Synthetic Data **Made to Order**: The Case of Parsing

Dingquan Wang and Jason Eisner
Department of Computer Science
JOHNS HOPKINS UNIVERSITY

Synthetic Data **Made to Order**: The Case of Dependency Parsing

Dingquan Wang and Jason Eisner
Department of Computer Science
JOHNS HOPKINS UNIVERSITY

Dependency Parsing

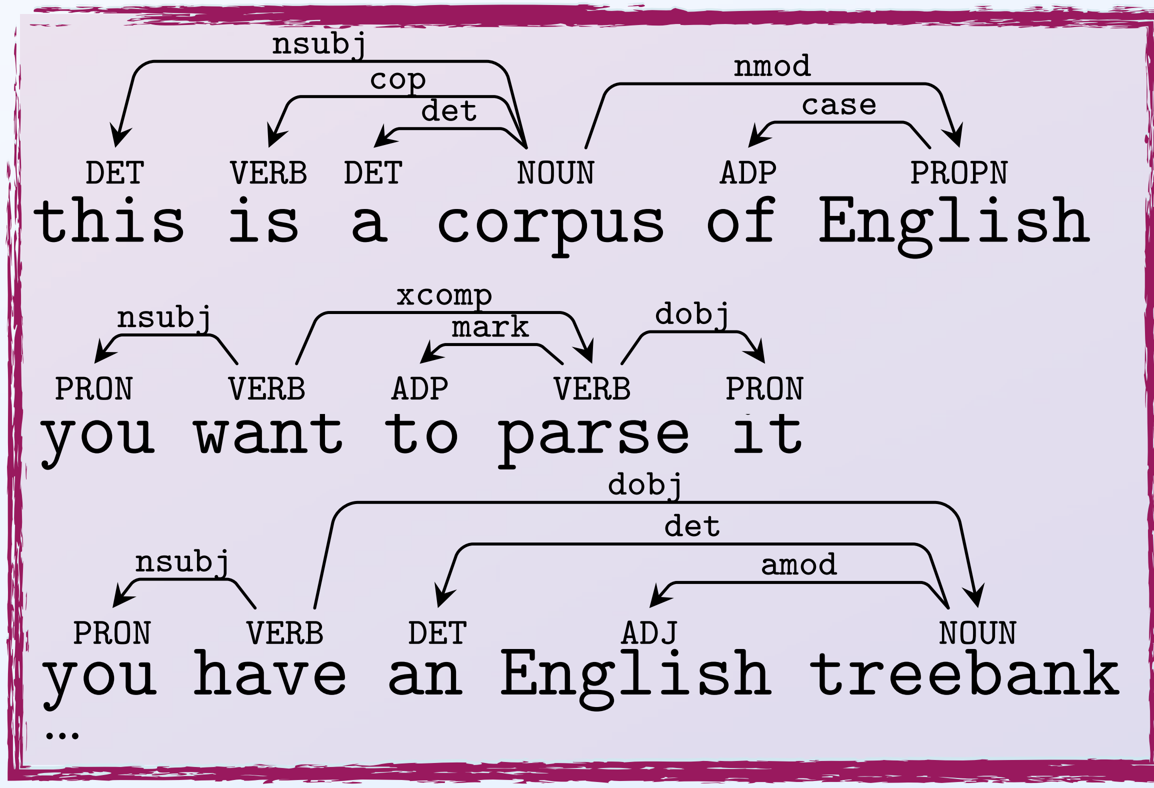
Dependency Parsing

English Corpus

```
this is a corpus of English
you want to parse it
you have an English treebank
…
```

