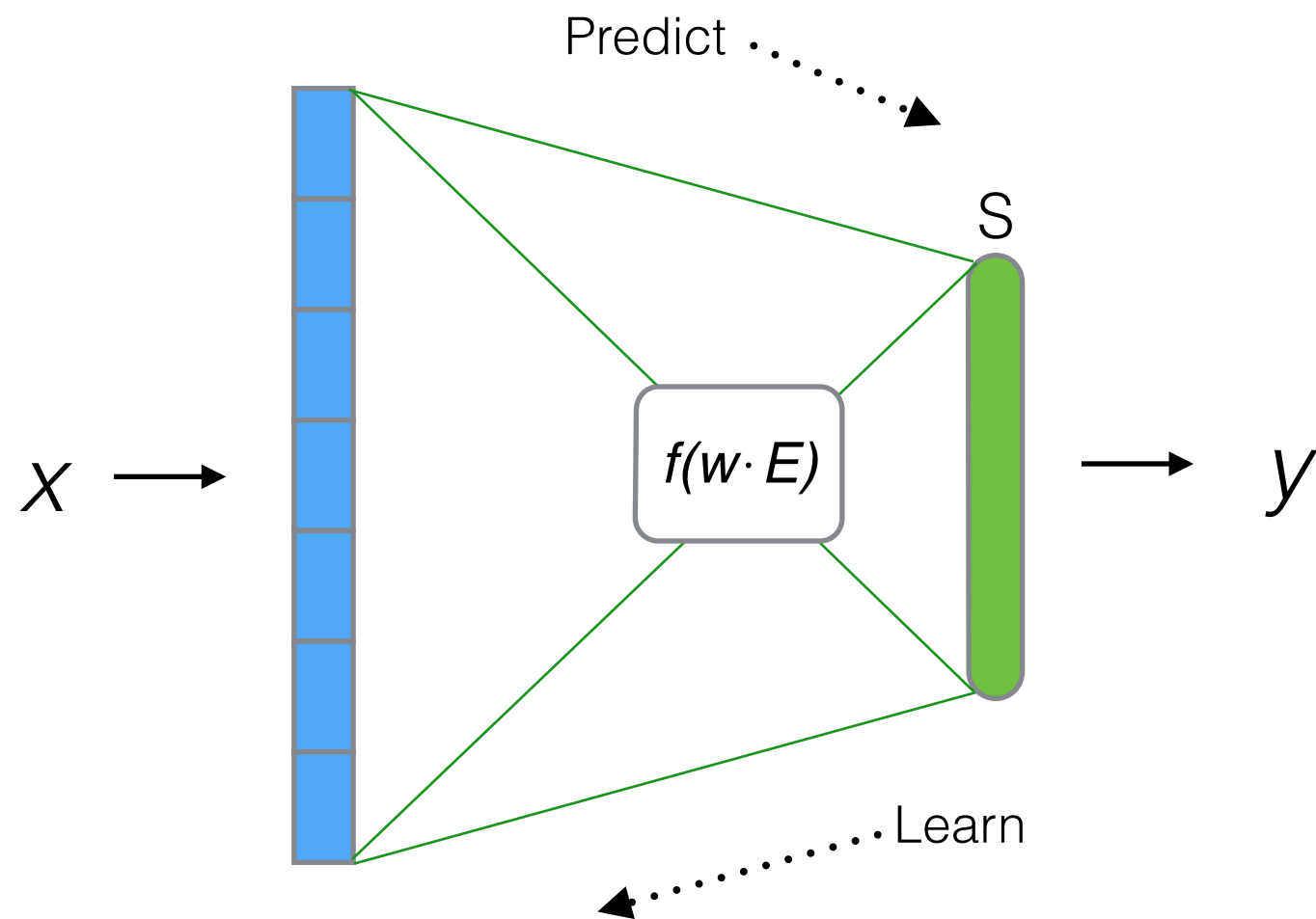
Nous avons cette bannière dans nos bureaux à Palo Alto