Dependency Parsing

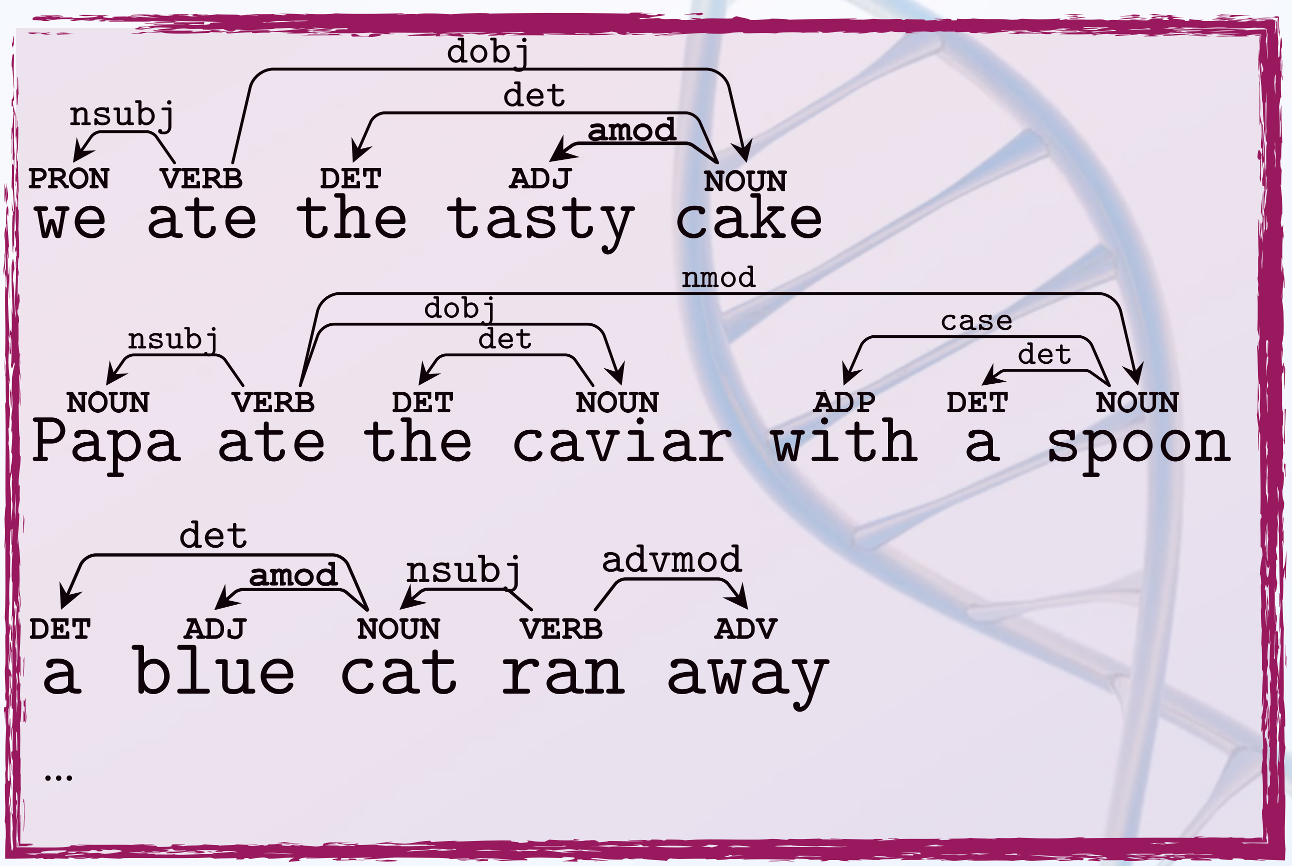
English Corpus



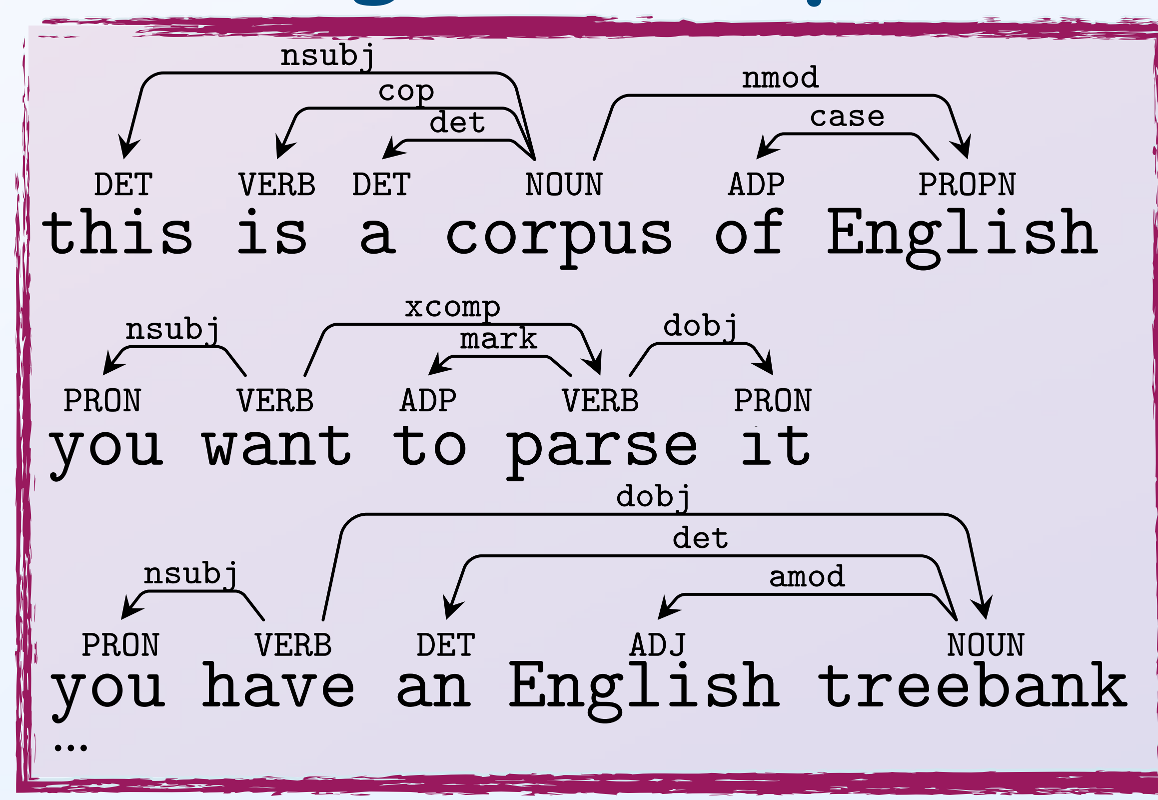
JOHNS HOPKINS
UNIVERSITY

Dependency Parsing

English Treebank

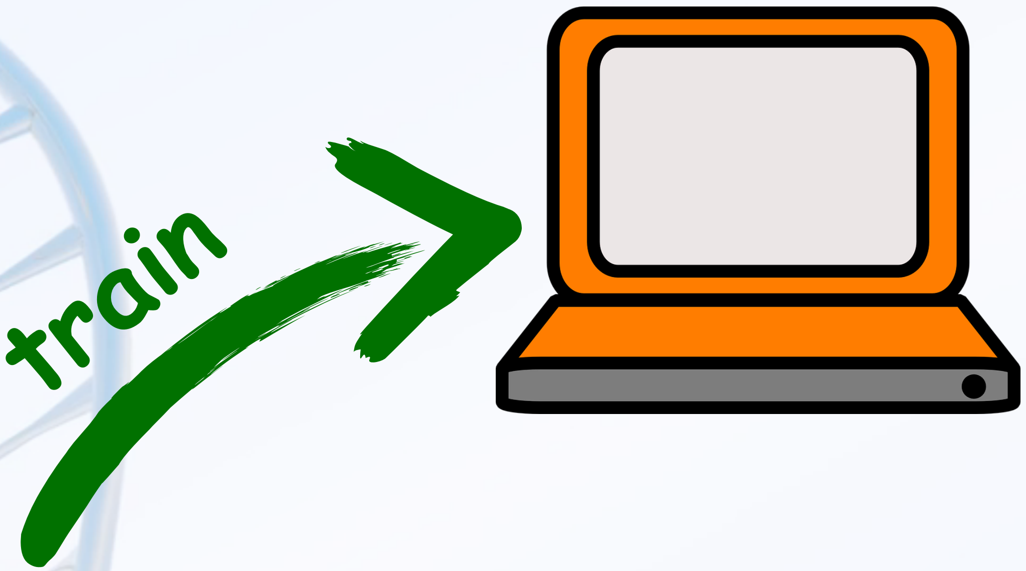


English Corpus

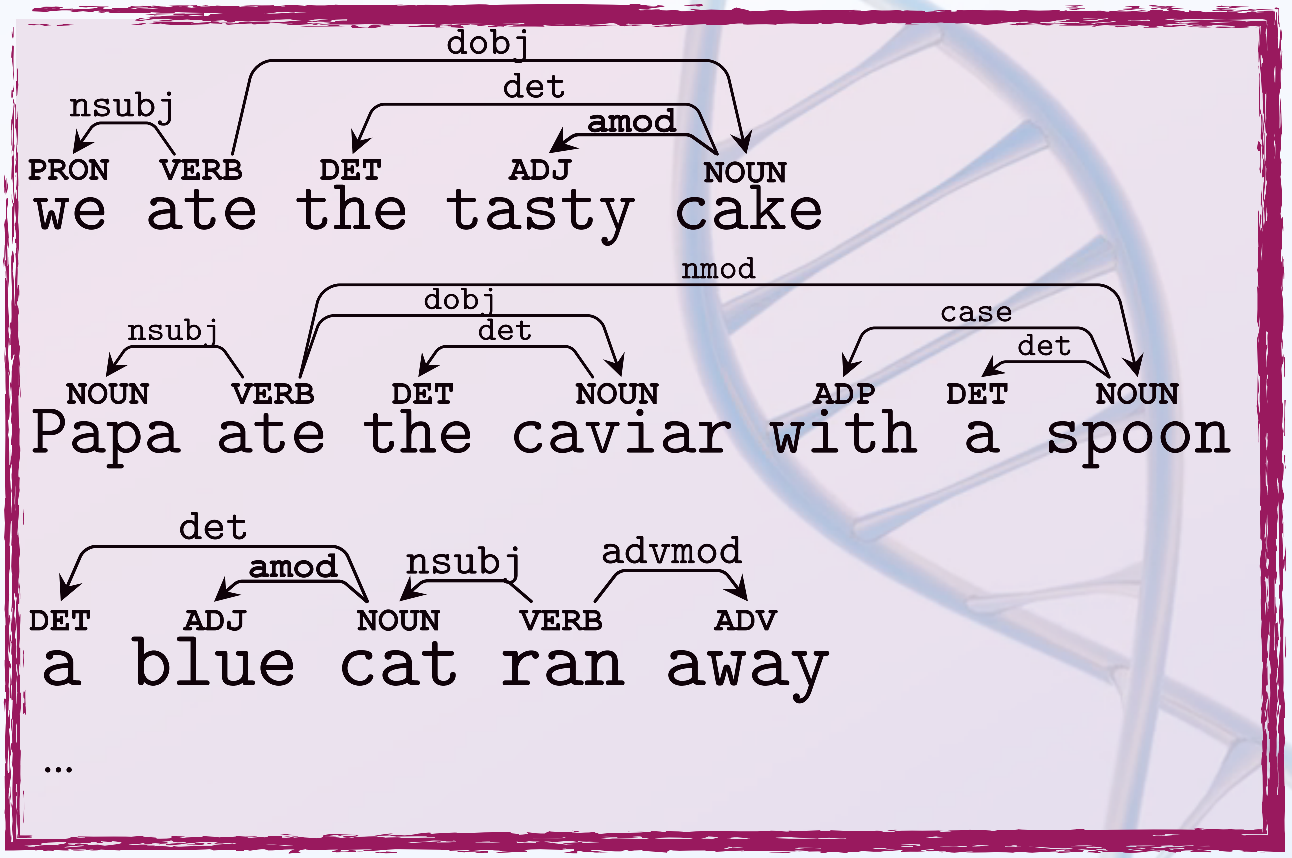


JOHNS HOPKINS UNIVERSITY

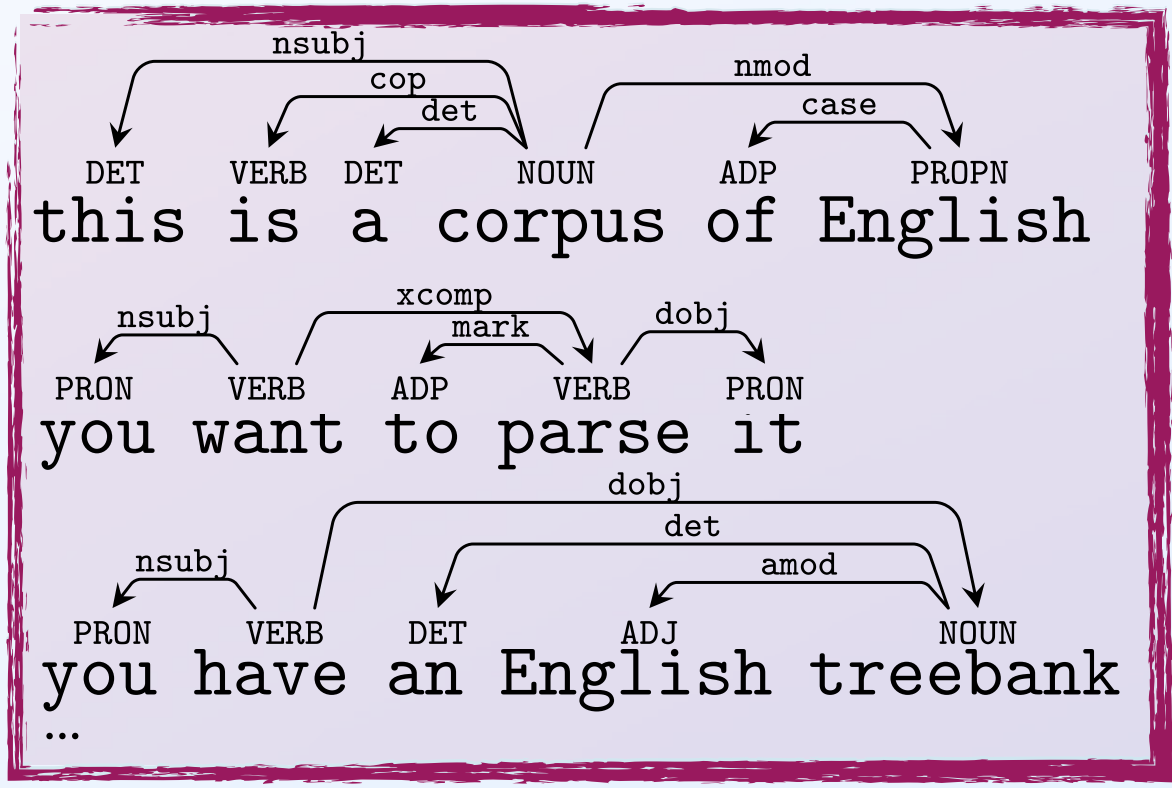
Dependency Parsing



English Treebank

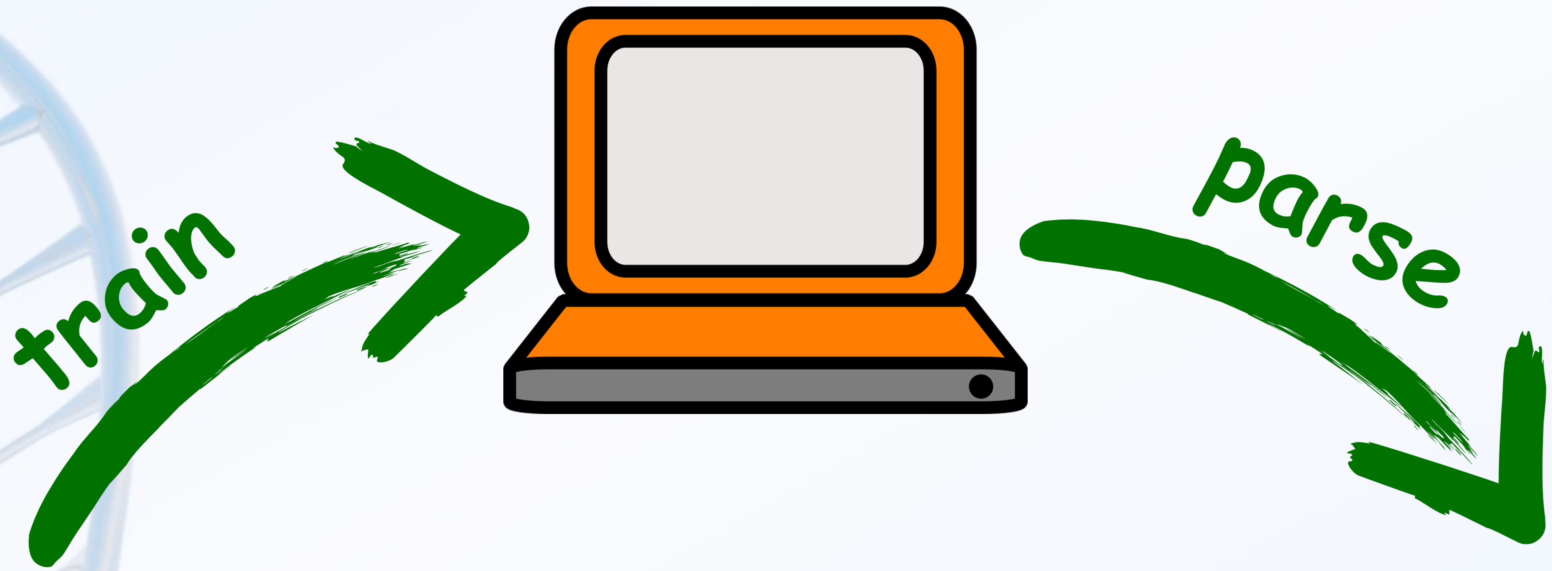


English Corpus

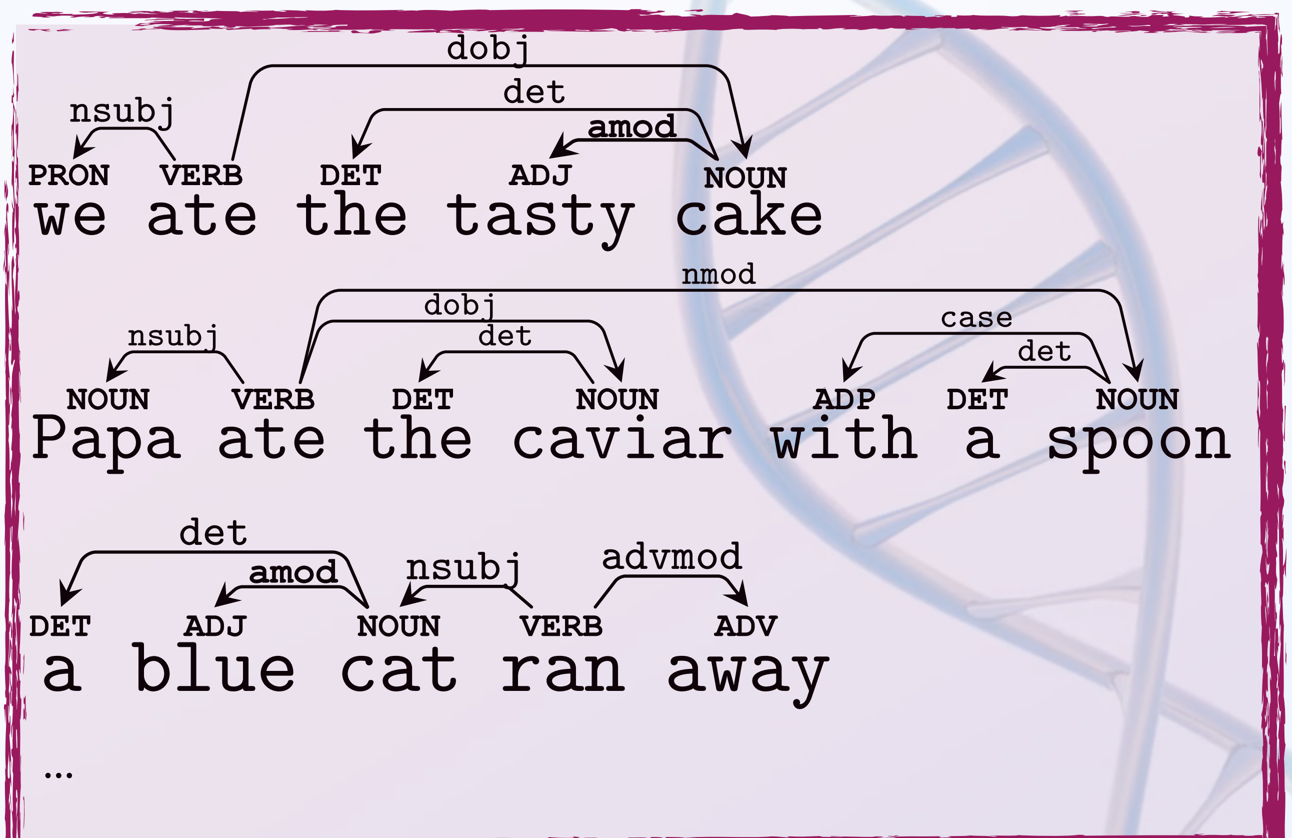


JOHNS HOPKINS
UNIVERSITY

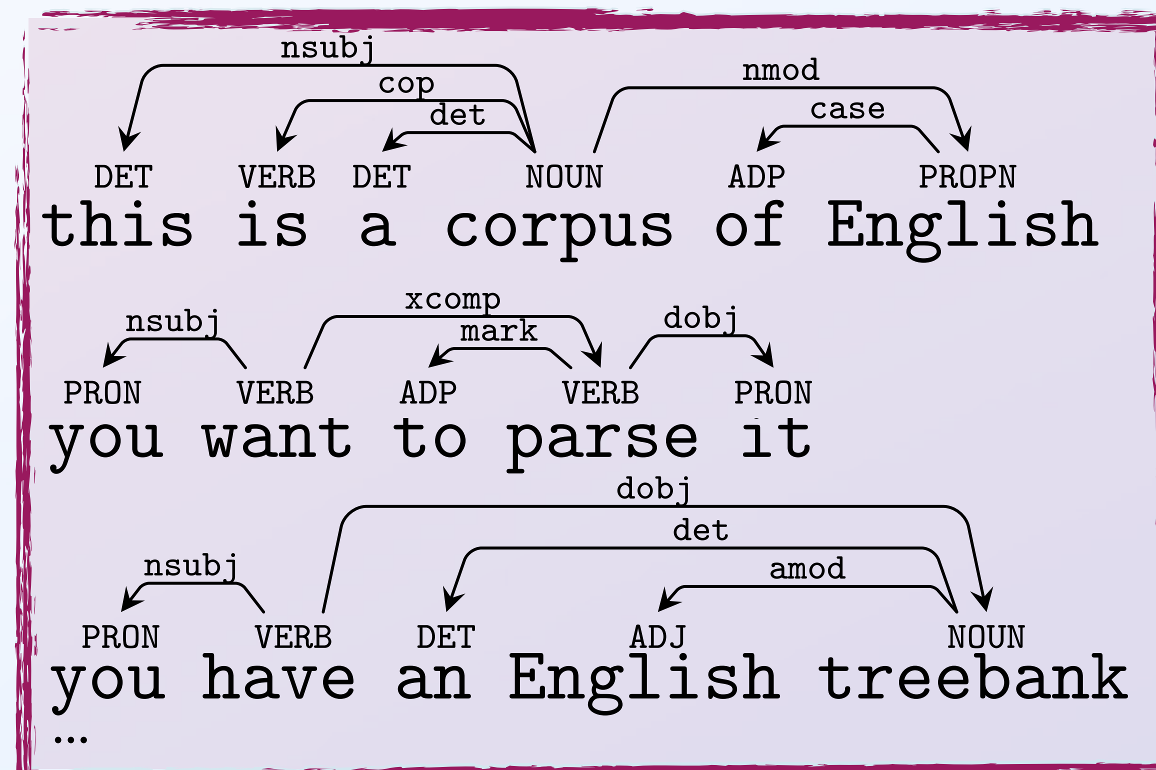
Dependency Parsing



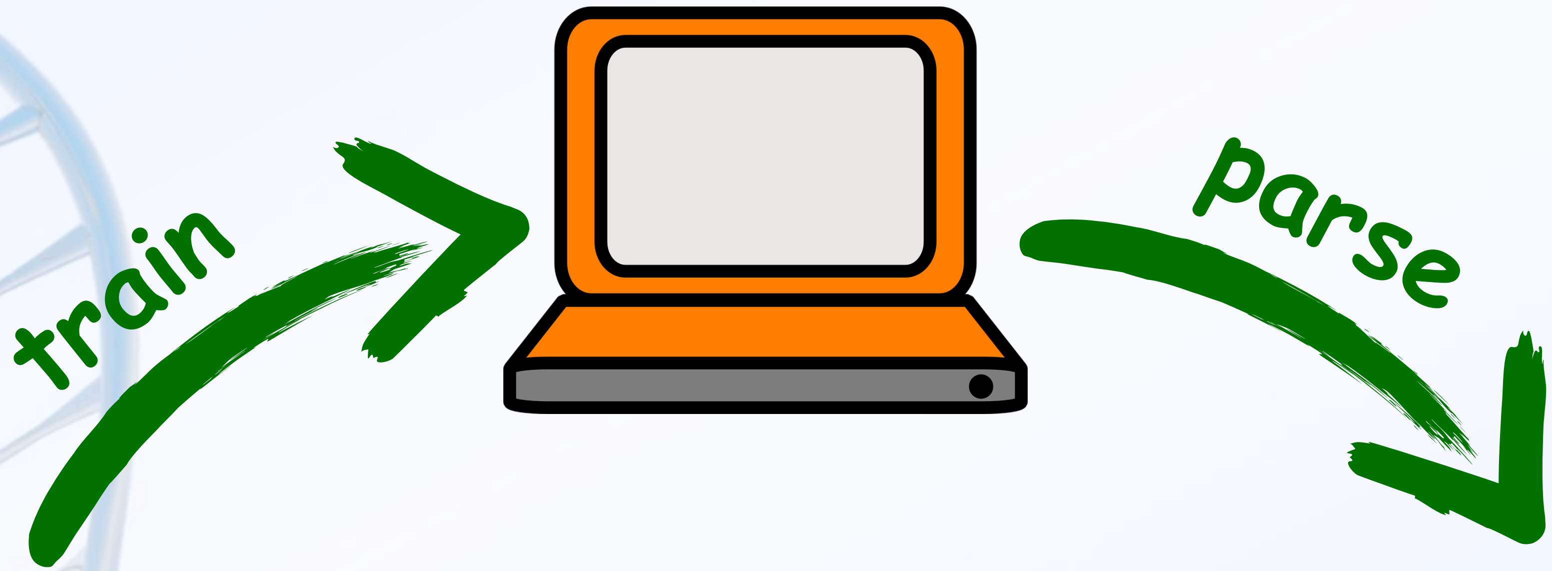
English Treebank



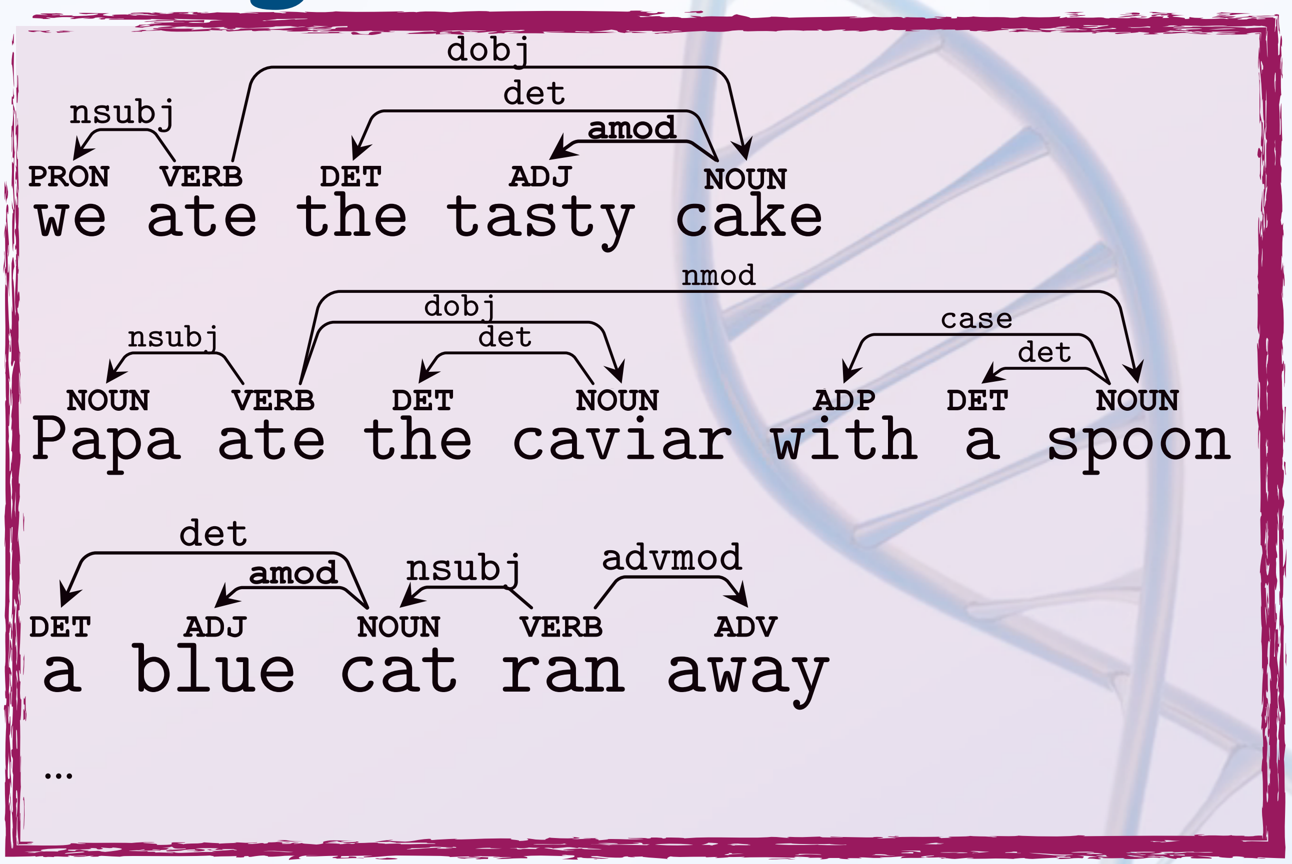
English Corpus



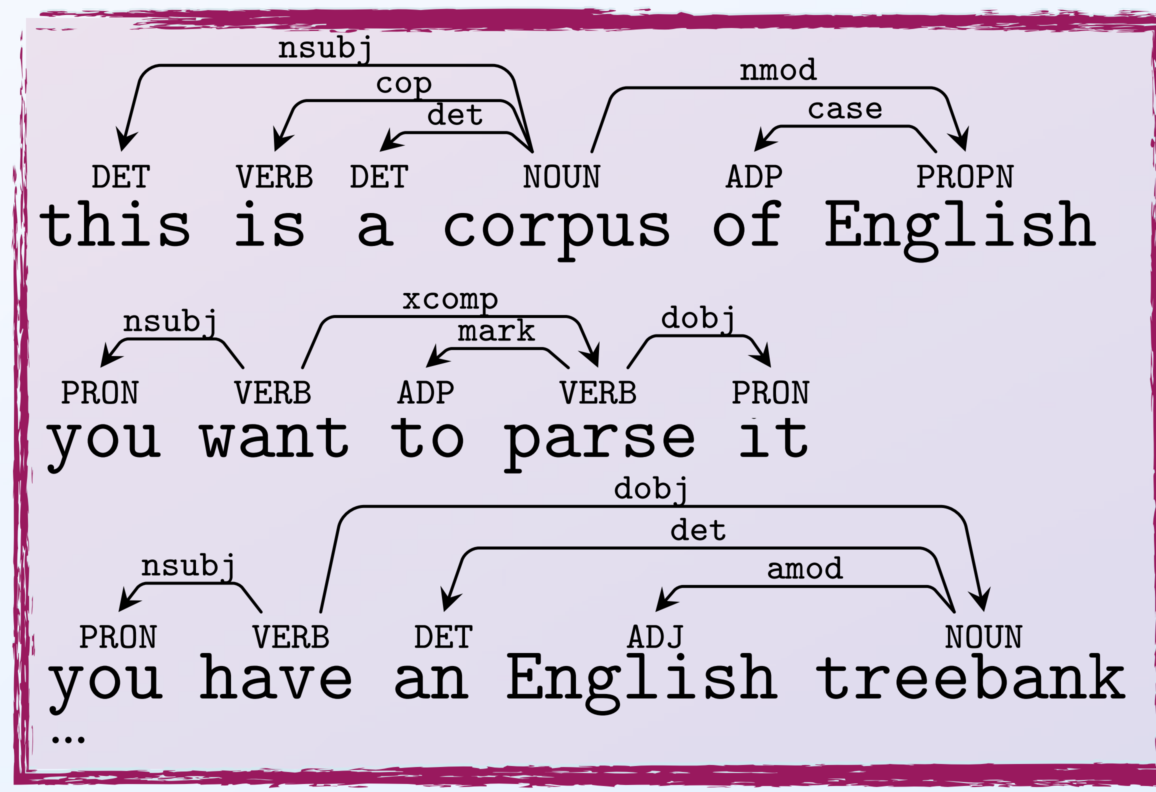
Dependency Parsing



English Treebank

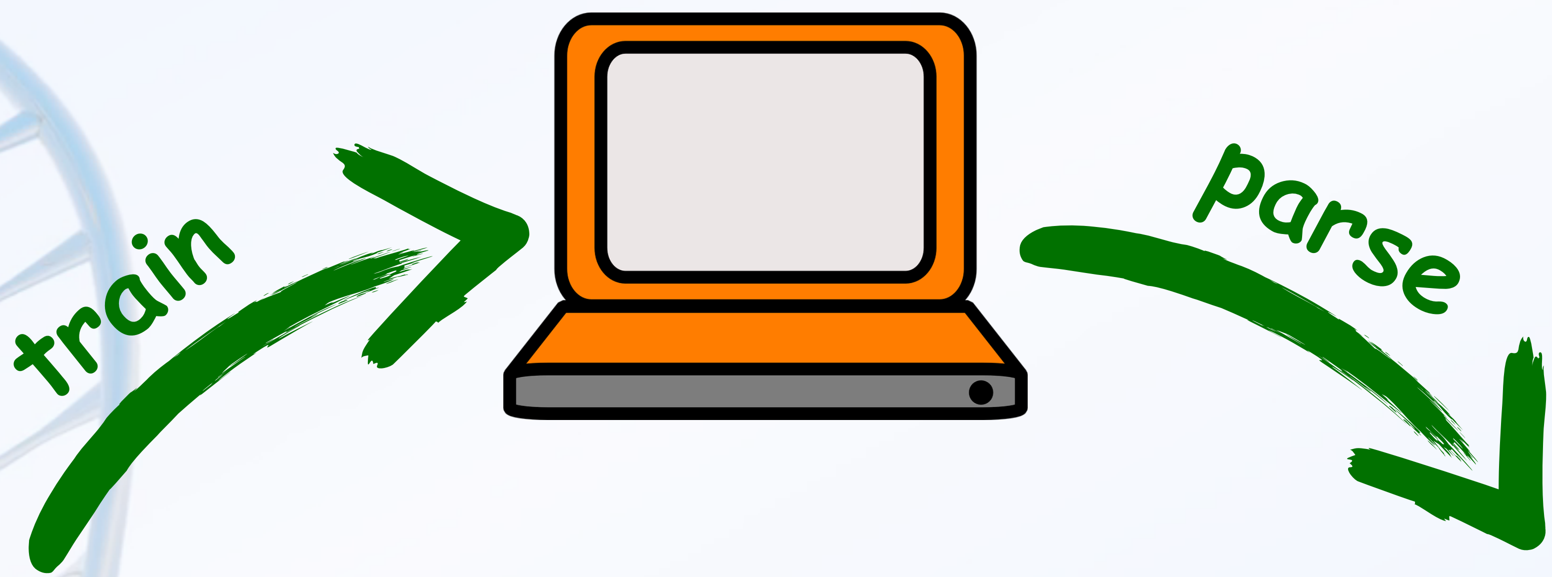


English Corpus

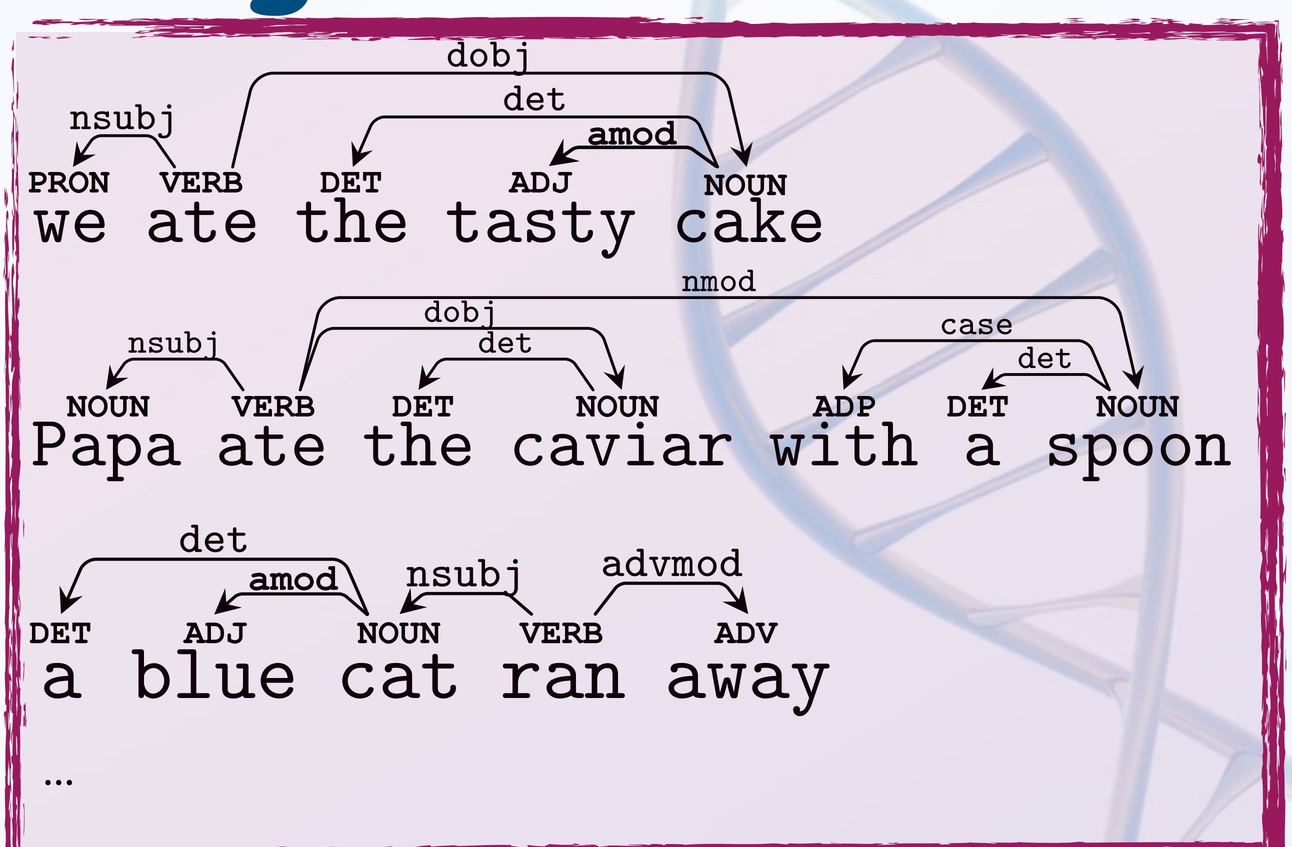


JOHNS HOPKINS UNIVERSITY

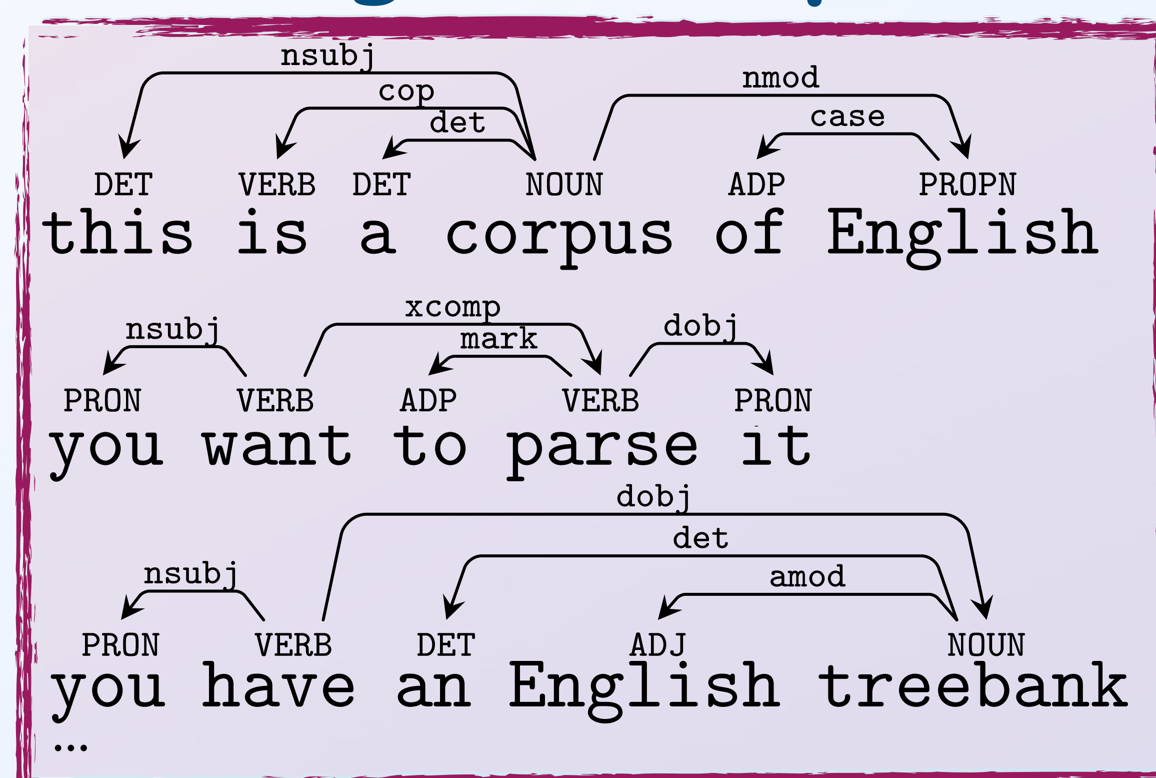
Supervised Dependency Parsing



English Treebank



English Corpus



Dependency Parsing

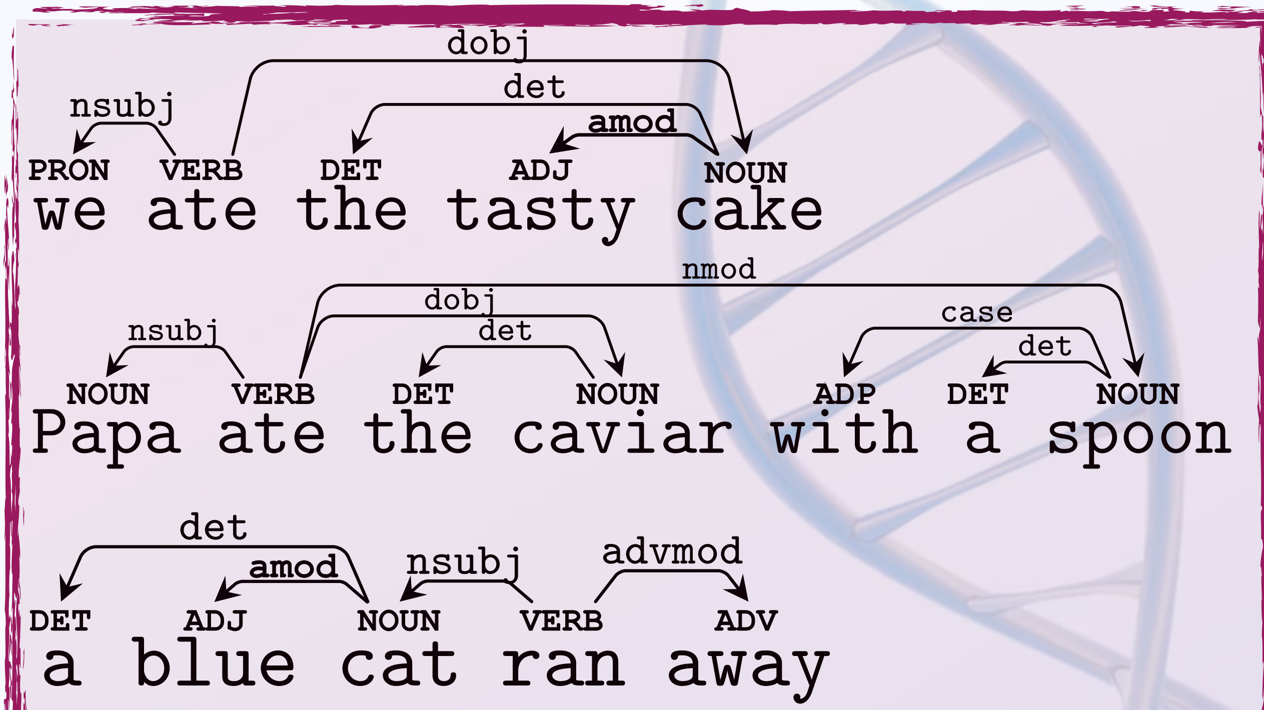
Dependency Parsing

French Corpus

```
Ma mère s'appelle Emilie Summer

Lundi, je retourne à l'école

C'est ma meilleure amie

J'aime beaucoup l'école

…
```

Dependency Parsing

English Treebank



…

French Corpus

Ma mère s'appelle Emilie Summer

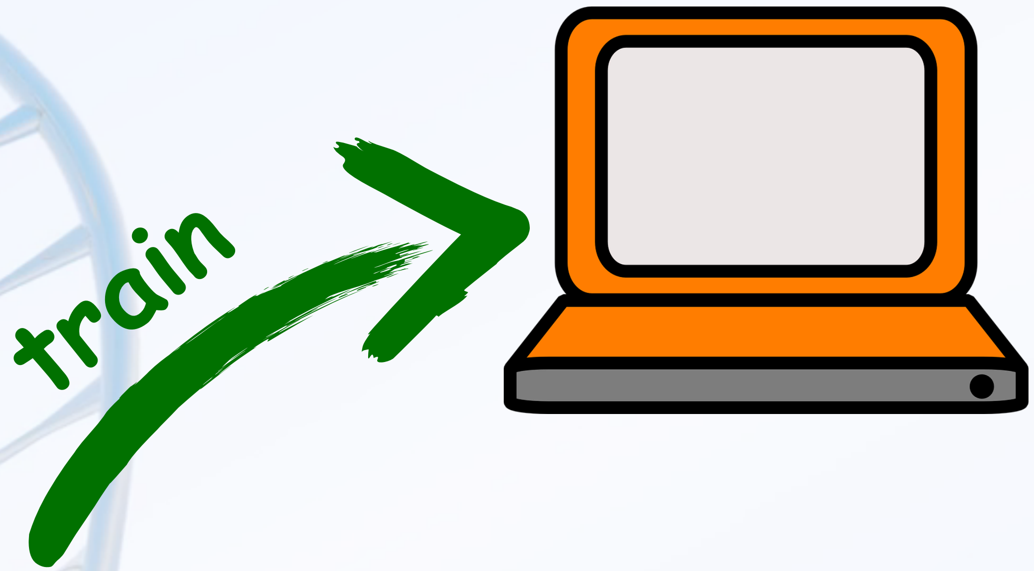
Lundi, je retourne à l’école

C’est ma meilleure amie

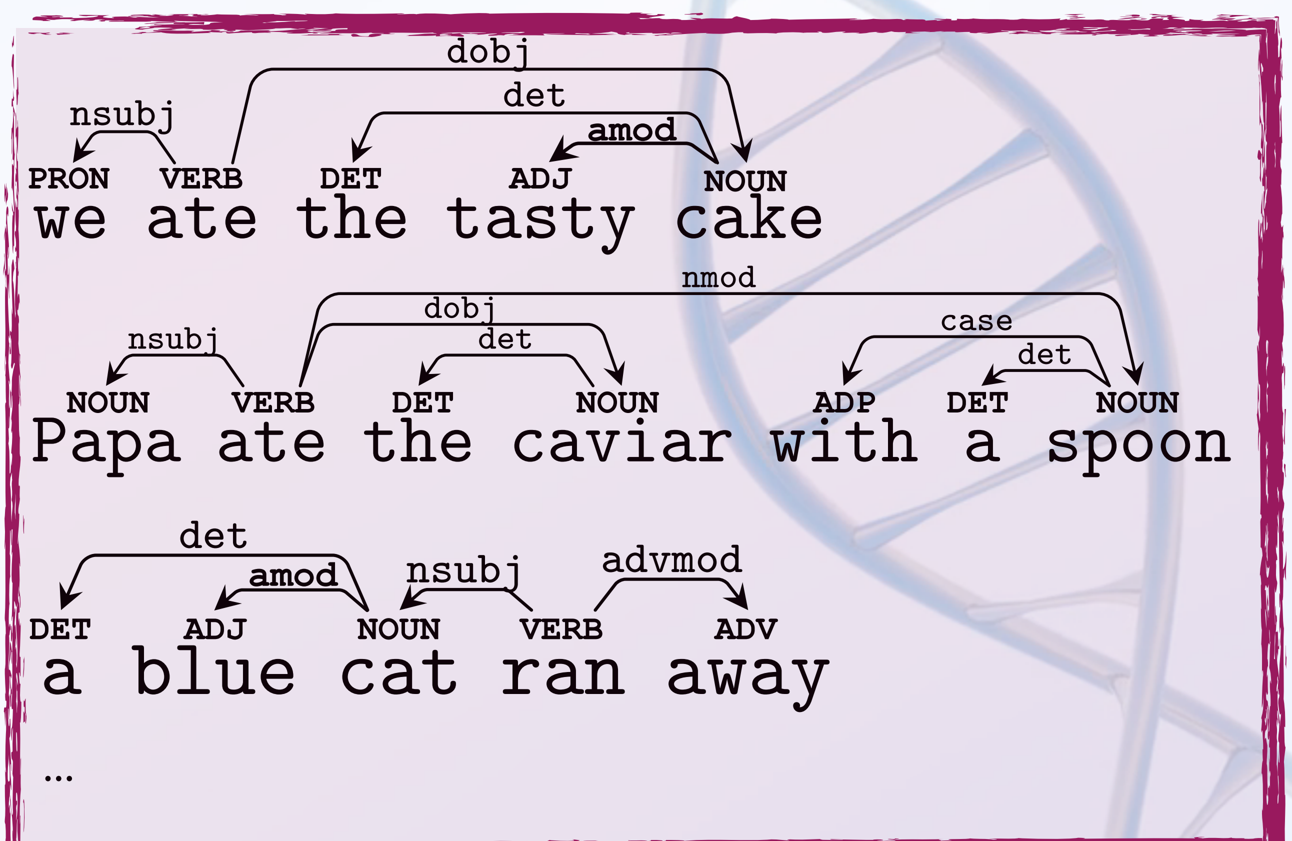
J’aime beaucoup l’école

…

Dependency Parsing



English Treebank



French Corpus

Ma mère s'appelle Emilie Summer

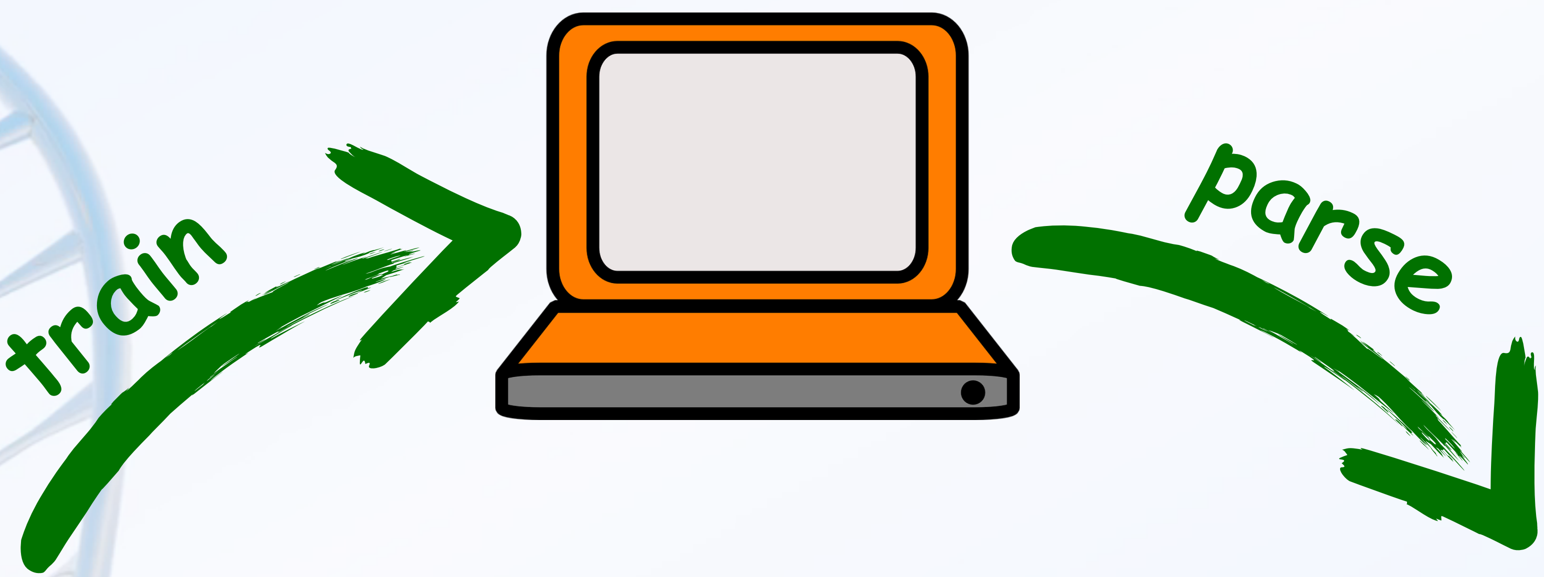
Lundi, je retourne à l’école

C’est ma meilleure amie

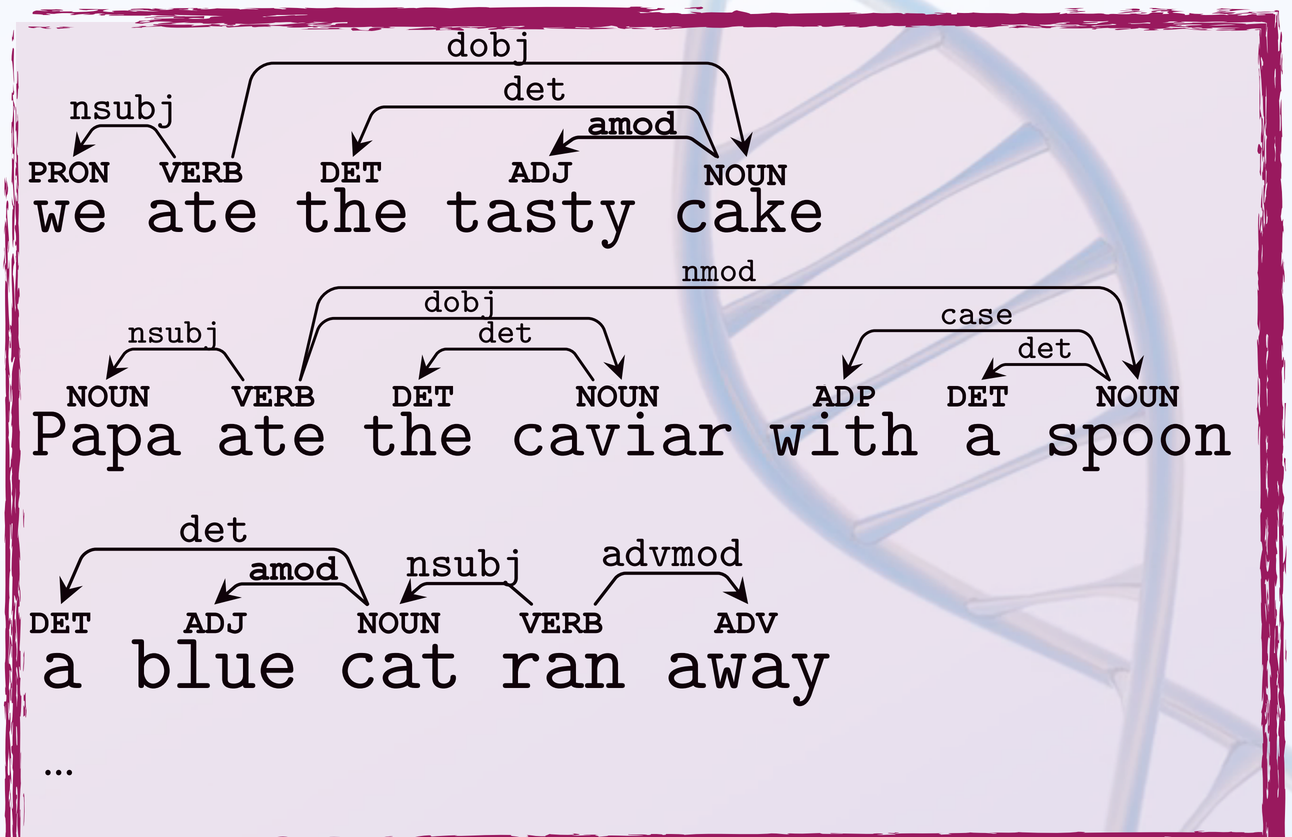
J’aime beaucoup l’école

…

Dependency Parsing



English Treebank



French Corpus

Ma mère s'appelle Emilie Summer

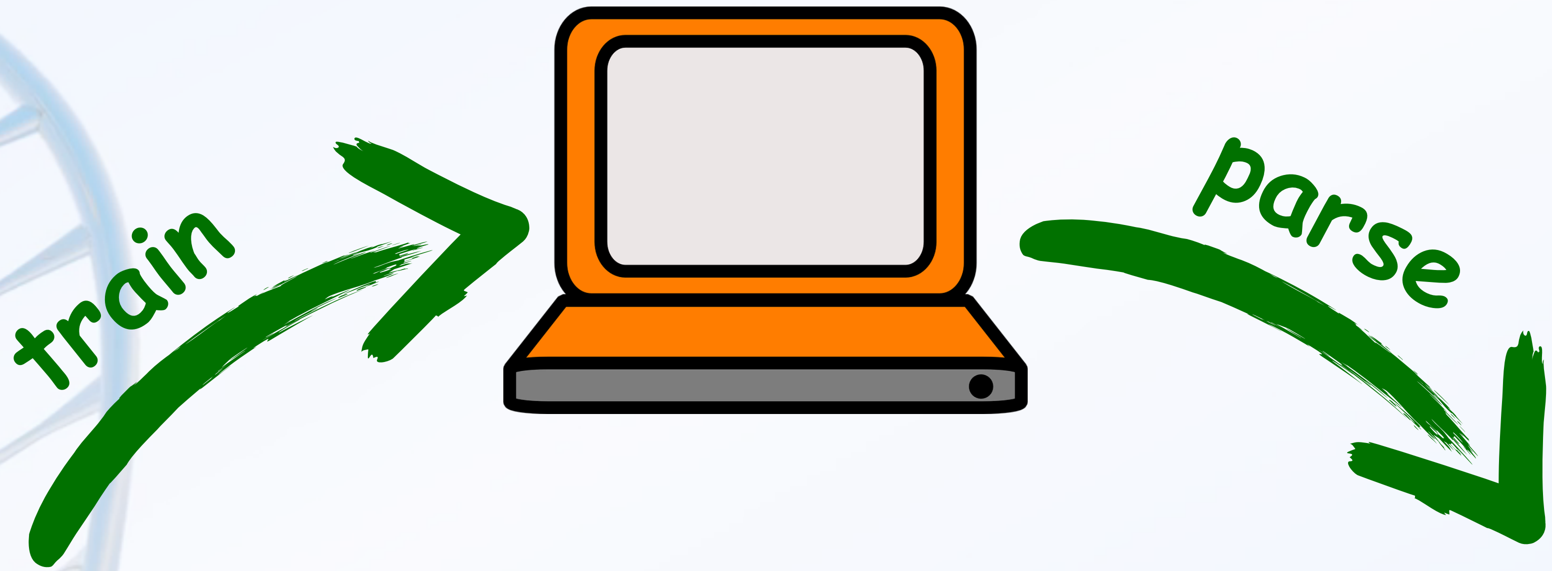
Lundi, je retourne à l’école

C’est ma meilleure amie

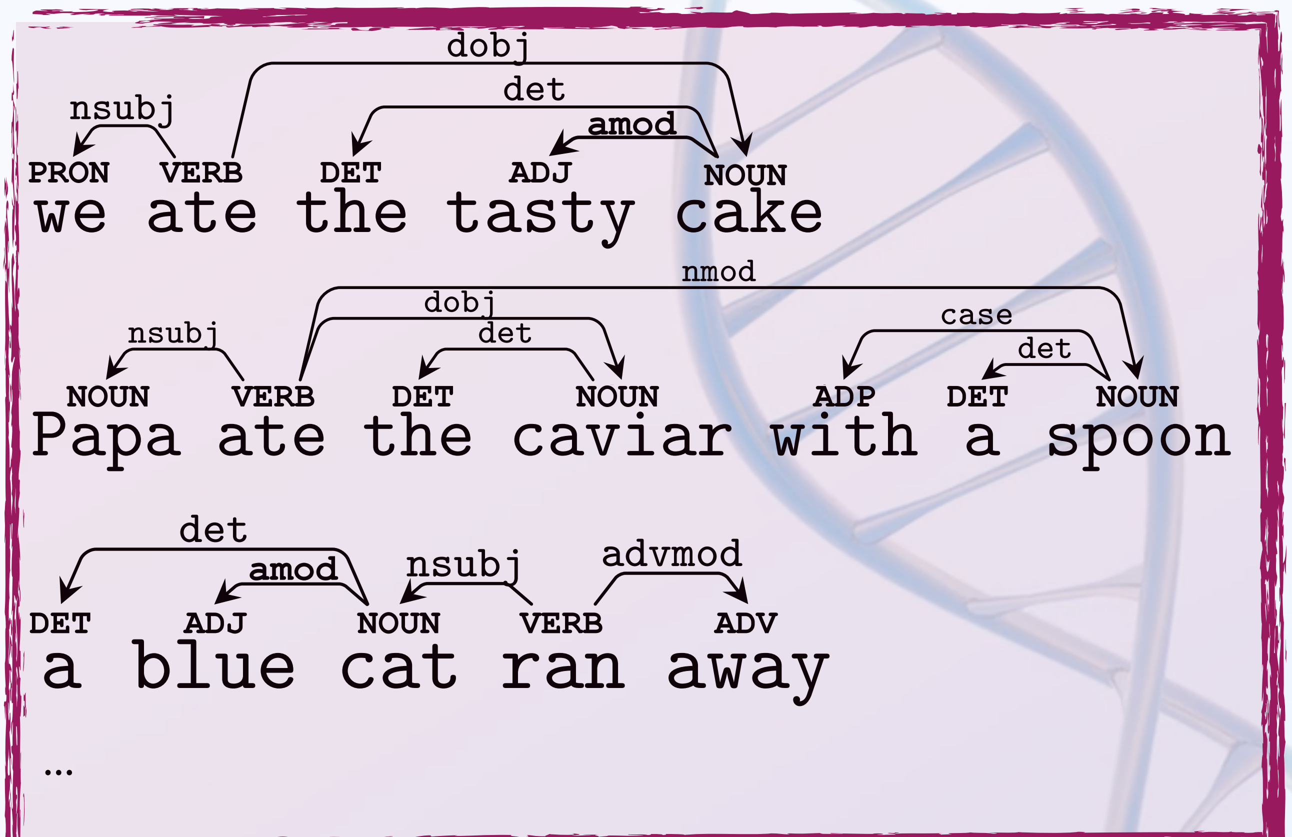
J’aime beaucoup l’école

…

Transfer Dependency Parsing



English Treebank



French Corpus

Ma mère s'appelle Emilie Summer

Lundi, je retourne à l’école

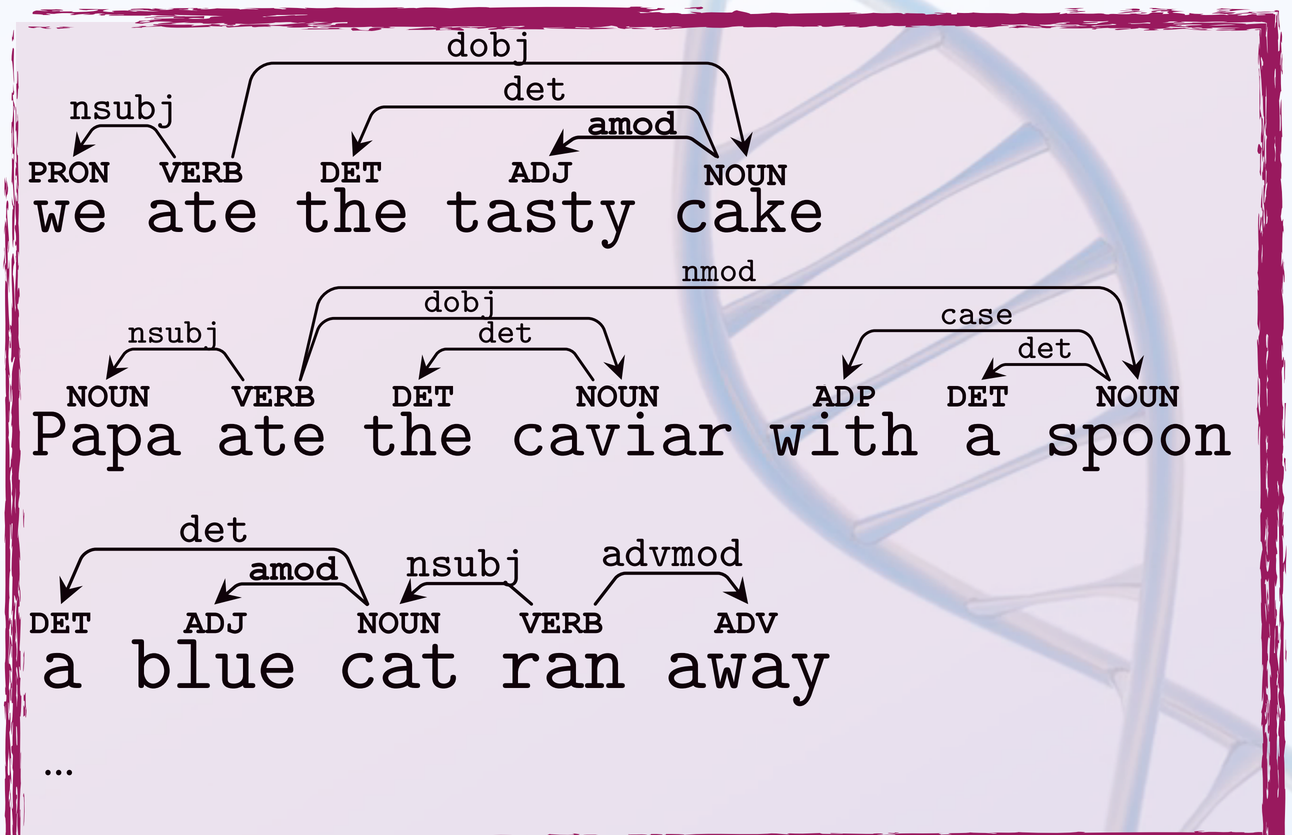
C’est ma meilleure amie

J’aime beaucoup l’école

…

Transfer Dependency Parsing

English Treebank



French Corpus

Ma mère s'appelle Emilie Summer

Lundi, je retourne à l'école

C'est ma meilleure amie

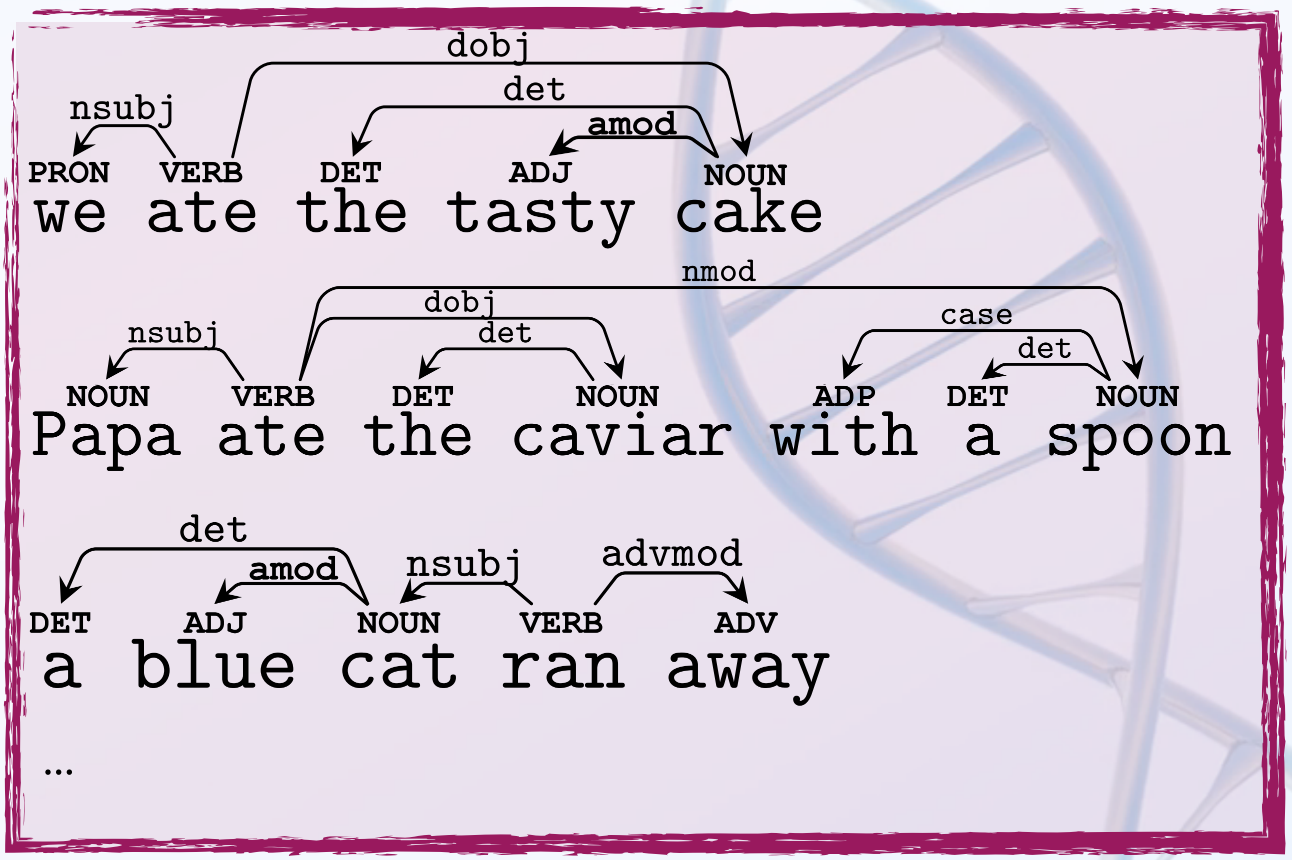
J'aime beaucoup l'école

…

JOHNS HOPKINS UNIVERSITY

Transfer Dependency Parsing

English Treebank



French Corpus