elle exprime notre manière d' aborder le futur

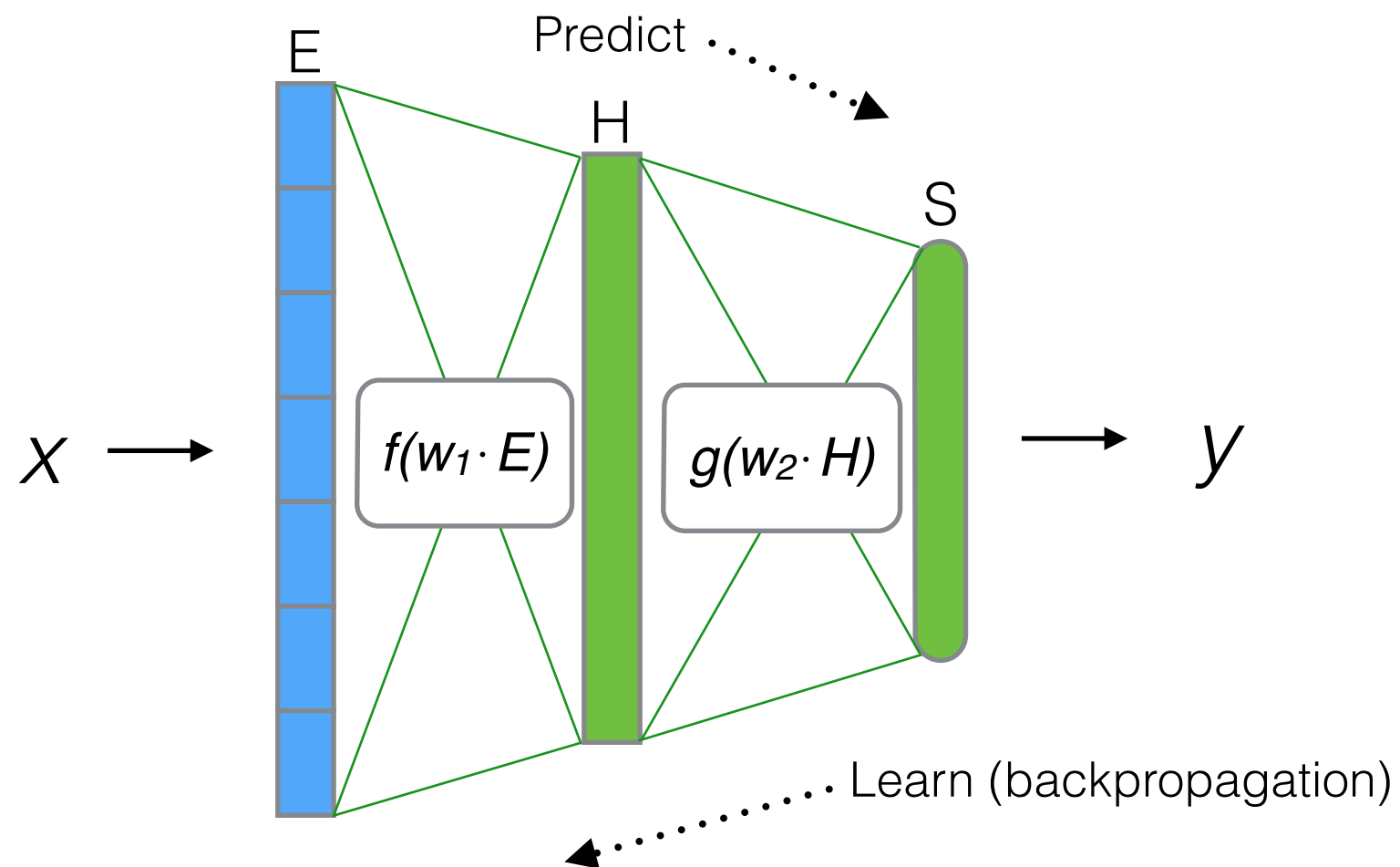
Artificial Neural Networks

in three minutes

Perceptron