Ma mère s'appelle Emilie Summer

Lundi, je retourne à l'école

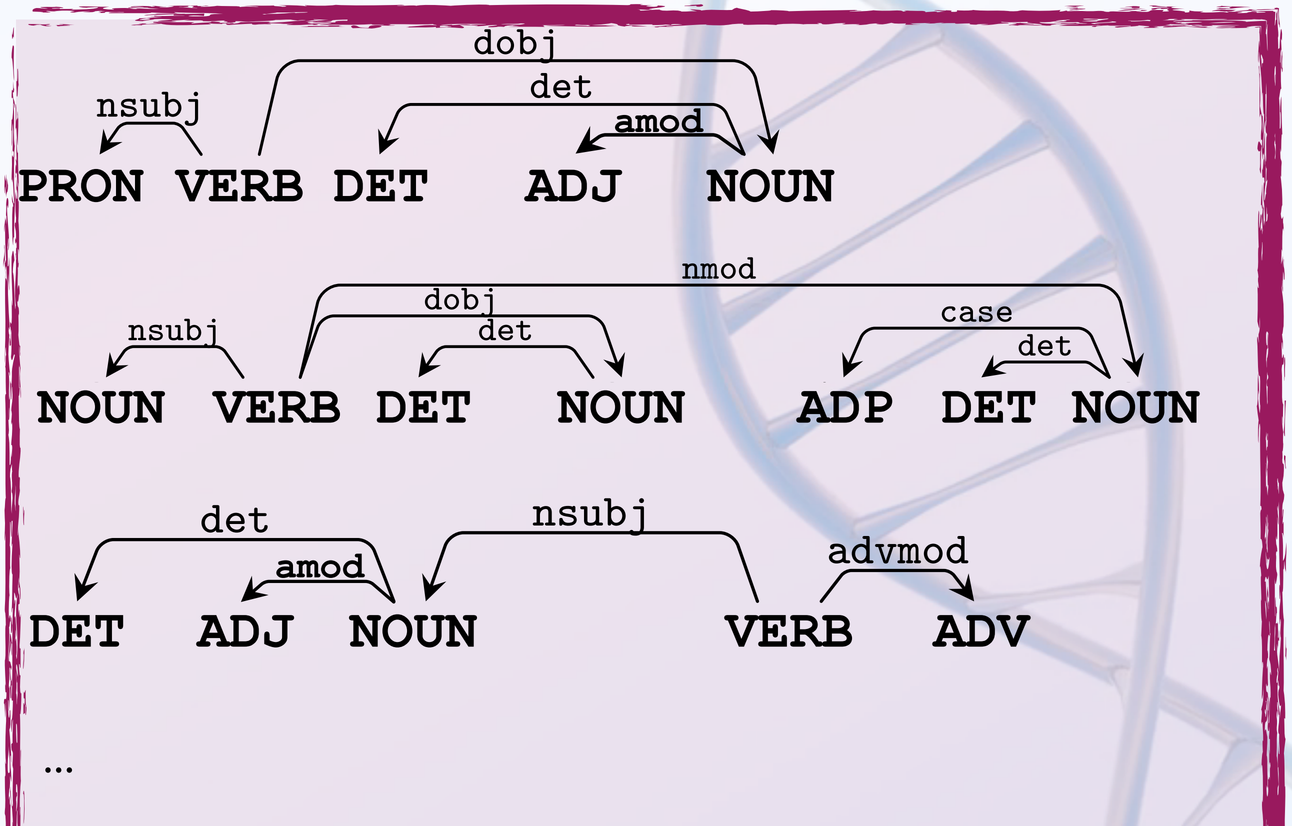
C'est ma meilleure amie

J'aime beaucoup l'école

…

Delexicalized Transfer Parsing

English Delex Treebank



French POS Corpus

NOUN VERB DET NOUN ADJ ADP NOUN

NOUN VERB PART NOUN

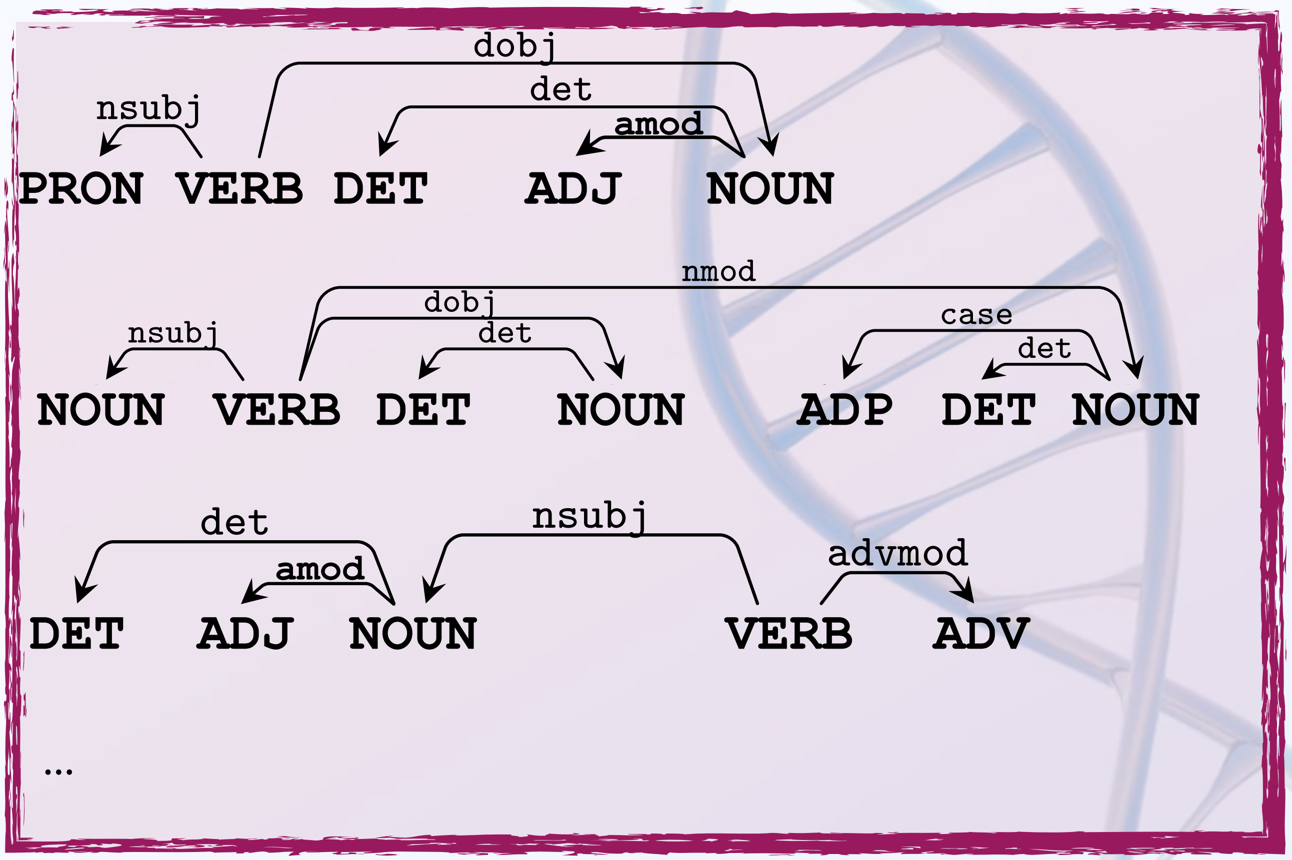
DET NOUN ADJ VERB

PRON VERB ADP DET NOUN

…

Delexicalized Transfer Parsing

English Delex Treebank



French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

NOUN VERB PART NOUN

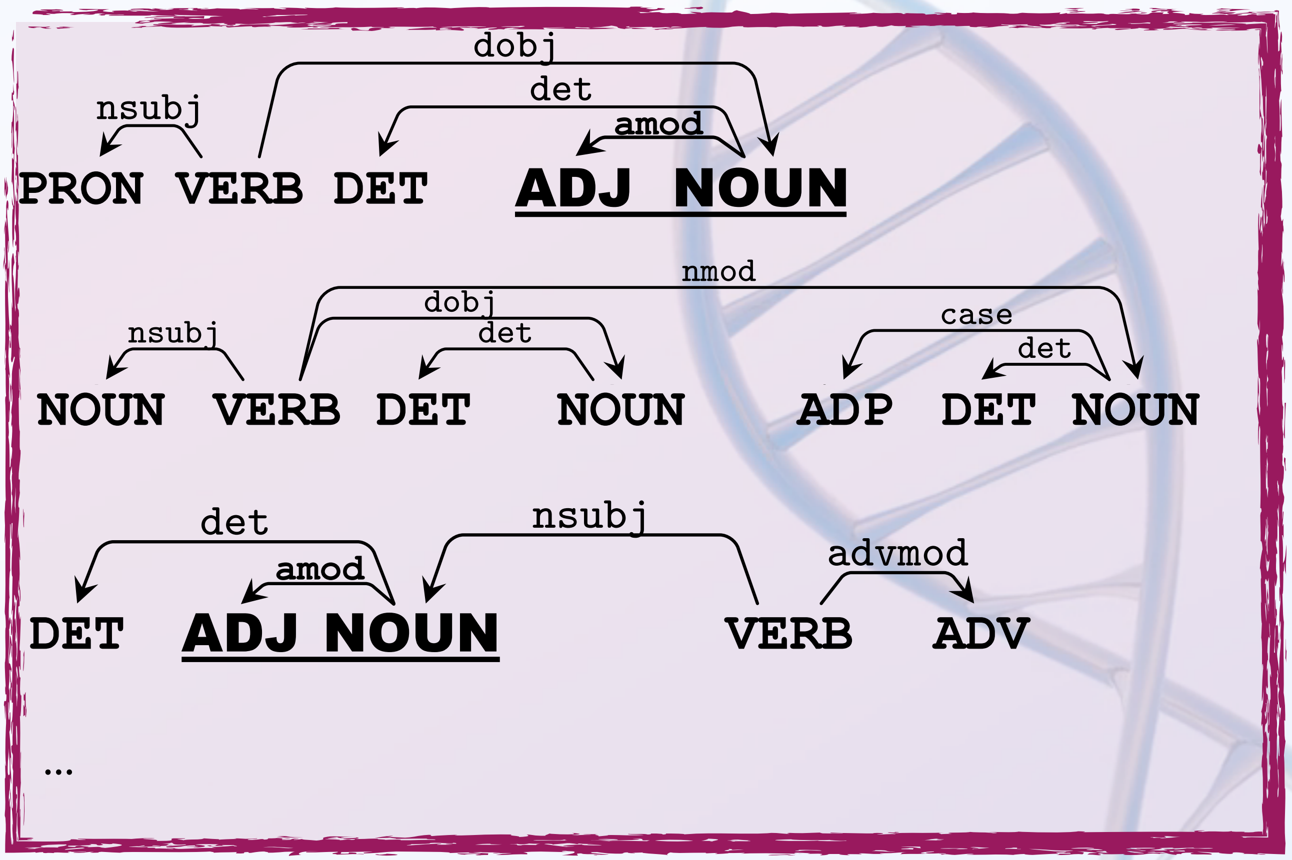
DET **NOUN ADJ** VERB

PRON VERB ADP DET NOUN

…

Delexicalized Transfer Parsing

English Delex Treebank



French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

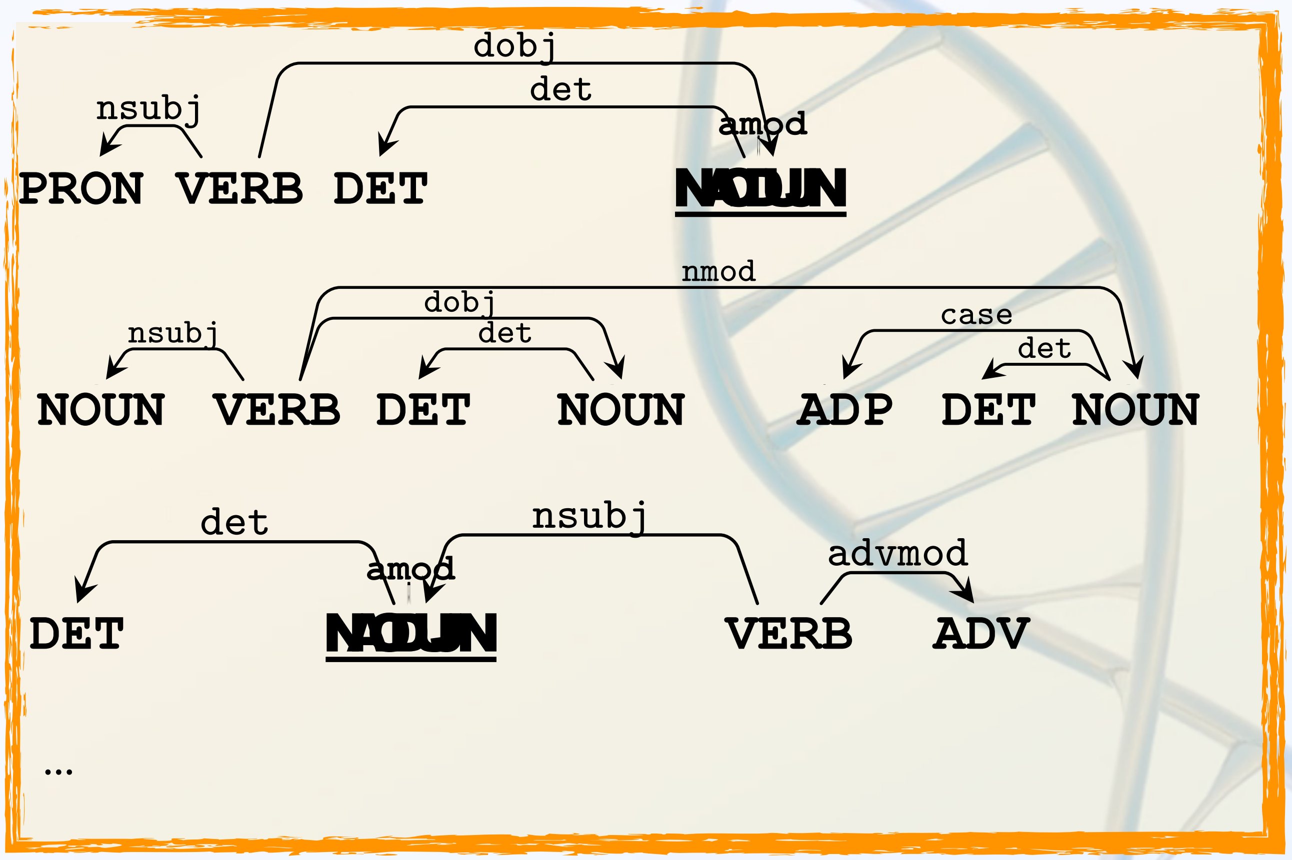
PRON VERB ADP DET NOUN

…

Johns Hopkins University

Delexicalized Transfer Parsing

English Delex Treebank



French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

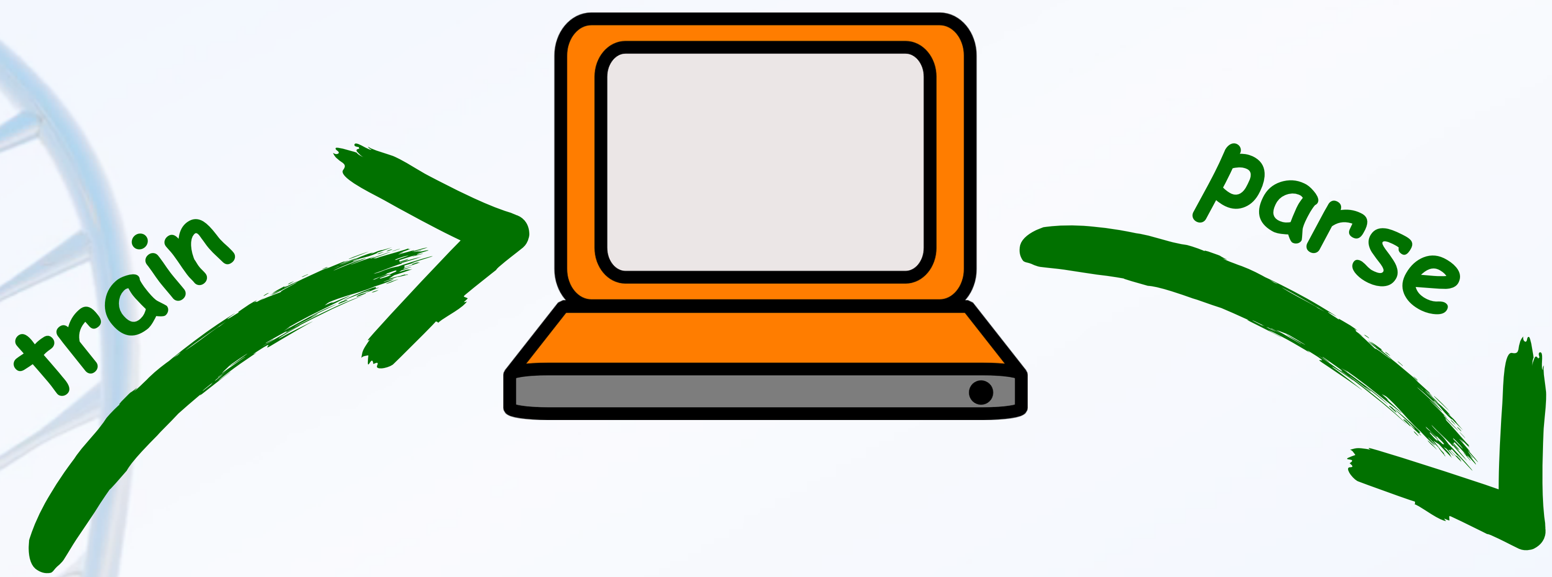
NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

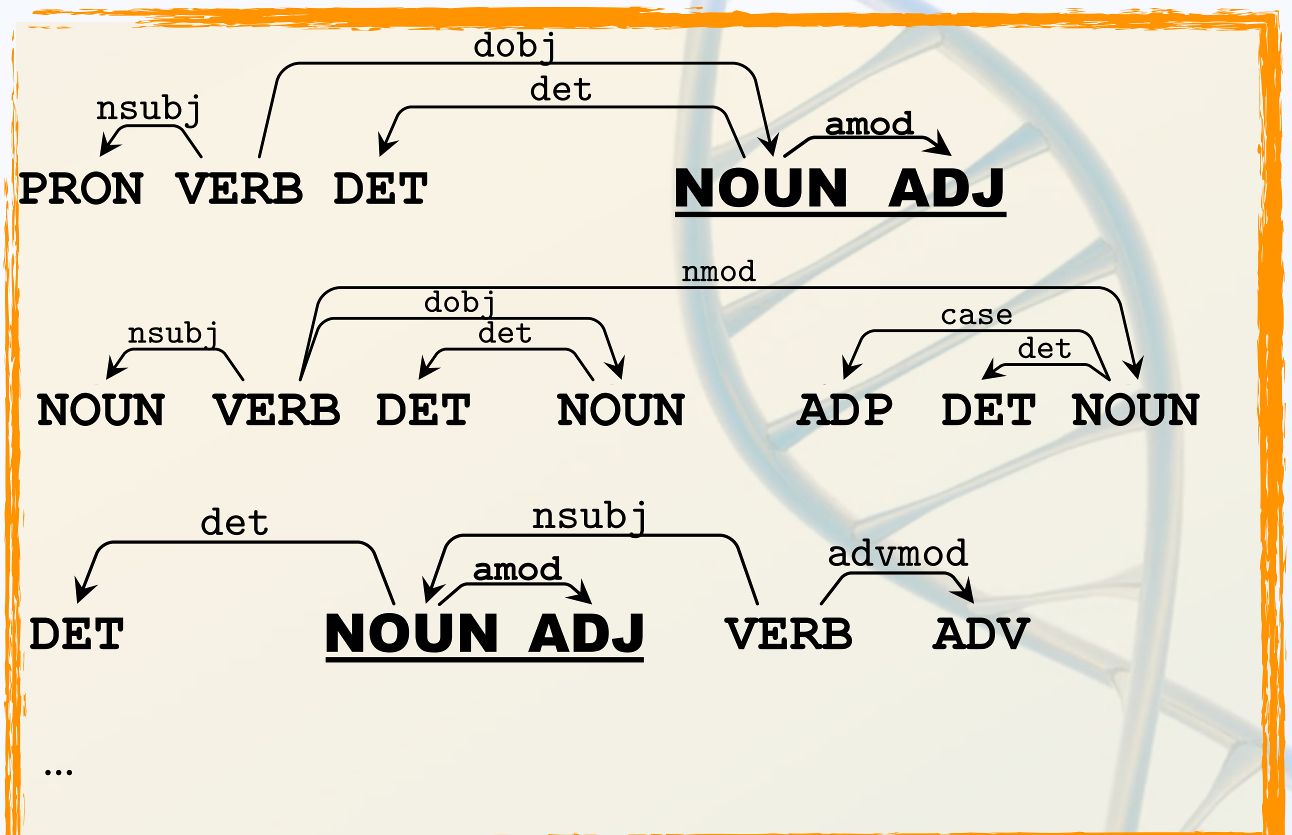
PRON VERB ADP DET NOUN

…

Delexicalized Transfer Parsing



English' Delex Treebank



French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

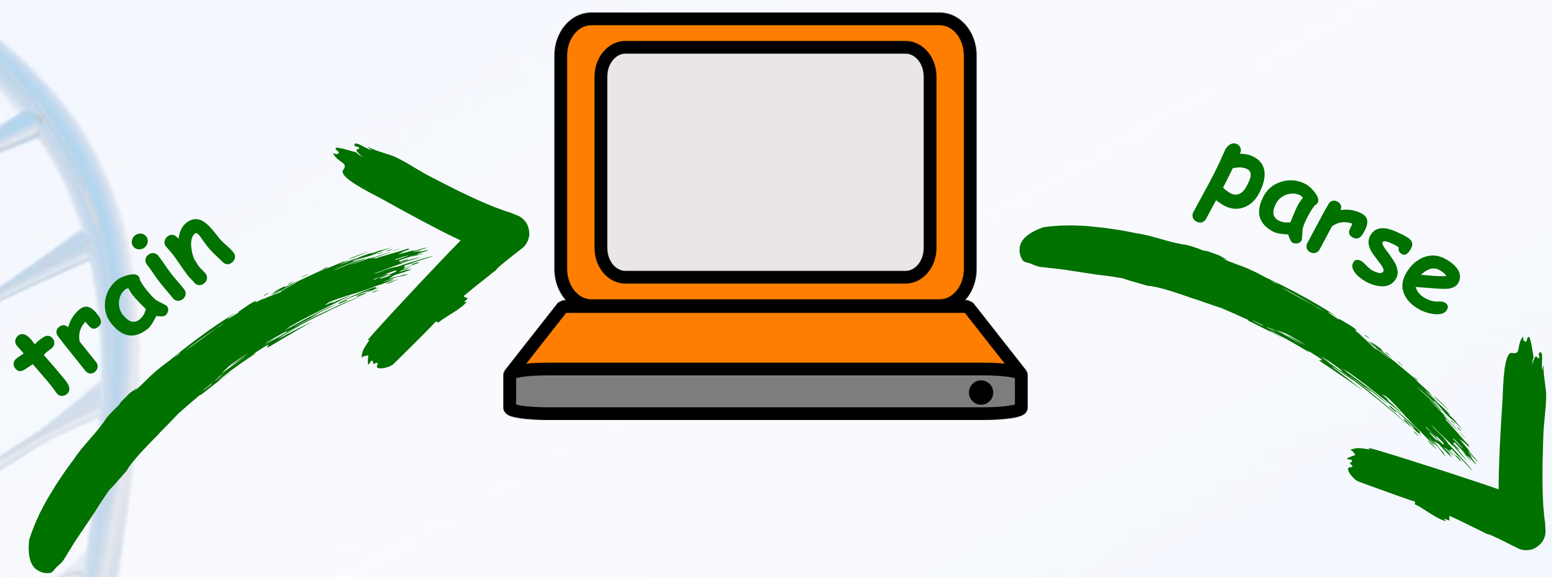
NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

PRON VERB ADP DET NOUN

…

Delexicalized Transfer Parsing



English' Delex Treebank

nsubj dobj det amod
PRON VERB DET **NOUN ADJ**

nsubj nmod dobj det case det
NOUN VERB DET NOUN ADP DET NOUN

det nsubj amod advmod
DET **NOUN ADJ** VERB ADV

…

French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

NOUN VERB PART NOUN

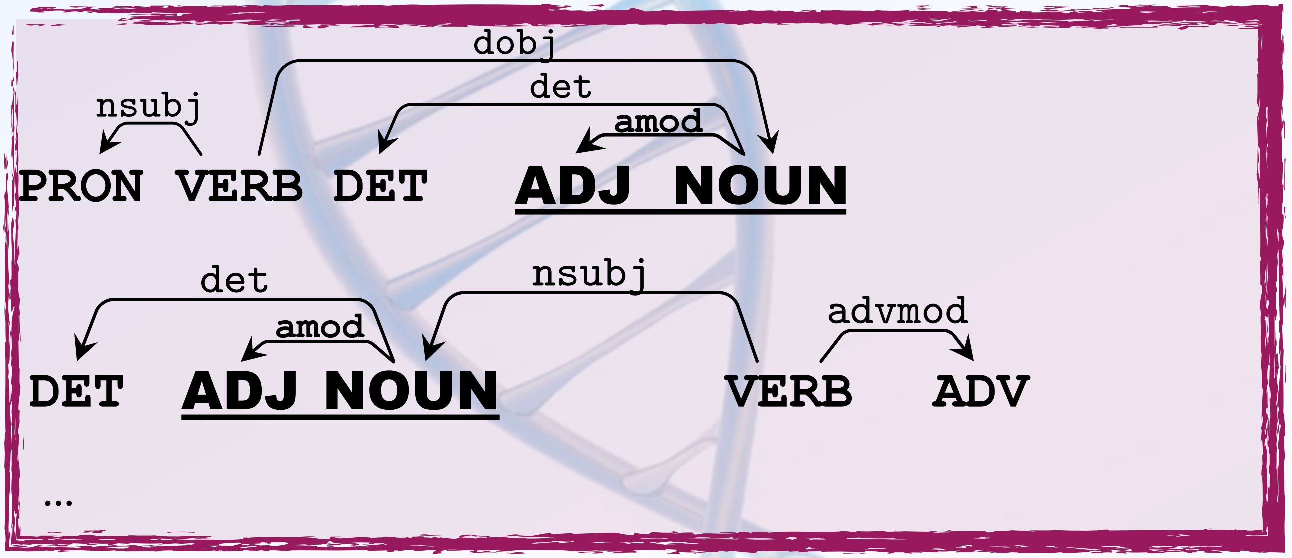
DET **NOUN ADJ** VERB

PRON VERB ADP DET NOUN

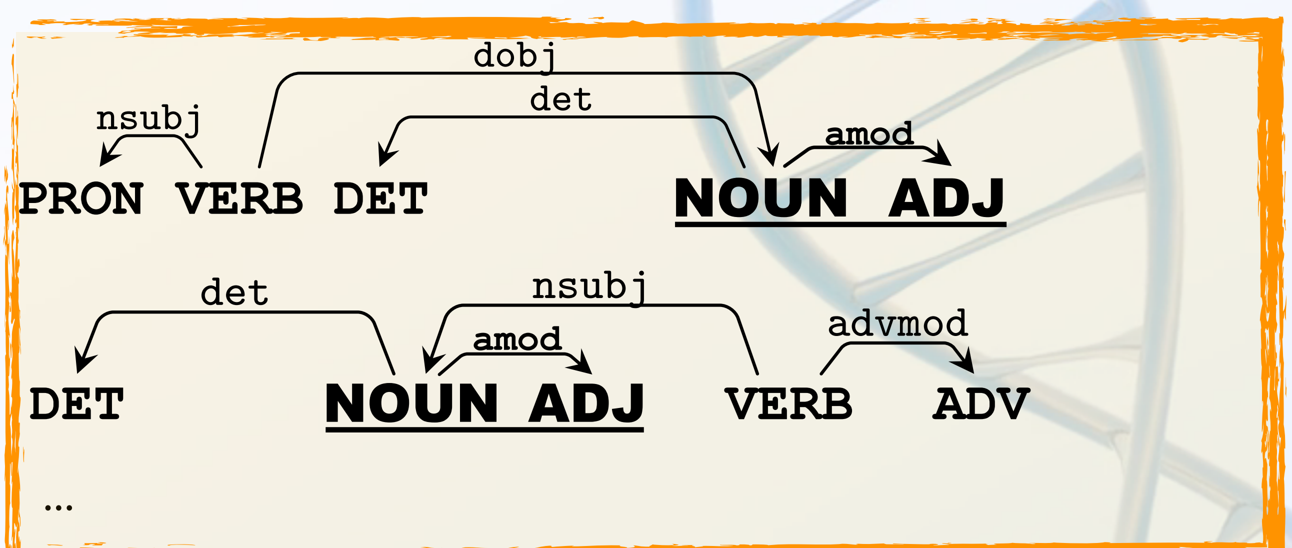
…

Improve the surface similarity

English



English'



French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

PRON VERB ADP DET NOUN

…

Improve the surface similarity

English

dobj det nsubj amod

PRON VERB DET **ADJ NOUN**

det nsubj amod advmod

DET **ADJ NOUN** VERB ADV

…

English'

dobj det nsubj amod

PRON VERB DET **NOUN ADJ**

det nsubj amod advmod

DET **NOUN ADJ** VERB ADV

…

French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

PRON VERB ADP DET NOUN

…

Improve the surface similarity

English

dobj det amod nsubj

PRON VERB DET **ADJ NOUN**

det amod nsubj advmod

DET **ADJ NOUN** VERB ADV

…

English'

dobj det amod nsubj

PRON VERB DET **NOUN ADJ**

det nsubj amod advmod

DET **NOUN ADJ** VERB ADV

…

French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

NOUN VERB PART NO

DET **NOUN ADJ** V

PRON VERB ADP I

…

Improve the surface similarity

English

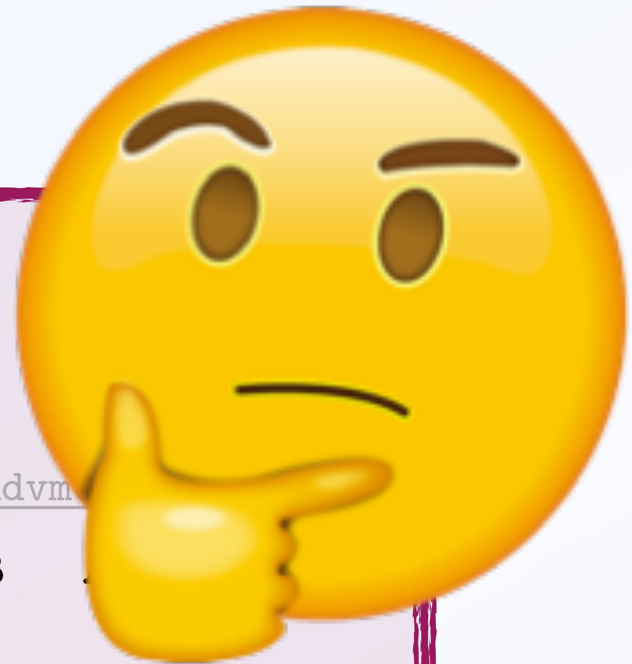
dobj det nsubj amod

PRON VERB DET **ADJ NOUN**

det nsubj amod advm

DET **ADJ NOUN** VERB

…



English'

dobj det nsubj amod

PRON VERB DET **NOUN ADJ**

det nsubj amod advmod

DET **NOUN ADJ** VERB ADV

…

French POS Corpus

NOUN VERB DET **NOUN ADJ** ADP NOUN

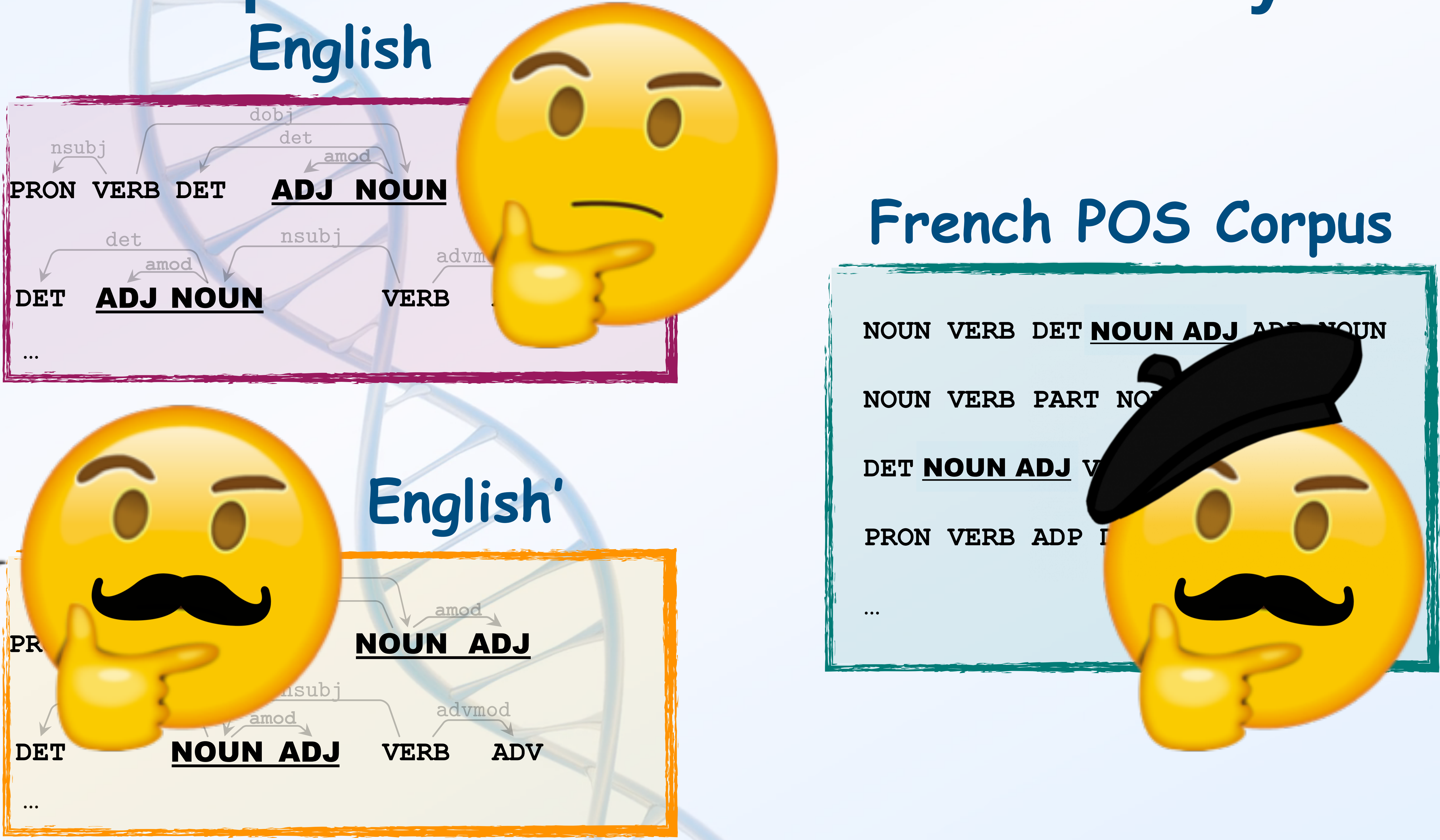
NOUN VERB PART NO

DET **NOUN ADJ** V

PRON VERB ADP I

…

Improve the surface similarity

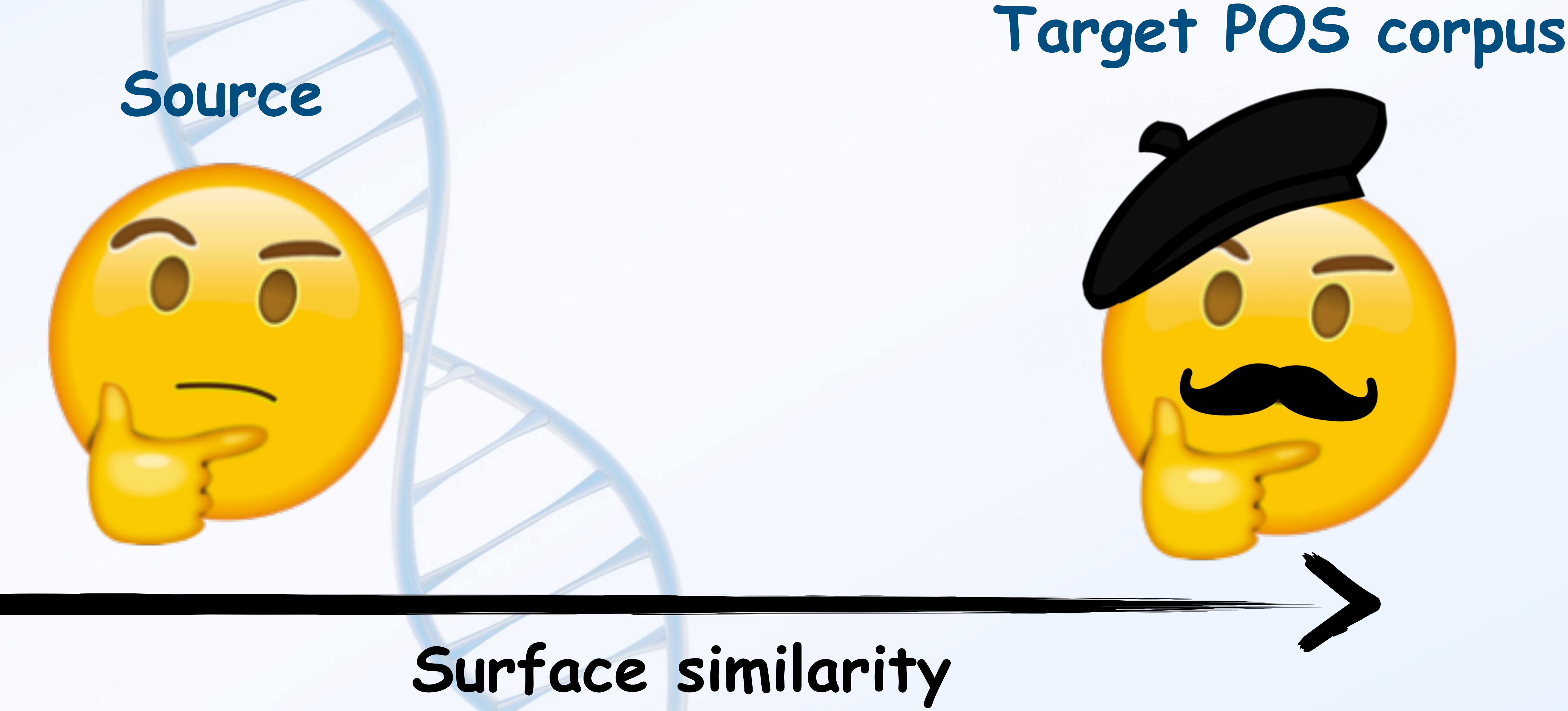


Improve the surface similarity

Source

Target POS corpus

Improve the surface similarity

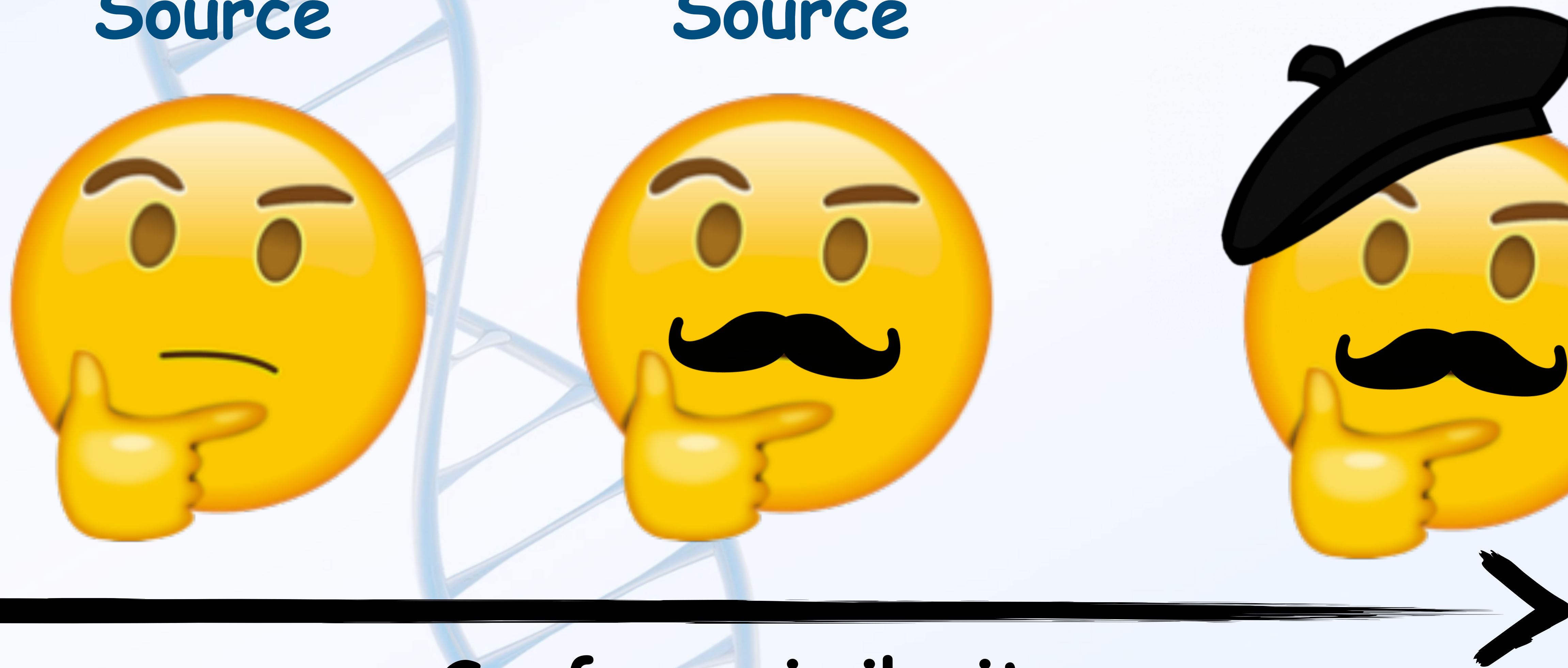


Improve the surface similarity

Source

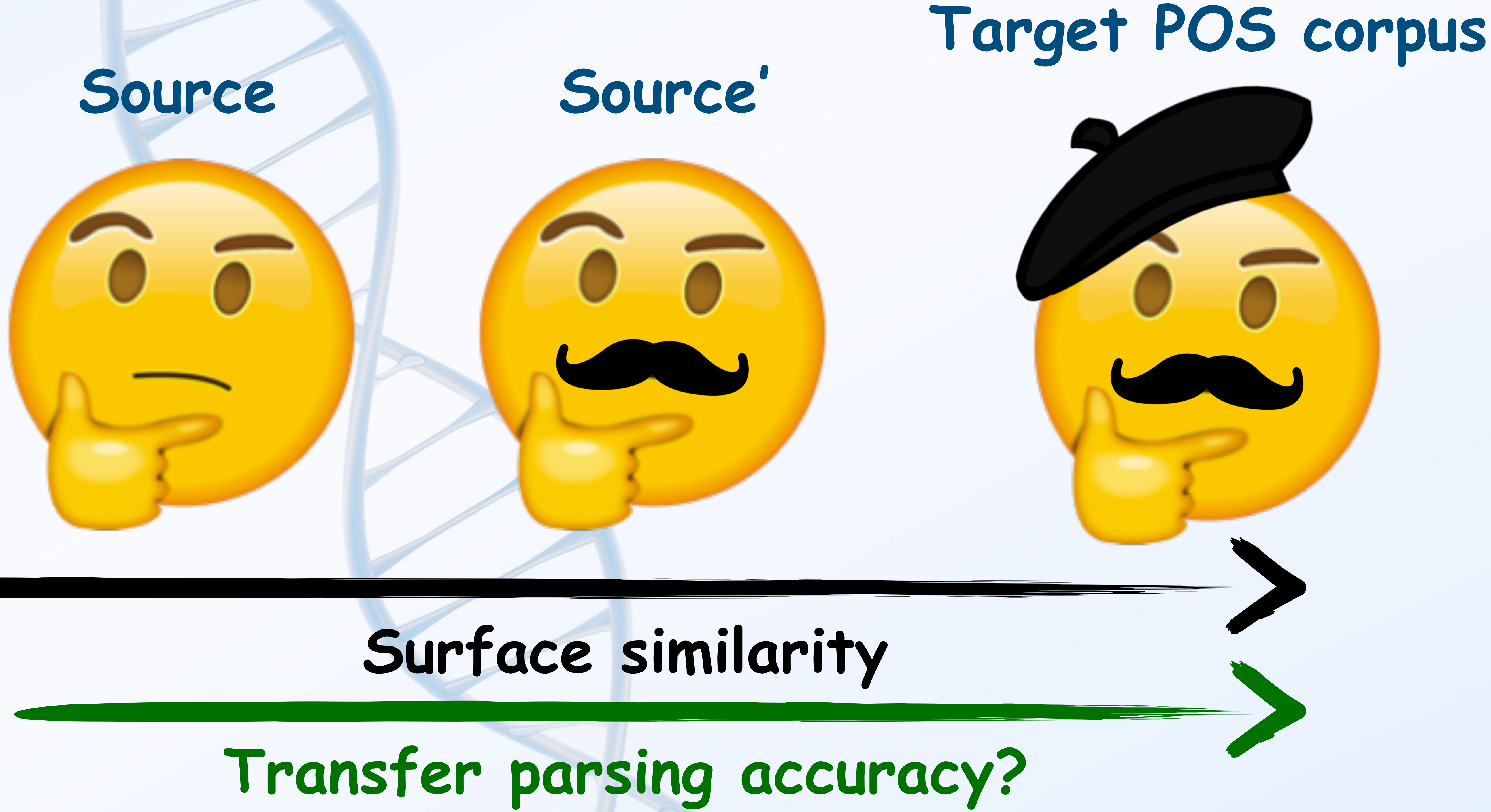
Source'