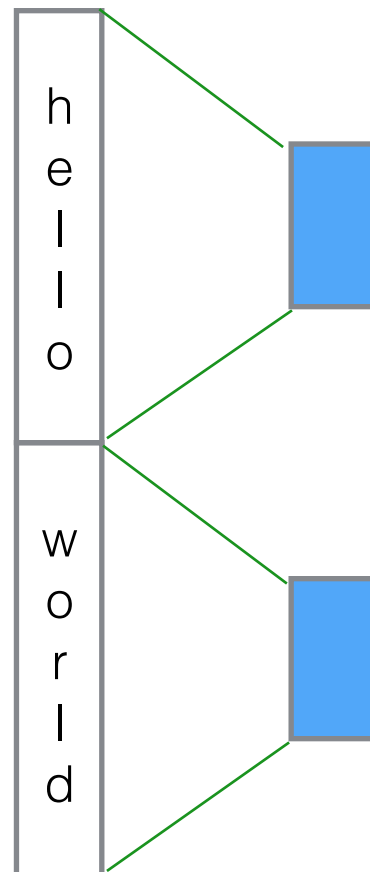


Multi-Layered Perceptron

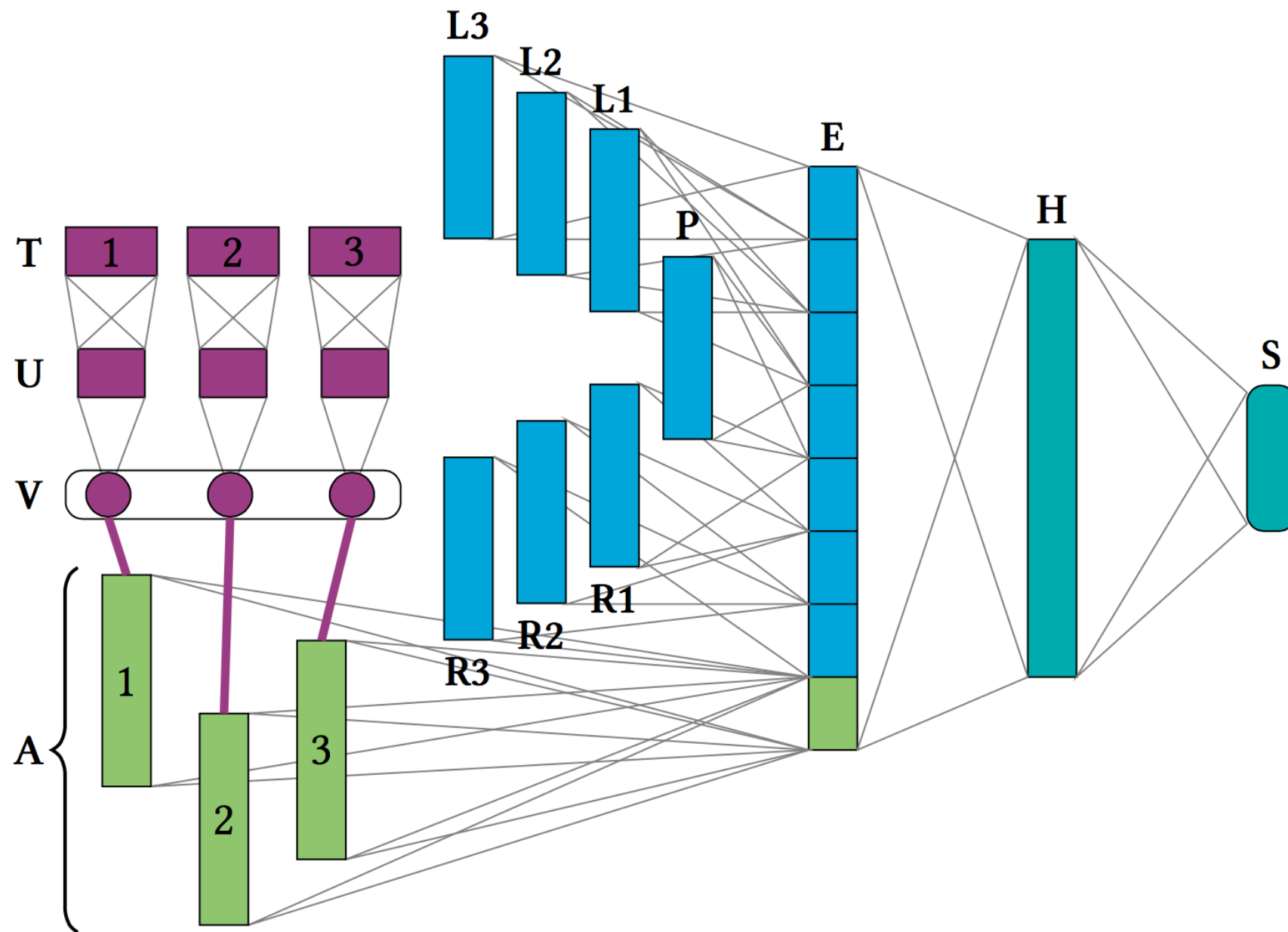


Word Embeddings

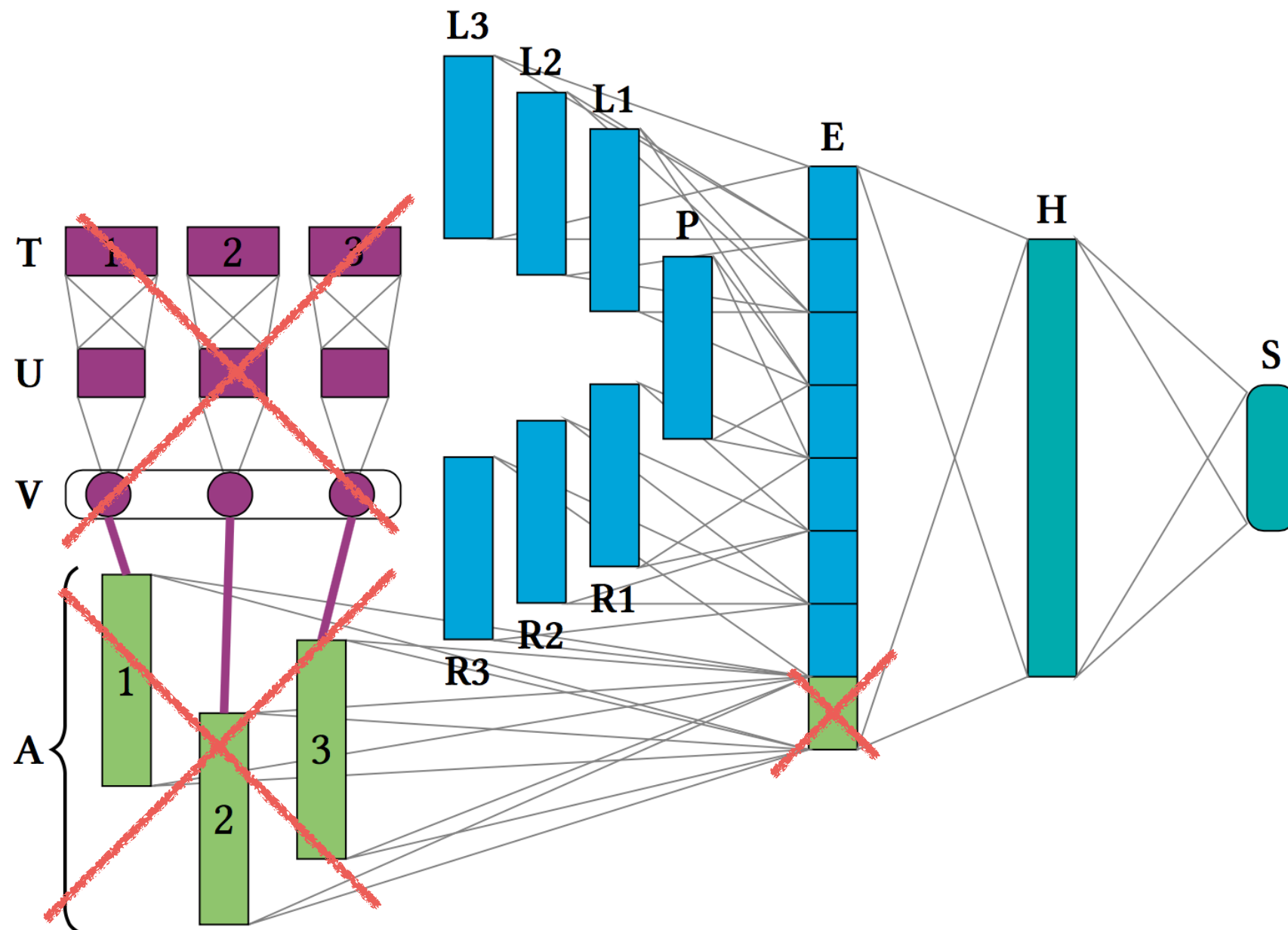
Words are mapped from one-hot vector to lower-dimensional vector. Hopefully this creates some sort of useful abstract representation.