Target POS corpus

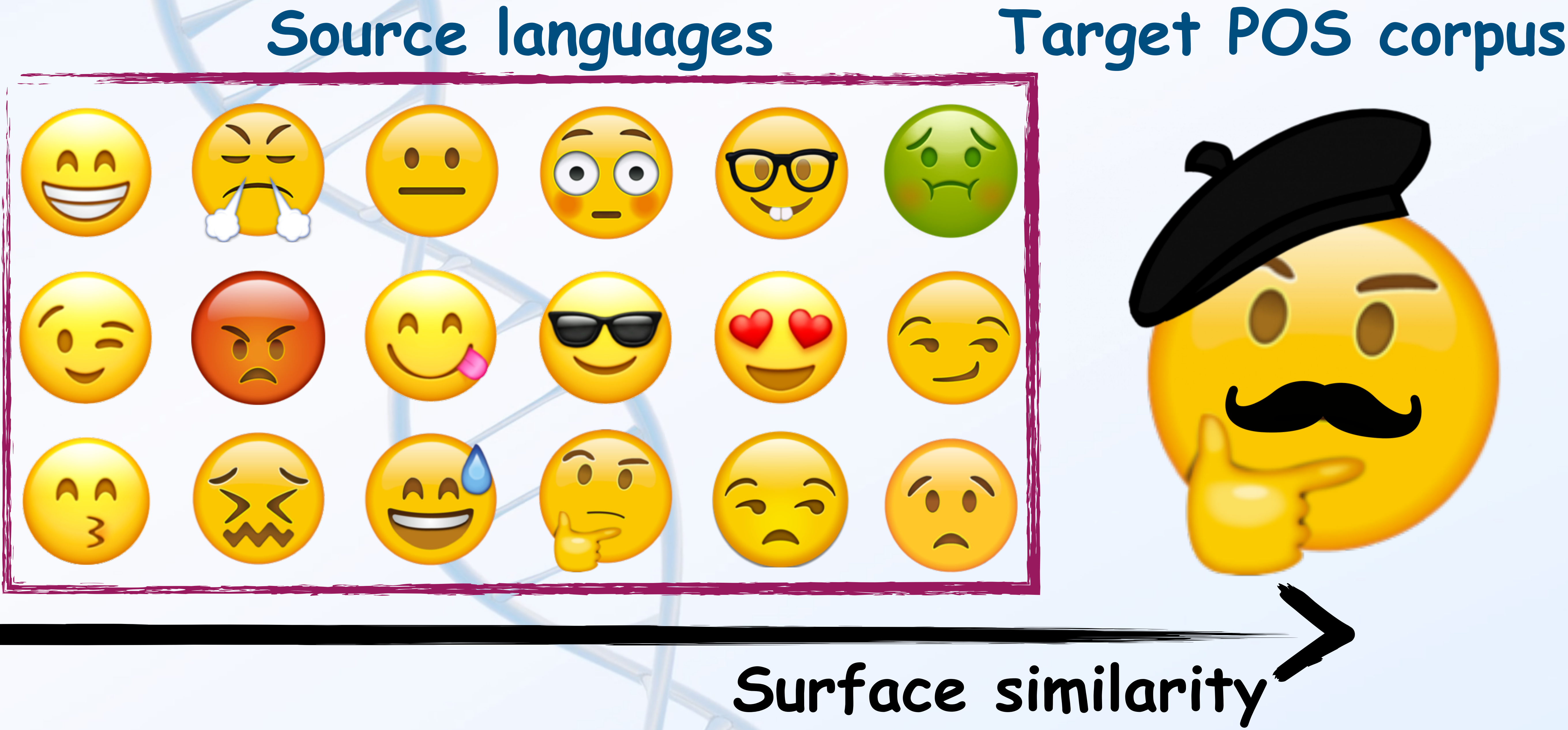


Surface similarity

Improve the surface similarity



Single-Source Selection



Source languages

Source languages

- Universal Dependencies

Source languages

- Universal Dependencies
  - Now has **146** treebanks, **83** languages

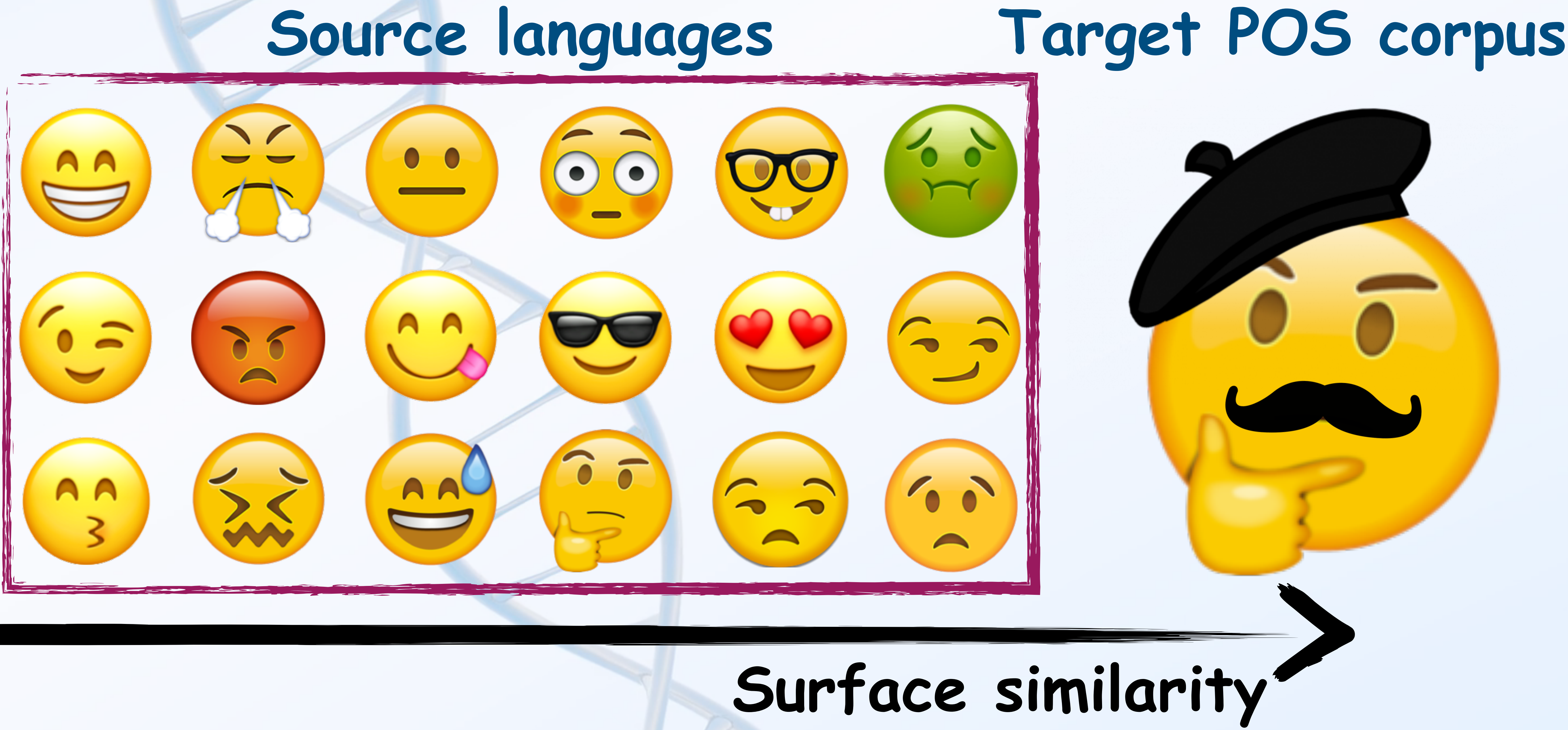
Source languages

- Universal Dependencies
  - Now has **146** treebanks, **83** languages
  - Empower the supervised methods to process these languages

Source languages

- Universal Dependencies
  - Now has **146** treebanks, **83** languages
  - Empower the supervised methods to process these languages
  - Help analyze novel languages

Single-Source Selection



Johns Hopkins University

Single-Source Selection

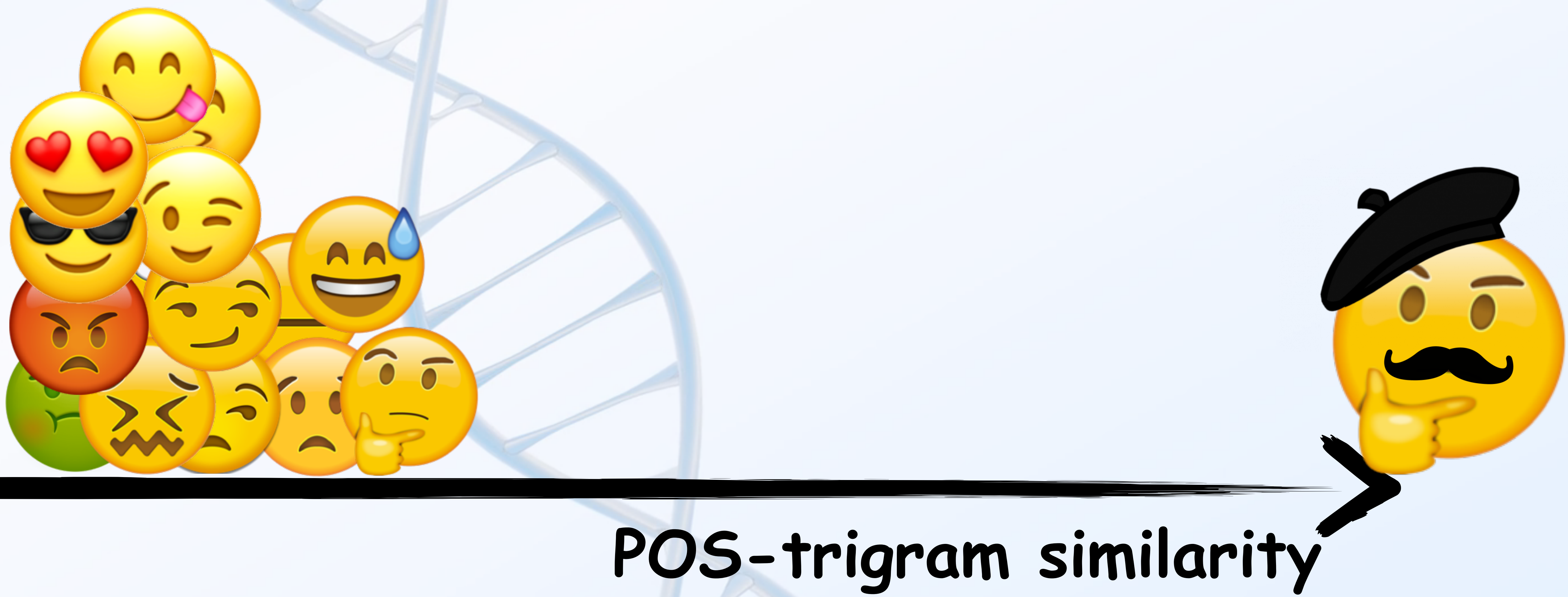
Source languages

Target POS corpus

POS-trigram similarity

Single-Source Selection

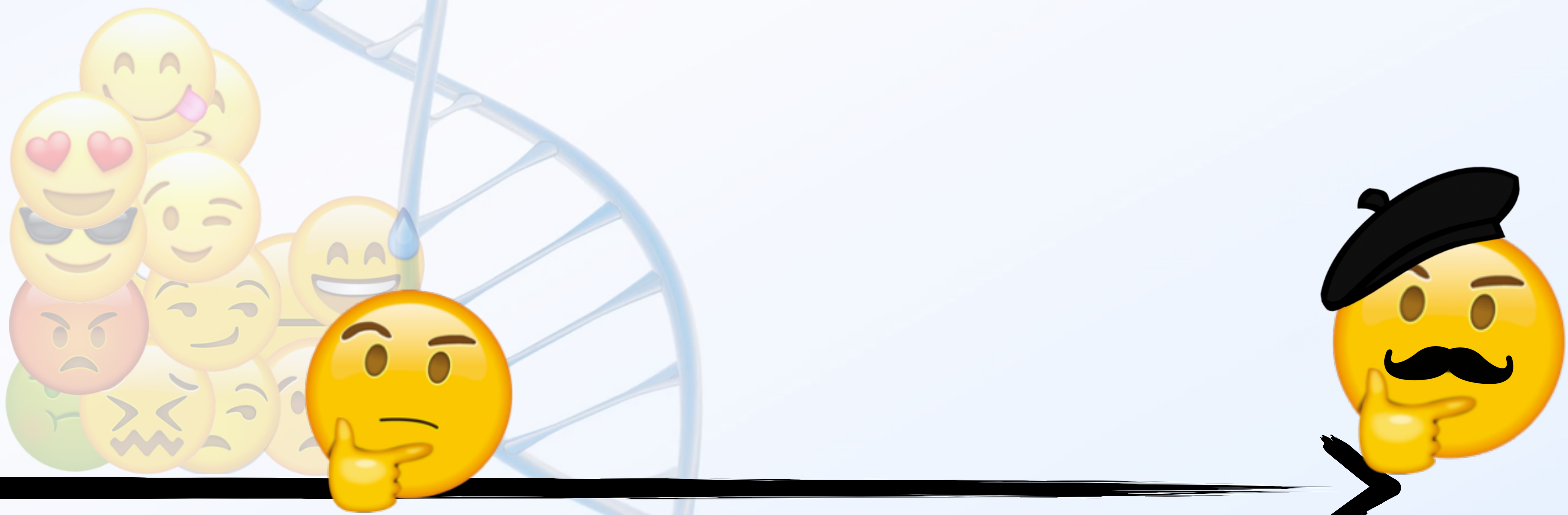
Source languages **Target POS corpus**



Single-Source Selection

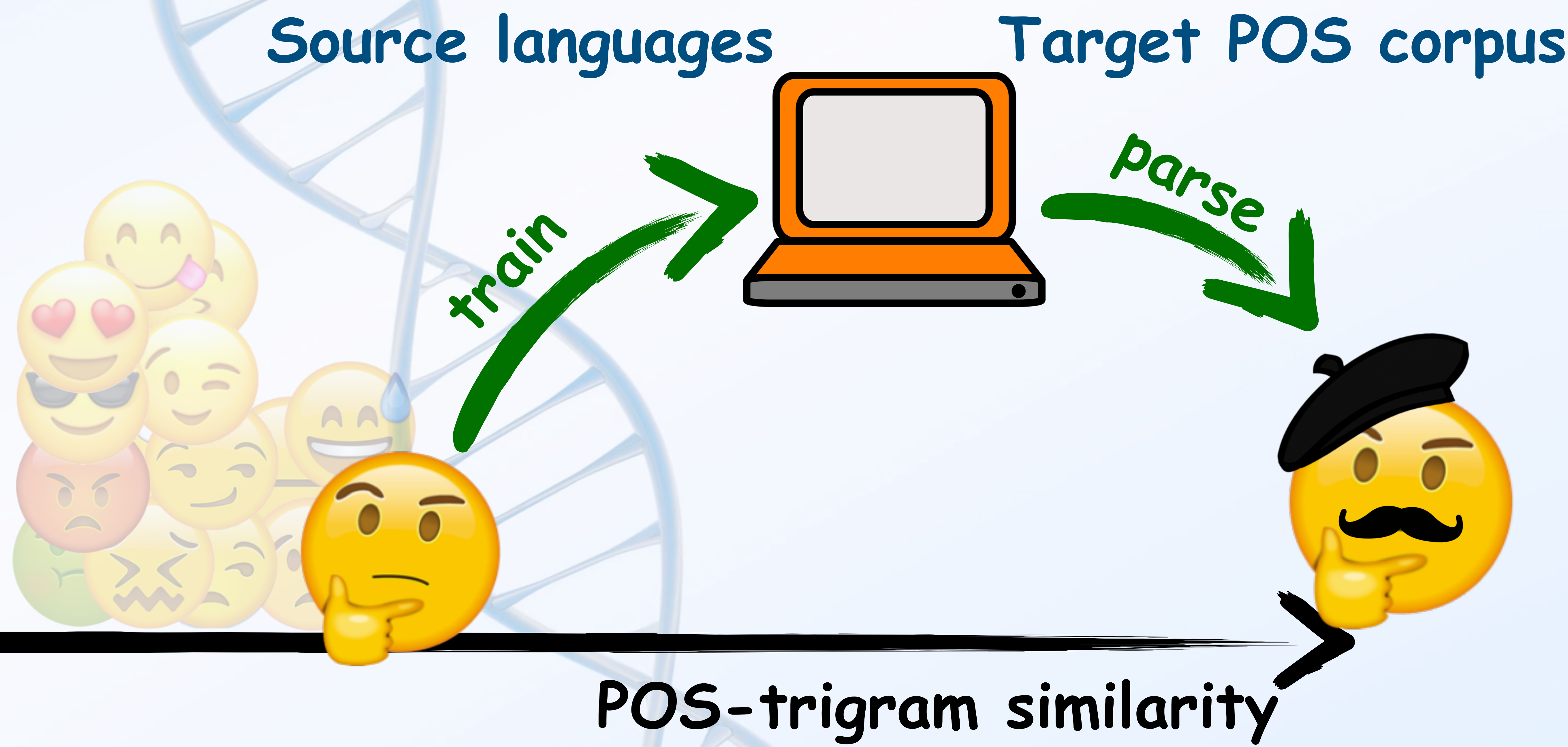
Source languages

Target POS corpus



POS-trigram similarity

Single-Source Selection



JOHNS HOPKINS
UNIVERSITY

POS-trigram similarity

Target POS corpus

NOUN VERB DET NOUN ADJ ADP NOUN

POS-trigram similarity

Target POS corpus

`p(NOUN VERB DET NOUN ADJ ADP NOUN)`

POS-trigram similarity

Target POS corpus

```
p(NOUN VERB DET NOUN ADJ ADP NOUN)
= p(NOUN | BOS BOS)
```

POS-trigram similarity

Target POS corpus

```
p(NOUN VERB DET NOUN ADJ ADP NOUN)
= p(NOUN | BOS BOS)
      * p(VERB | BOS NOUN)
```

POS-trigram similarity

Target POS corpus

```
p(NOUN VERB DET NOUN ADJ ADP NOUN)
= p(NOUN | BOS BOS)
       * p(VERB | BOS NOUN)
            * p(DET | NOUN VERB)
```

JOHNS HOPKINS
UNIVERSITY

POS-trigram similarity

Target POS corpus

```
p( NOUN VERB DET NOUN ADJ ADP NOUN )
= p(NOUN | BOS BOS)
      * p(VERB | BOS NOUN)
          * p(DET | NOUN VERB)
              * p(NOUN | VERB DET)
```

POS-trigram similarity

Target POS corpus

```
p( NOUN VERB DET NOUN ADJ ADP NOUN )
= p(NOUN | BOS BOS)
      * p(VERB | BOS NOUN)
           * p(DET | NOUN VERB)
                * p(NOUN | VERB DET)
                     * p(ADJ | DET NOUN)
```

POS-trigram similarity

Target POS corpus

```
p( NOUN VERB DET NOUN ADJ ADP NOUN )
= p(NOUN | BOS BOS)
      * p(VERB | BOS NOUN)
            * p(DET | NOUN VERB)
                  * p(NOUN | VERB DET)
                        * p(ADJ | DET NOUN)
                              . . .
```

Johns Hopkins University

POS-trigram similarity

Target POS corpus

```
p( NOUN VERB DET NOUN ADJ ADP NOUN )
= p(NOUN | BOS BOS)
      * p(VERB | BOS NOUN)
            * p(DET | NOUN VERB)
                  * p(NOUN | VERB DET)
                        * p(ADJ | DET NOUN)
                              . . .

p(NOUN | VERB DET)=
```

POS-trigram similarity

Target POS corpus

p(NOUN VERB DET NOUN ADJ ADP NOUN)
= p(NOUN | BOS BOS)
* p(VERB | BOS NOUN)
* p(DET | NOUN VERB)
* p(NOUN | VERB DET)
* p(ADJ | DET NOUN)
. . .

$$p(\text{NOUN} \mid \text{VERB DET}) = \frac{\#\text{VERB DET NOUN}}{\#\text{VERB DET}}$$

JOHNS HOPKINS UNIVERSITY

POS-trigram similarity

Target POS corpus

p(NOUN VERB DET NOUN ADJ ADP NOUN)

= p(NOUN | BOS BOS)

* p(VERB | BOS NOUN)

* p(DET | NOUN VERB)

* p(NOUN | VERB DET)

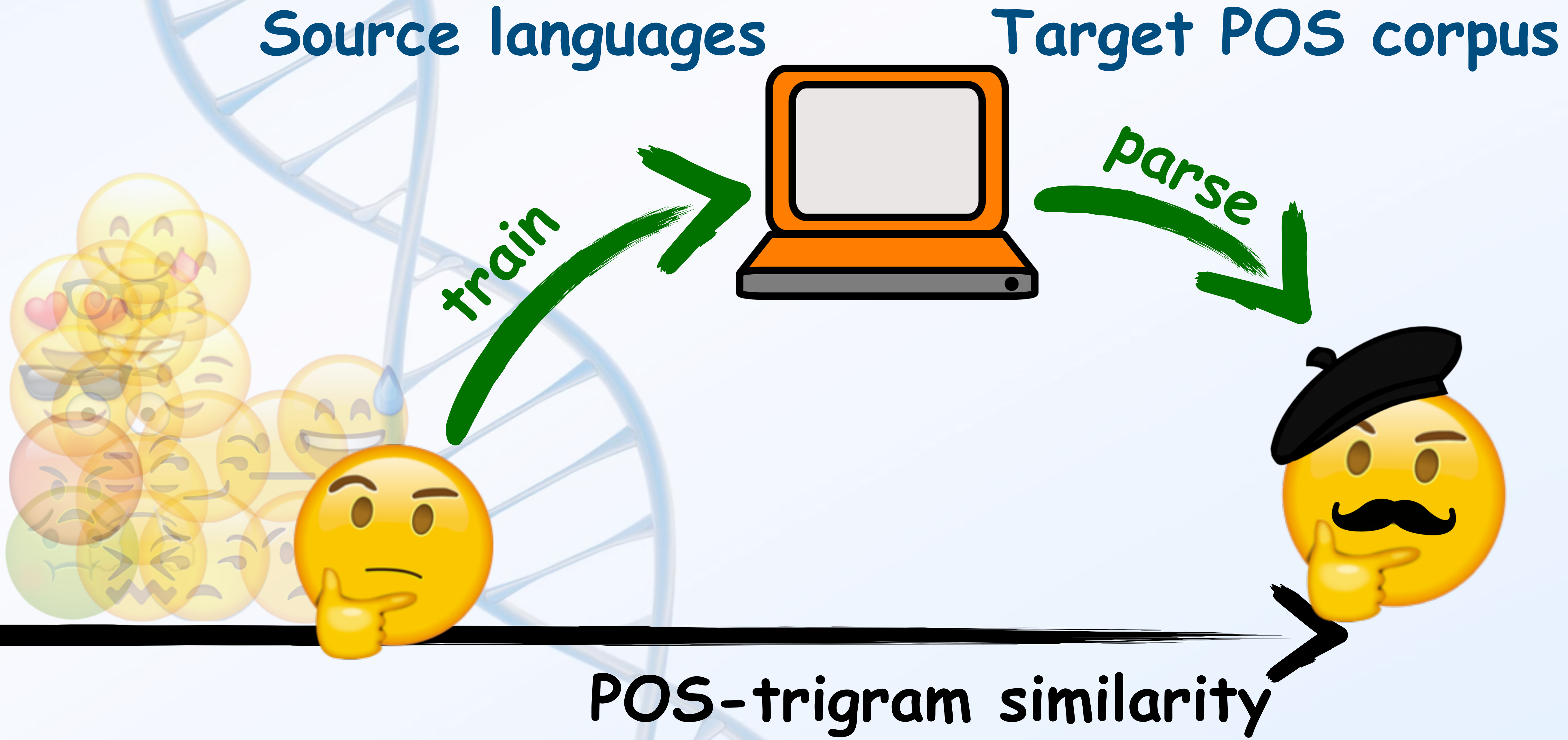
* p(ADJ | DET NOUN)

. . .

$$p(\text{NOUN} \mid \text{VERB DET}) = \frac{\#\text{VERB DET NOUN}}{\#\text{VERB DET}}$$

Count from Source language!