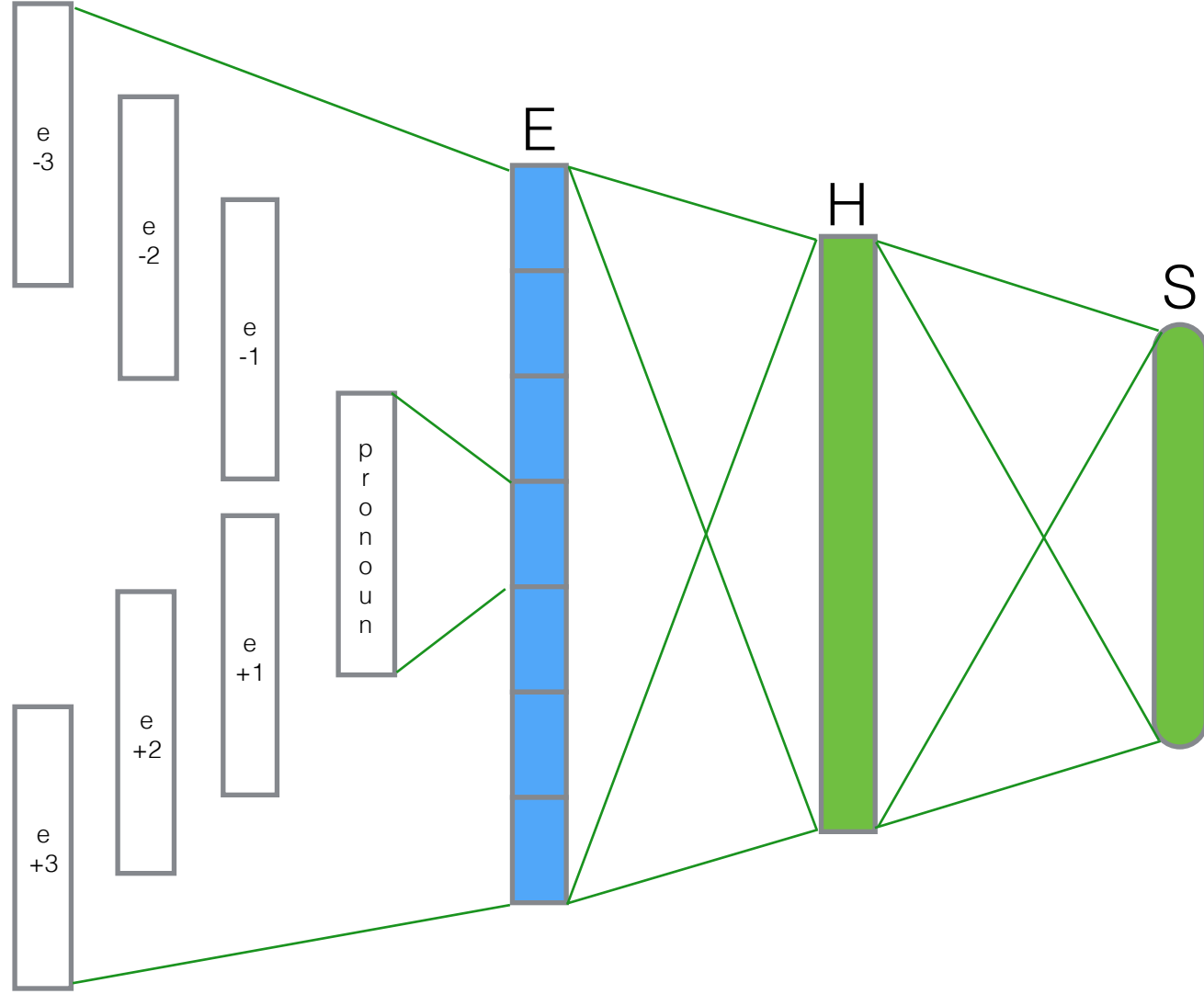


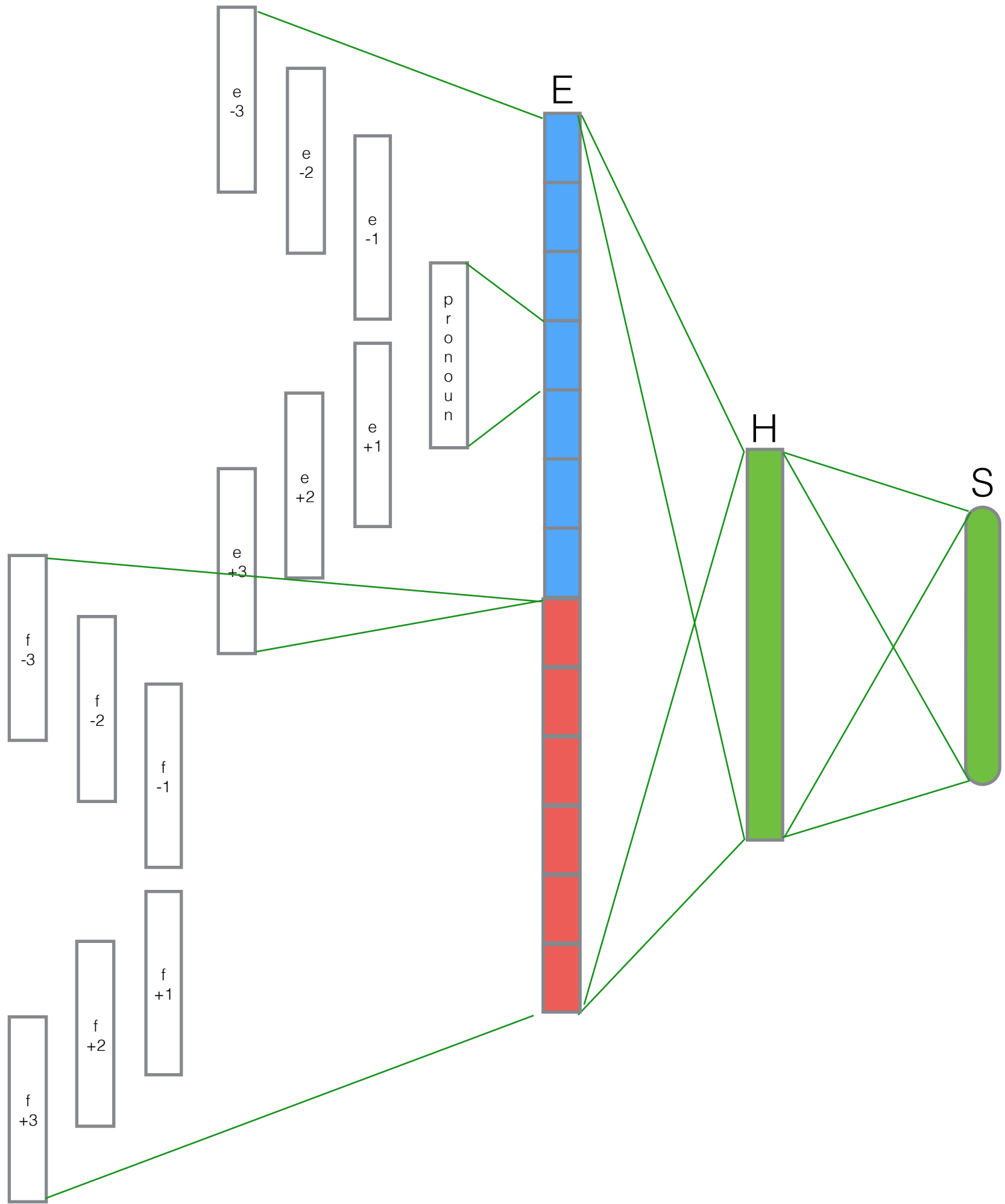
Hardmeier (2013)



Hardmeier (2013)





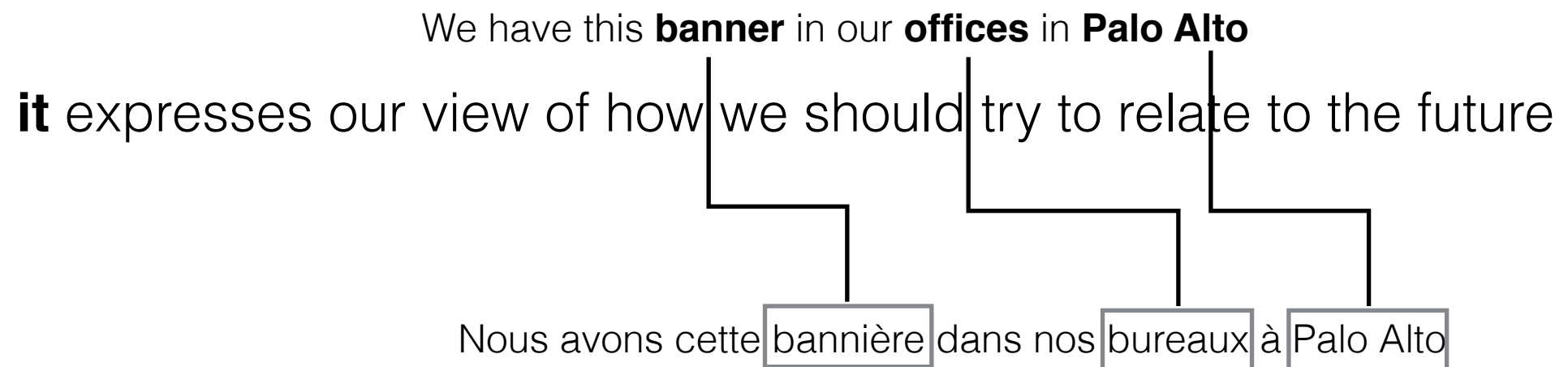


Adding nouns

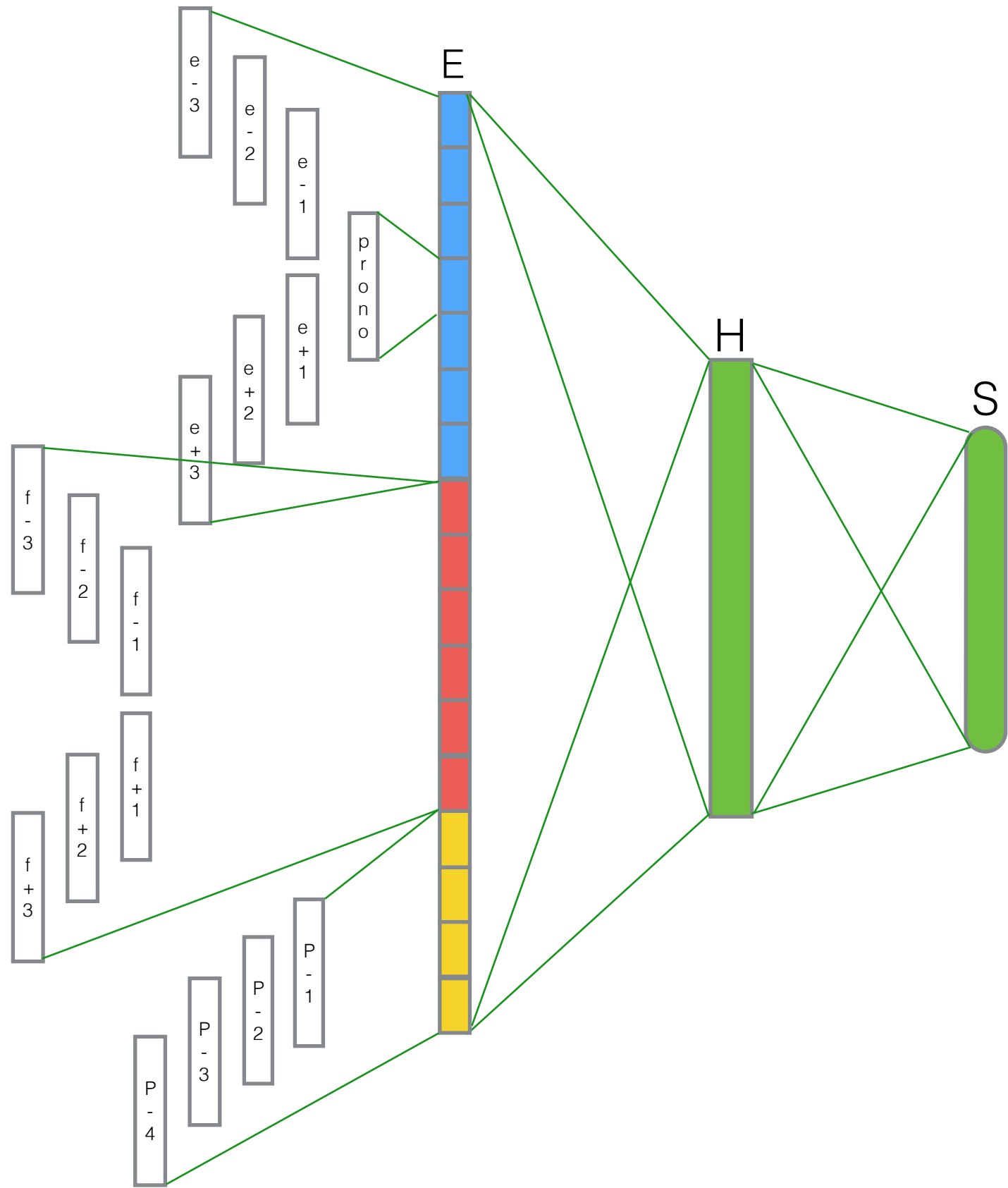


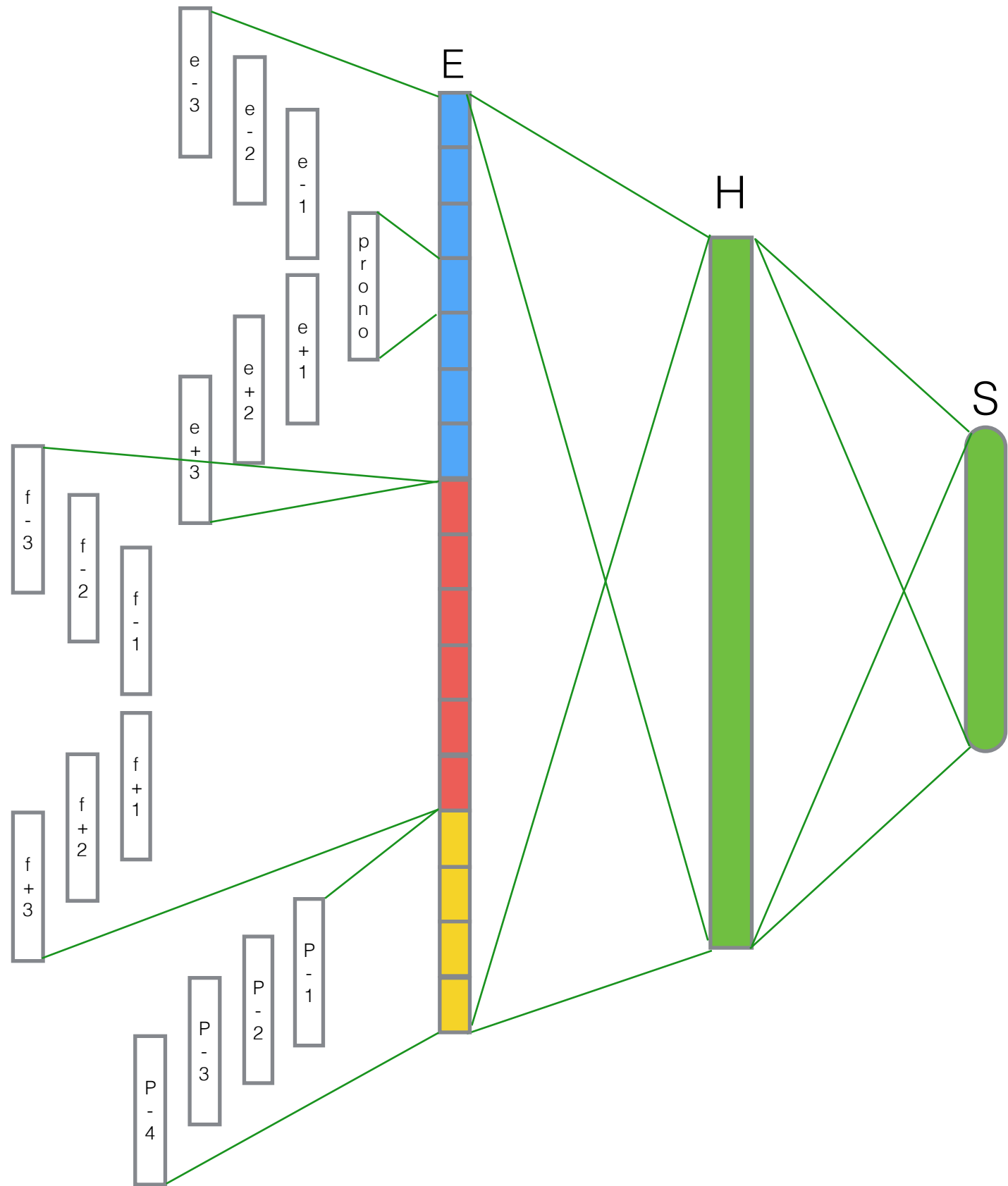
??? exprime notre manière d' aborder le futur

Adding nouns



??? exprime notre manière d' aborder le futur





Final features

<S> <S> <S> **it expresses our view** of how we should try to relate to the future