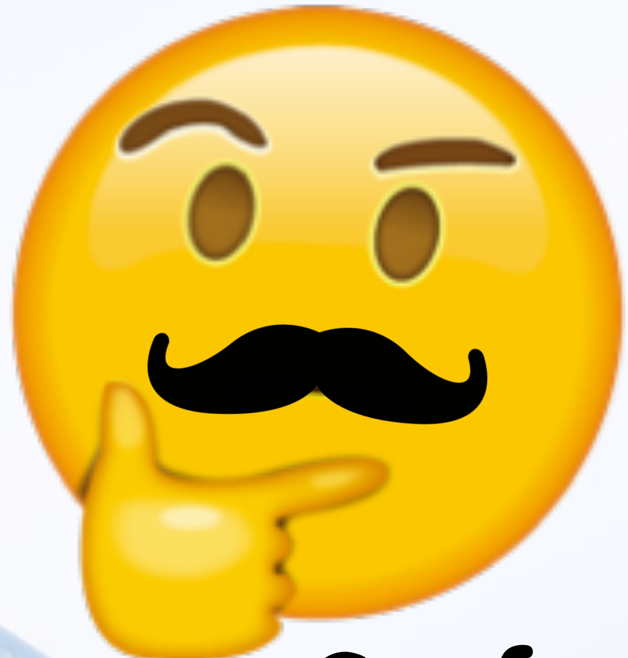
Single-Source Selection



Synthetic Data

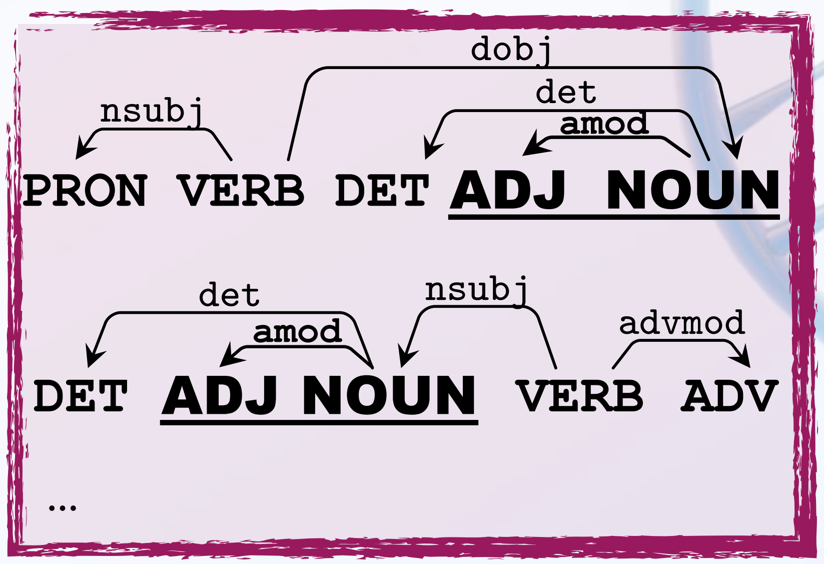
Source

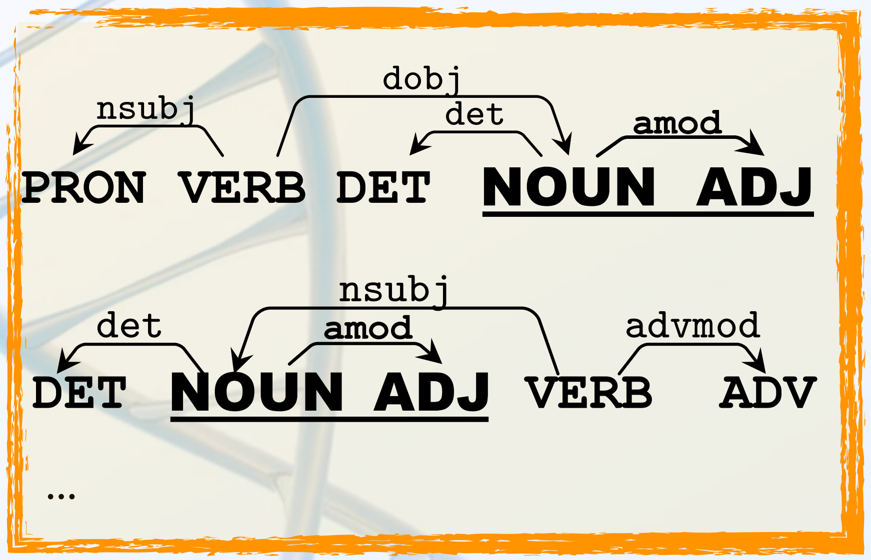
Source'



Target POS corpus

Surface similarity





NOUN VERB DET **NOUN ADJ**

NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

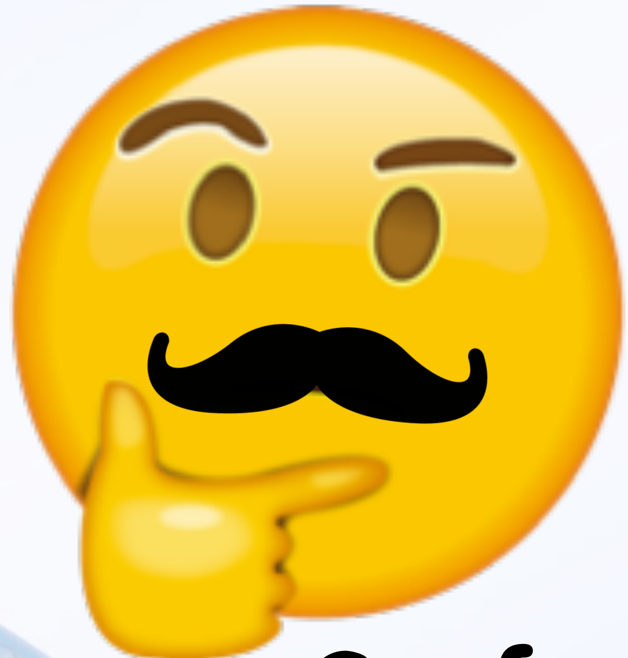
PRON VERB ADP DET NOUN

…

Synthetic Data

Source

Synthetic Source'



Target
POS corpus

Surface similarity

dobj, nsubj, det, amod
PRON VERB DET **ADJ NOUN**

det, amod, nsubj, advmod
DET **ADJ NOUN** VERB ADV

...

dobj, nsubj, det, amod
PRON VERB DET **NOUN ADJ**

nsubj, det, amod, advmod
DET **NOUN ADJ** VERB ADV

...

NOUN VERB DET **NOUN ADJ**

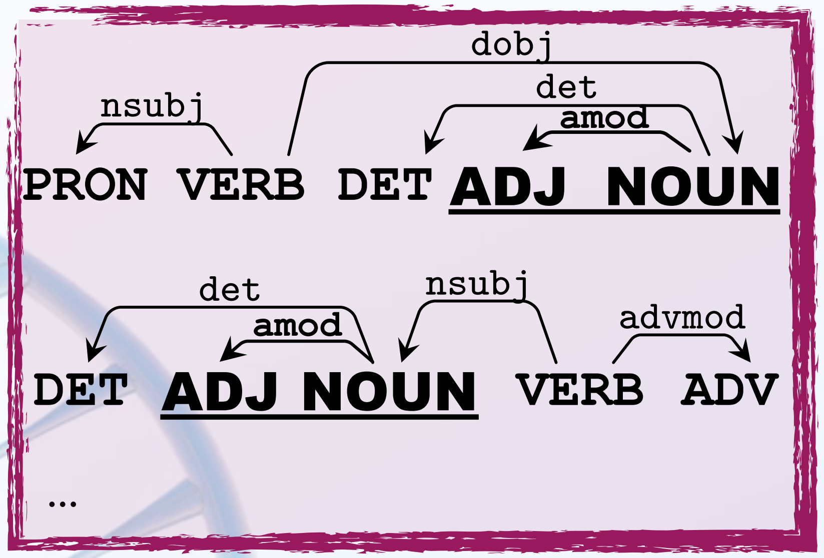
NOUN VERB PART NOUN

DET **NOUN ADJ** VERB

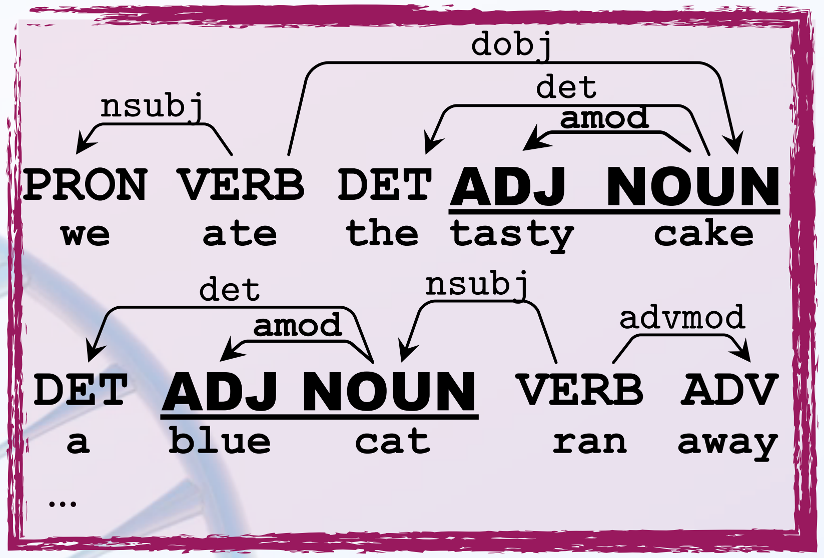
PRON VERB ADP DET NOUN

...

Synthetic Data

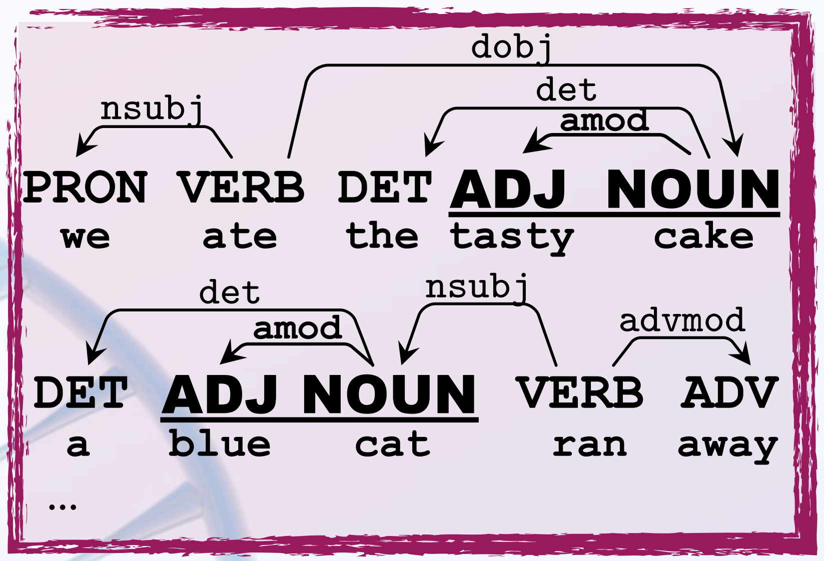


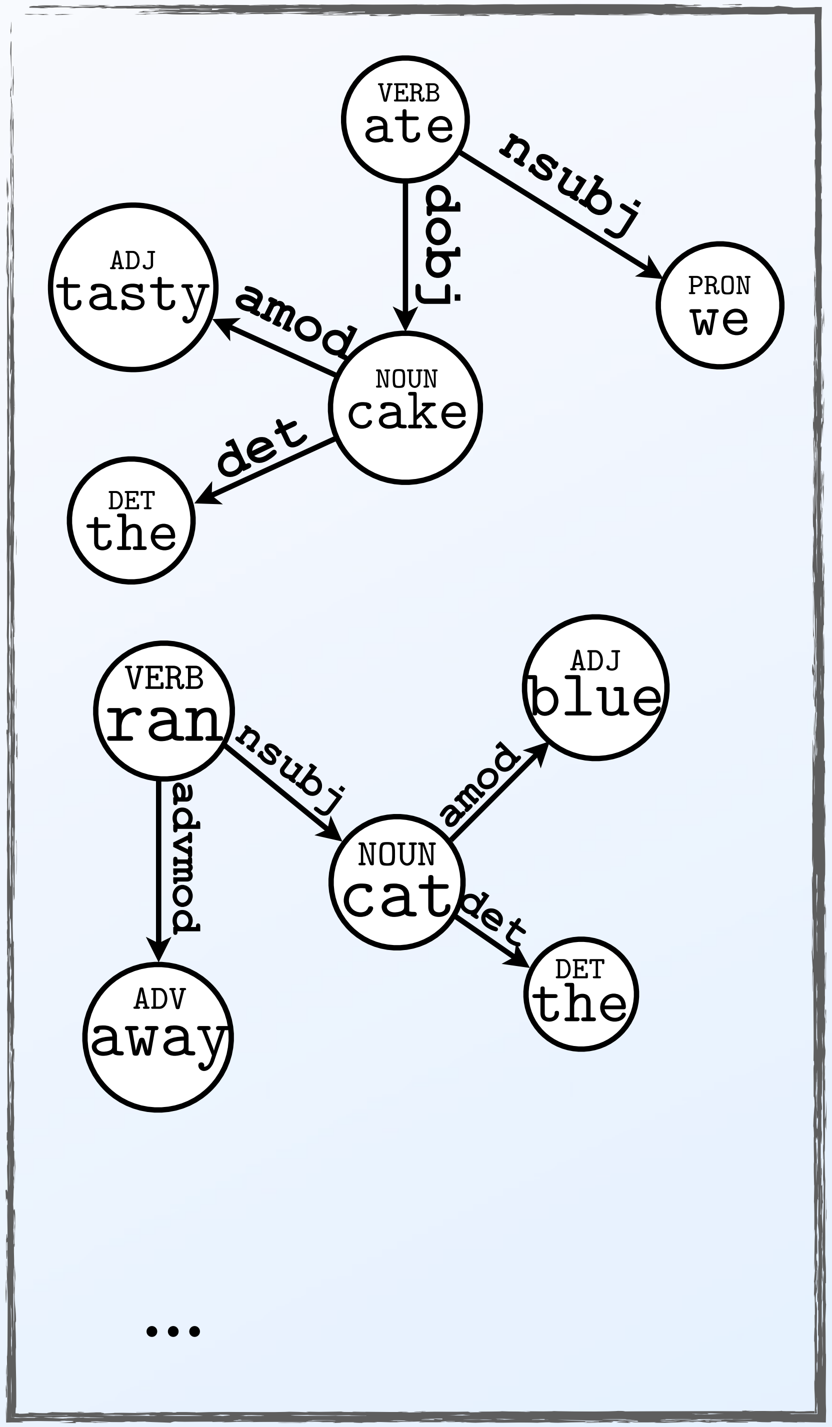
Synthetic Data



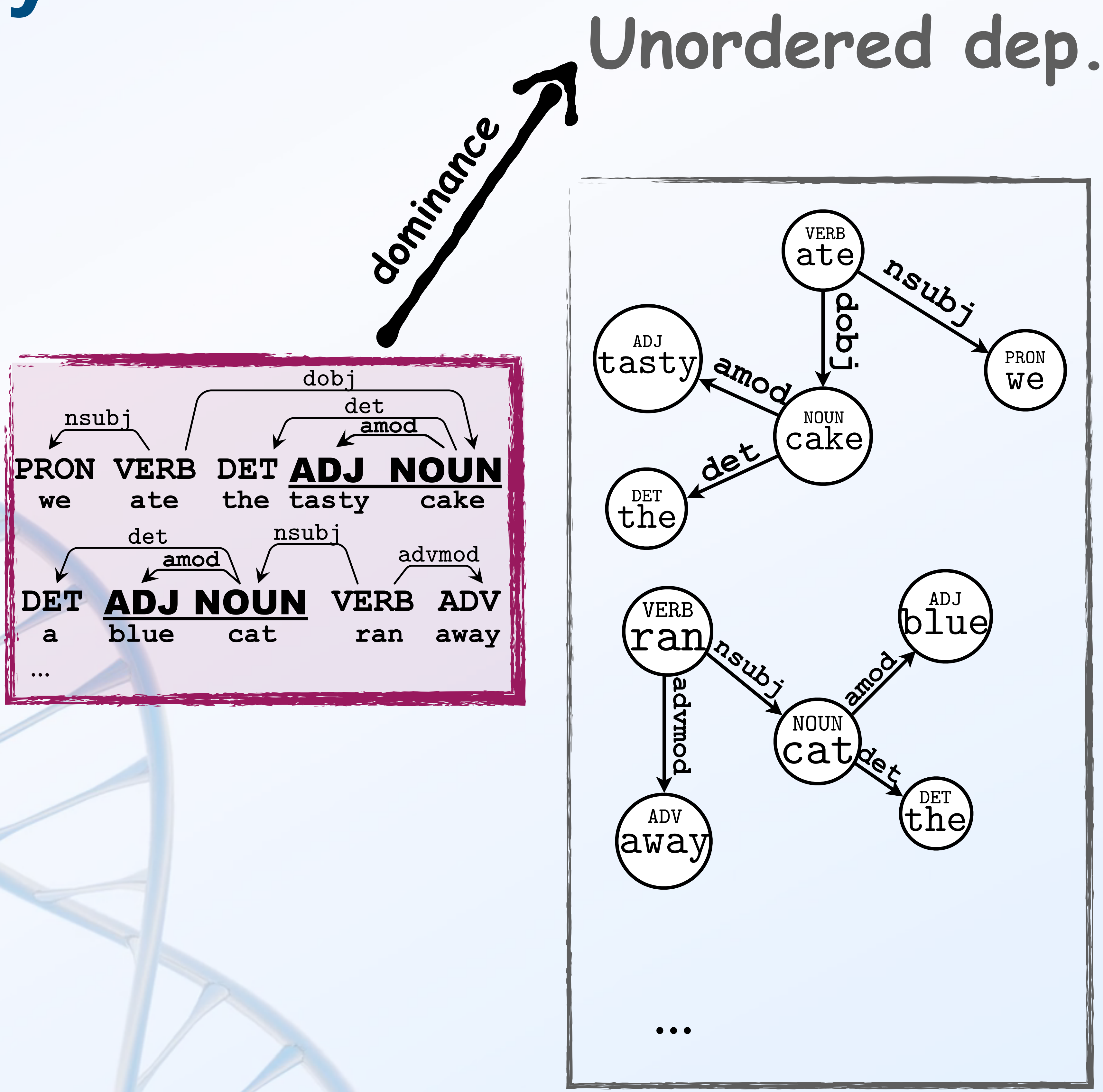
Synthetic Data

Unordered dep.





Synthetic Data



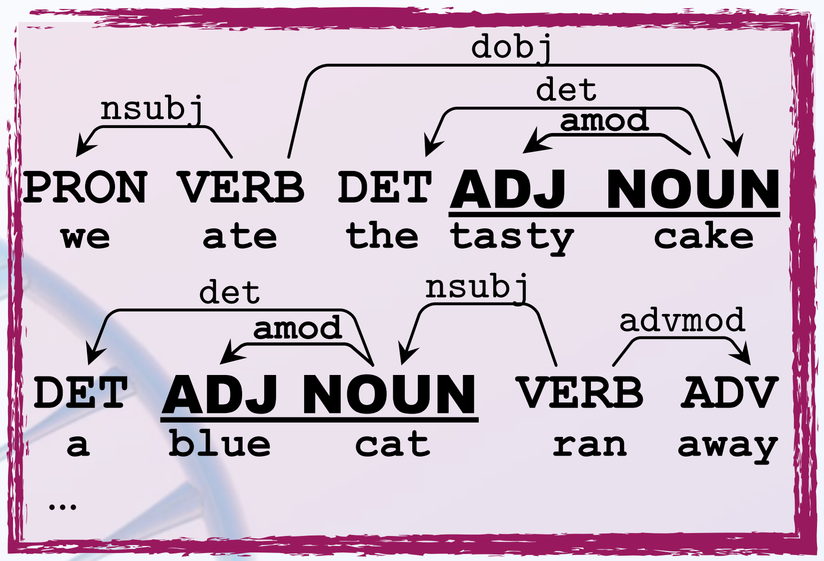
Synthetic Data

Surface order

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

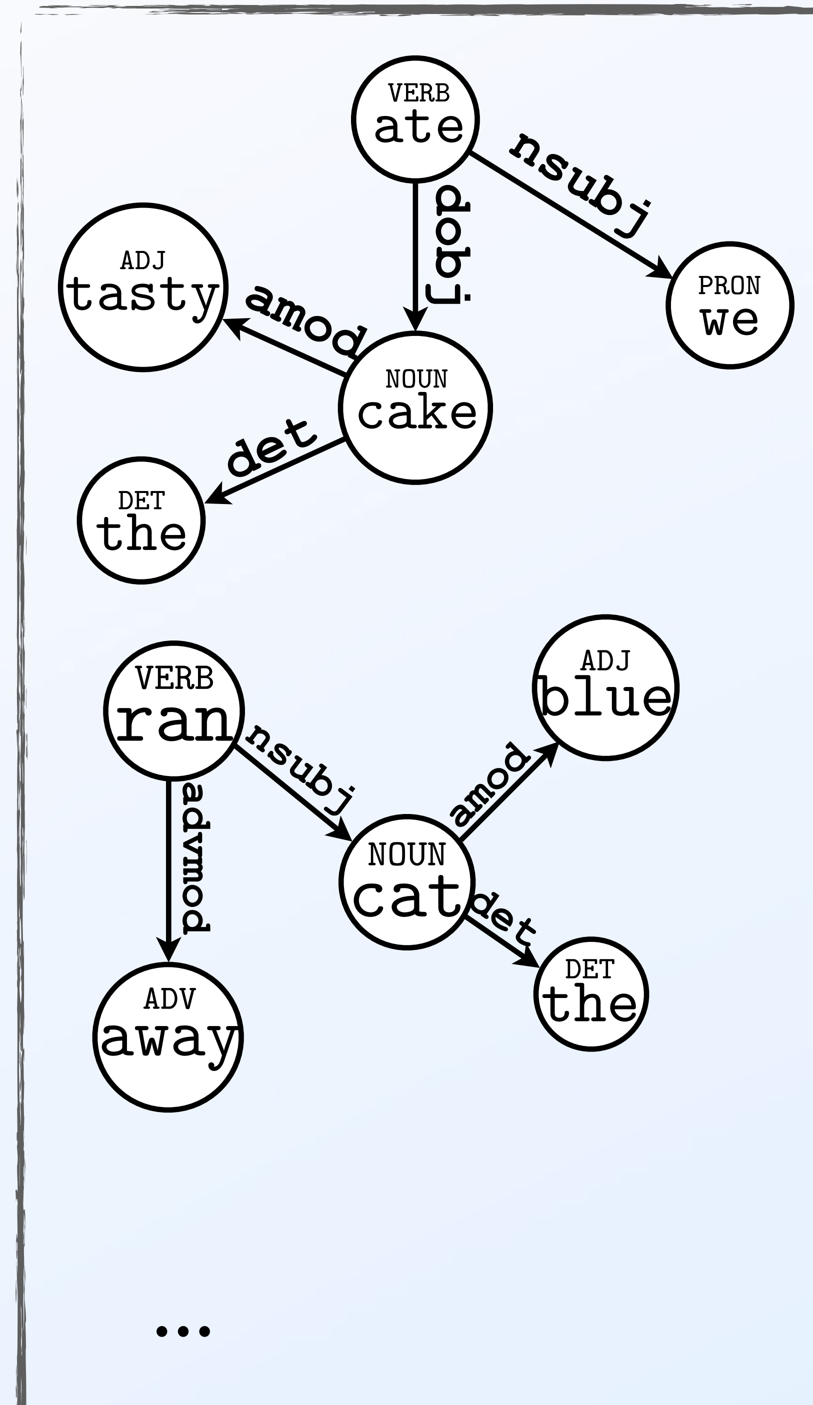
DET **ADJ NOUN** VERB ADV
a blue cat ran away

...