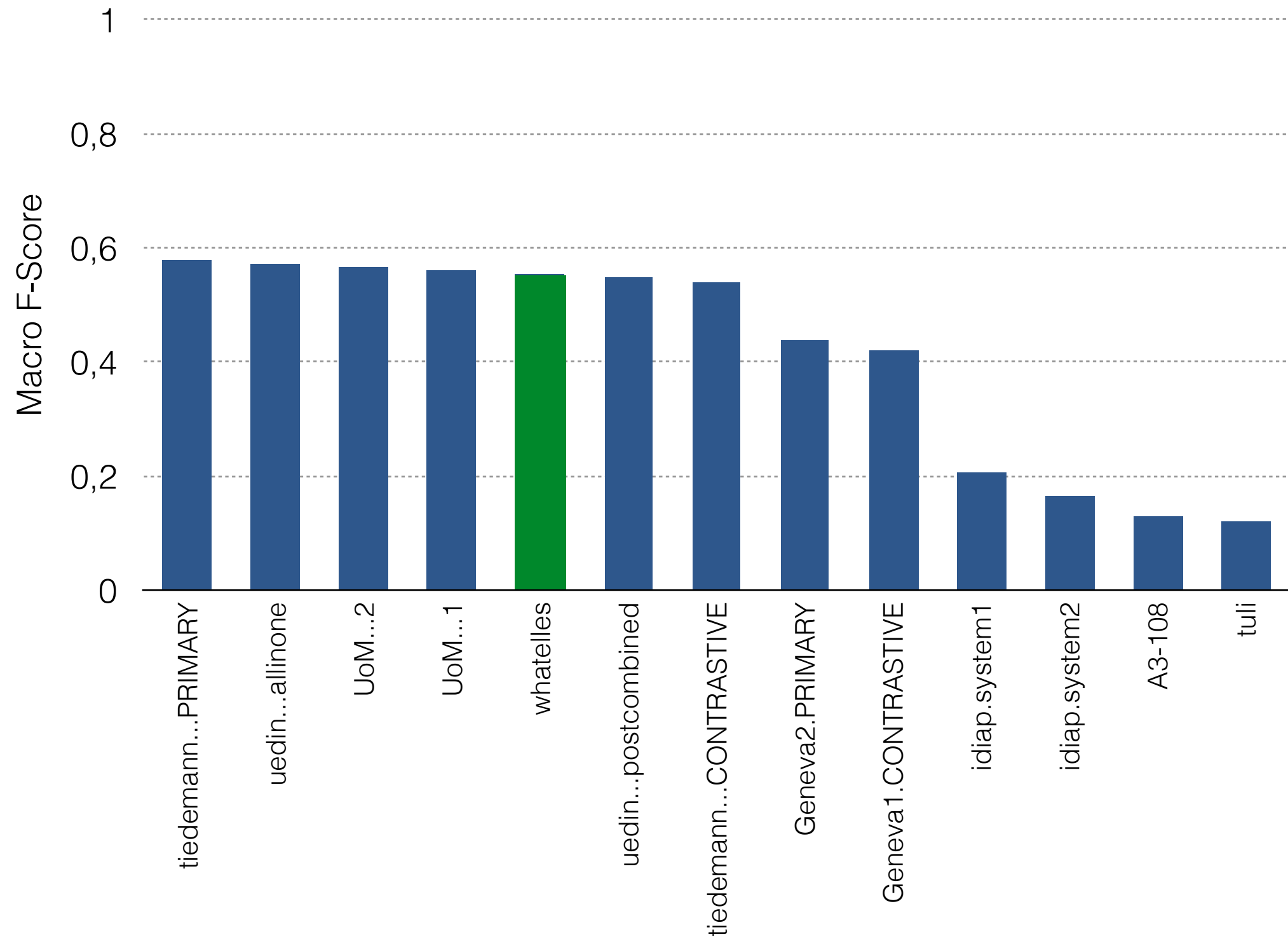
<S> <S> <S> **__ exprime notre manière** d' aborder le futur

We have this **banner** in our **offices** in **Palo Alto**

Nous avons cette **bannière** dans nos **bureaux** à **Palo** Alto



Results



Scores

	Precision	Recall	F1	Macro-F1	Micro-F1
ce	82.91%	89.67%	86.16%	55.30%	72.13%
cela	71.43%	62.02%	66.39%		
elle	50.00%	26.51%	34.65%		
elles	62.96%	33.33%	43.59%		
il	51.61%	61.54%	56.14%		
ils	74.87%	93.12%	83.01%		
OTHER	84.50%	85.79%	85.14%		

Confusion Matrix

	ce	cela	elle	elles	il	ils	OTHER	SUM
ce	165	3	0	1	8	1	6	184
cela	5	80	4	1	21	0	18	129
elle	7	10	22	2	22	2	18	83
elles	0	0	0	18	0	31	3	51
il	11	7	9	0	64	1	12	104
ils	1	0	0	5	0	149	5	160
OTHER	10	12	9	1	9	15	338	394
SUM	199	112	44	27	124	199	400	1105

Code available at
github.com/jimmycallin/whatelles

Please don't try it out until I've cleaned it up a bit.

(Thanks)