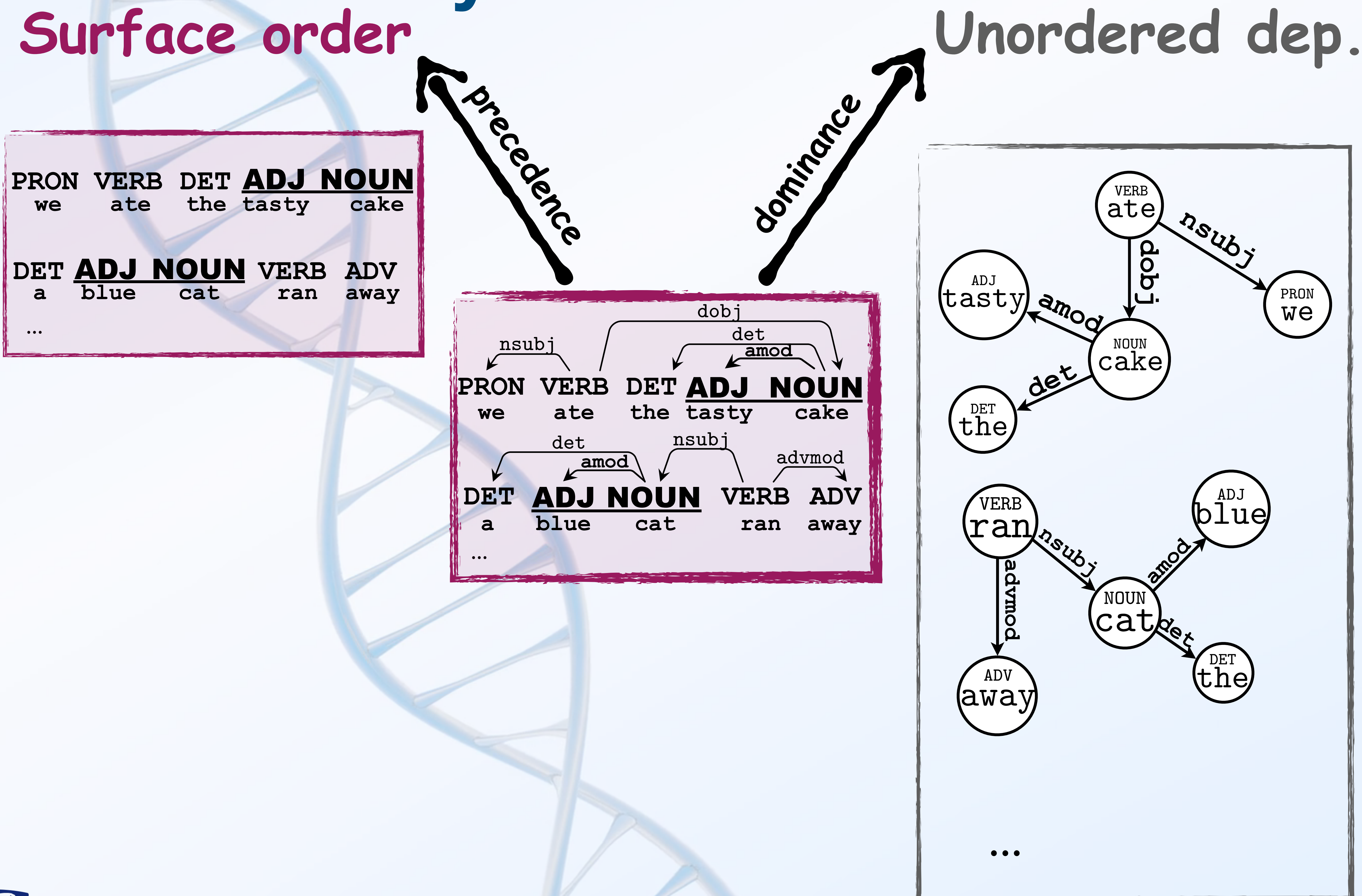


dominance

Unordered dep.



Synthetic Data



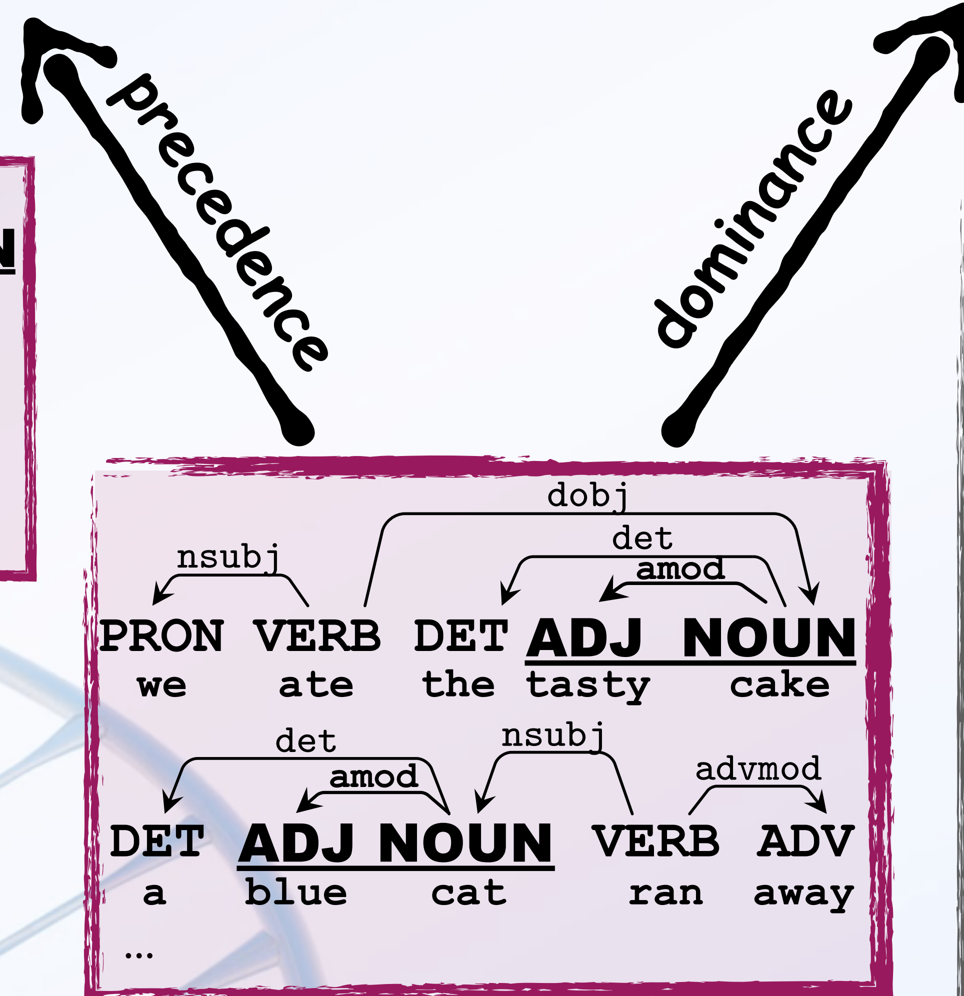
Synthetic Data

Surface order

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

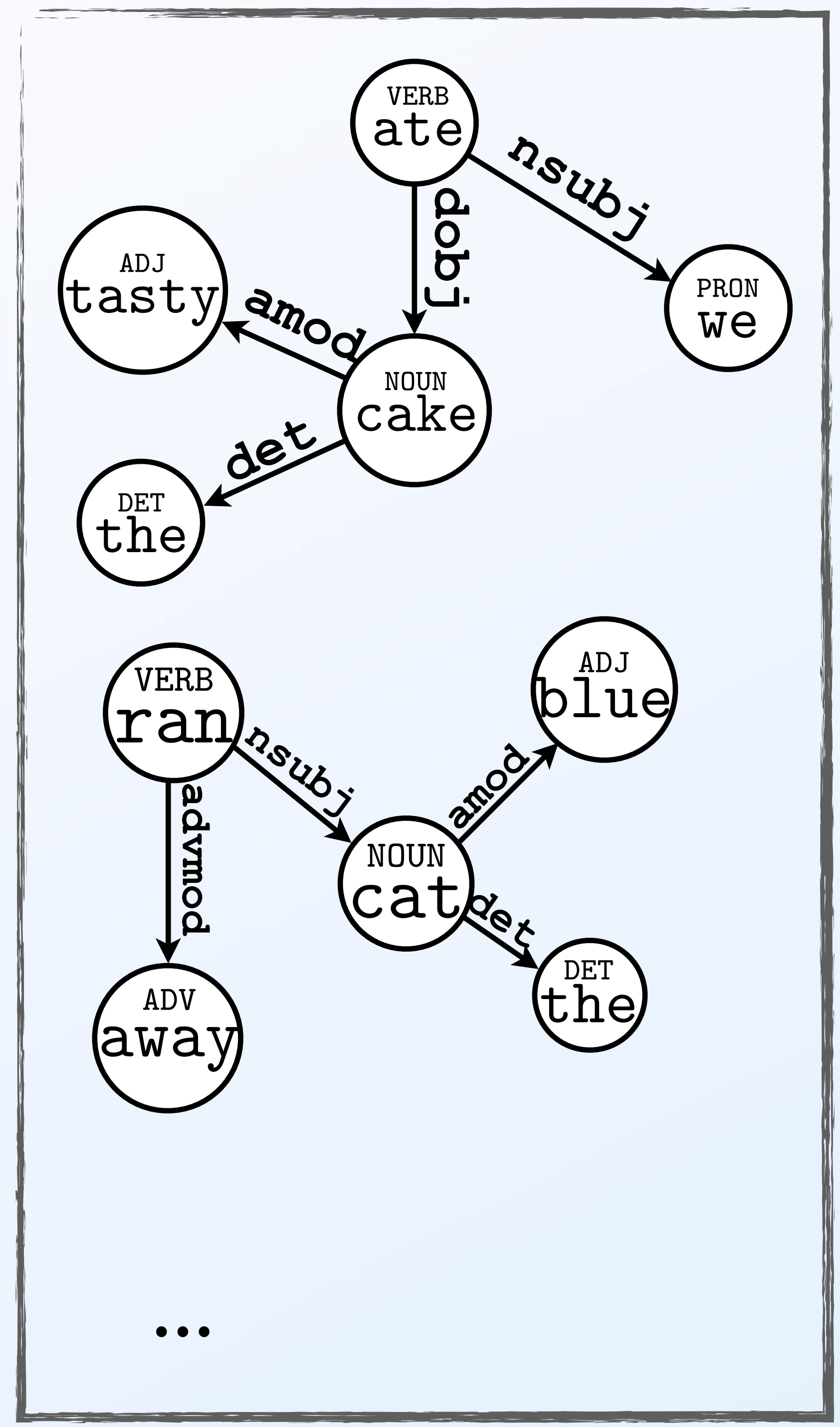
DET **ADJ NOUN** VERB ADV
a blue cat ran away

…



Unordered dep.

(Lang. universal)



Synthetic Data

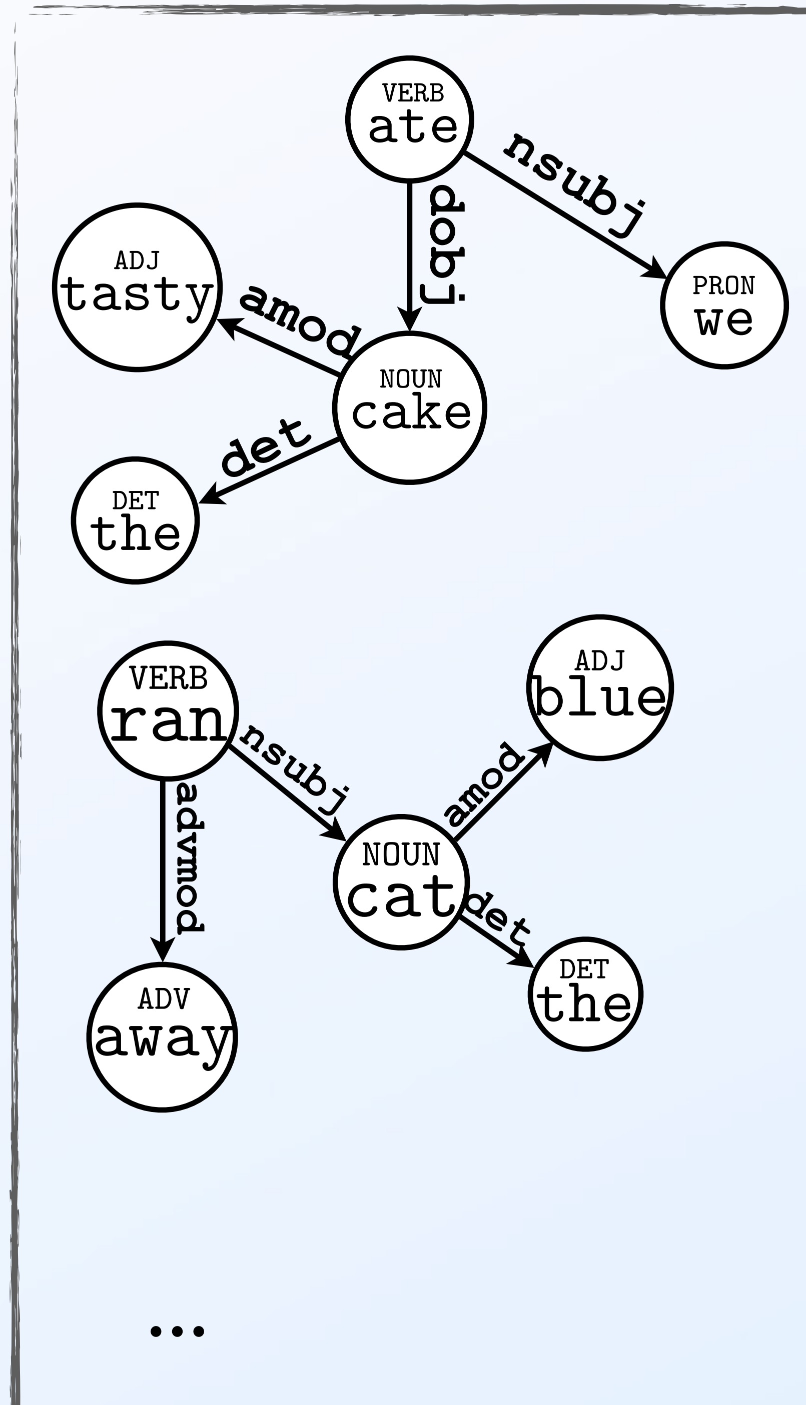
Surface order
(Lang. specific)

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

DET **ADJ NOUN** VERB ADV
a blue cat ran away

…

Unordered dep.
(Lang. universal)



Synthetic Data

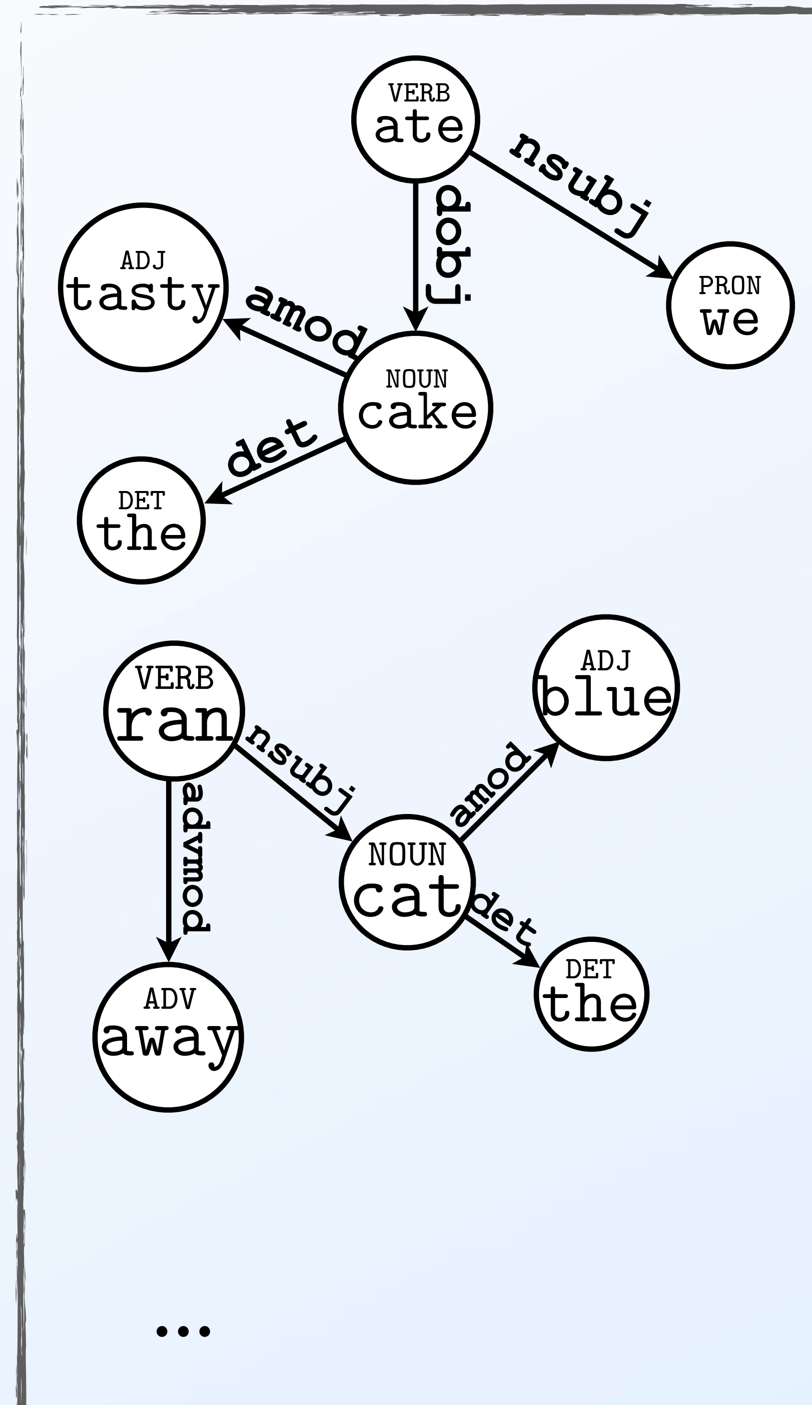
Surface order
(Lang. specific)

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

DET **ADJ NOUN** VERB ADV
a blue cat ran away

…

Unordered dep.
(Lang. universal)



Synthetic Data

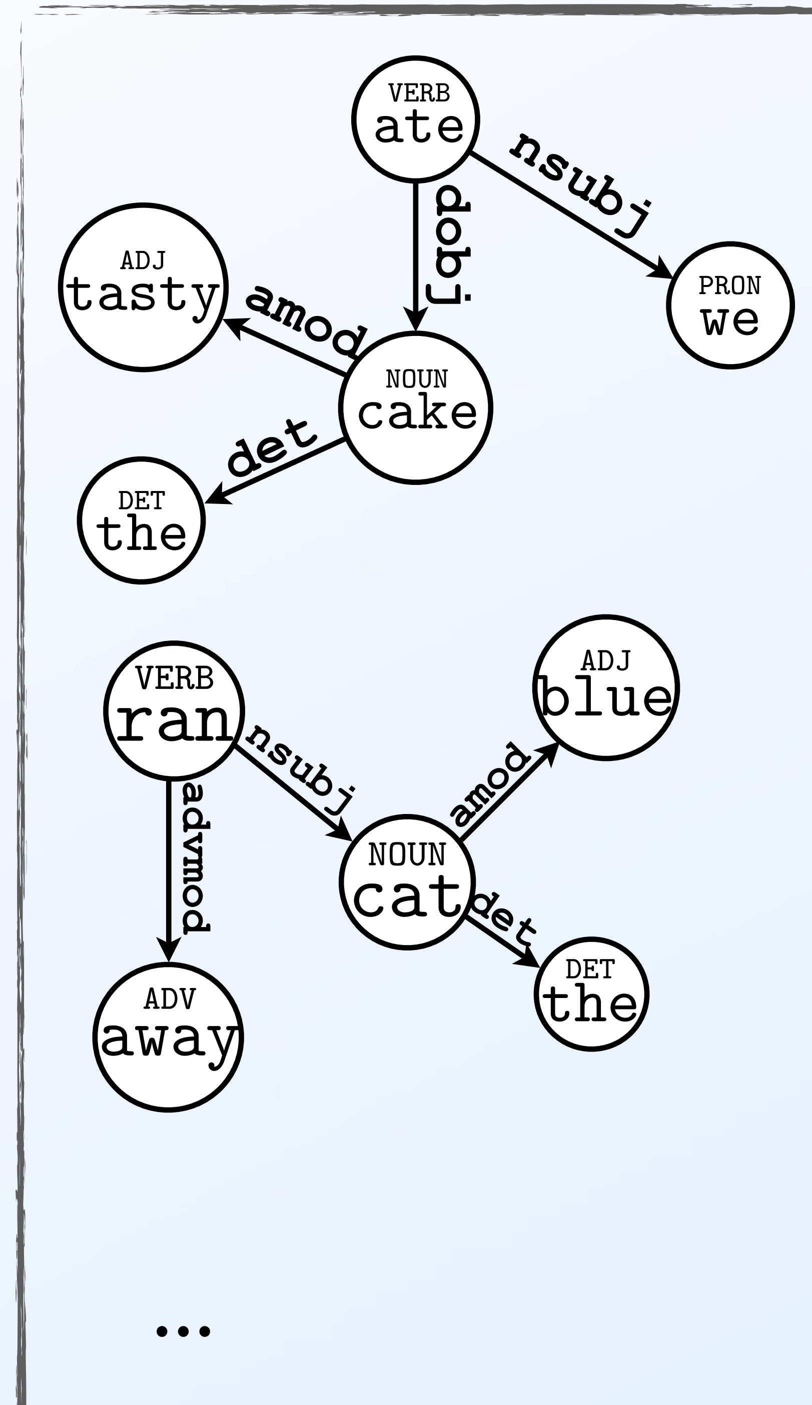
Surface order
(Lang. specific)

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

DET **ADJ NOUN** VERB ADV
a blue cat ran away

…

Unordered dep.
(Lang. universal)



Synthetic Data

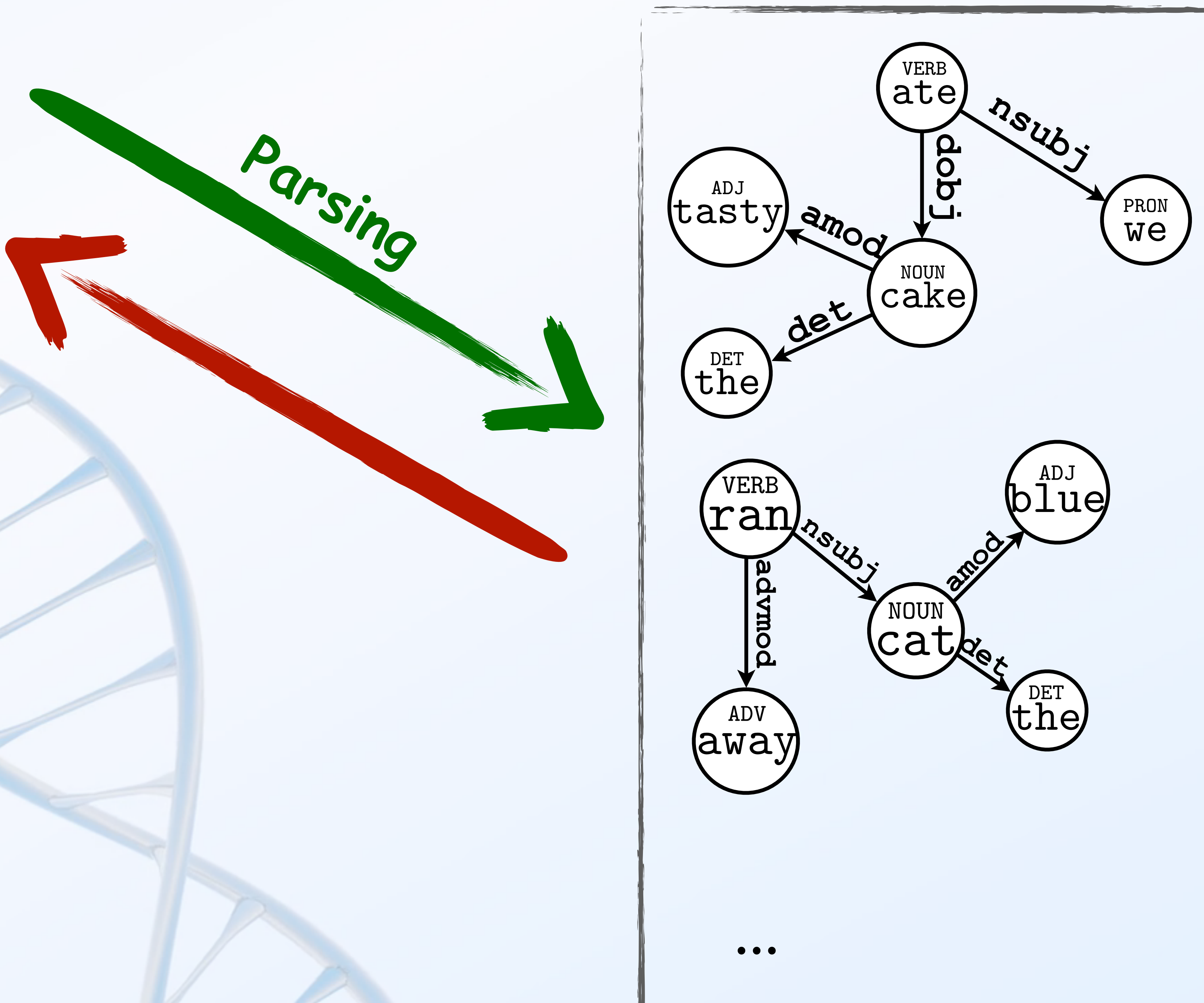
Surface order
(Lang. specific)

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

DET **ADJ NOUN** VERB ADV
a blue cat ran away

…

Unordered dep.
(Lang. universal)

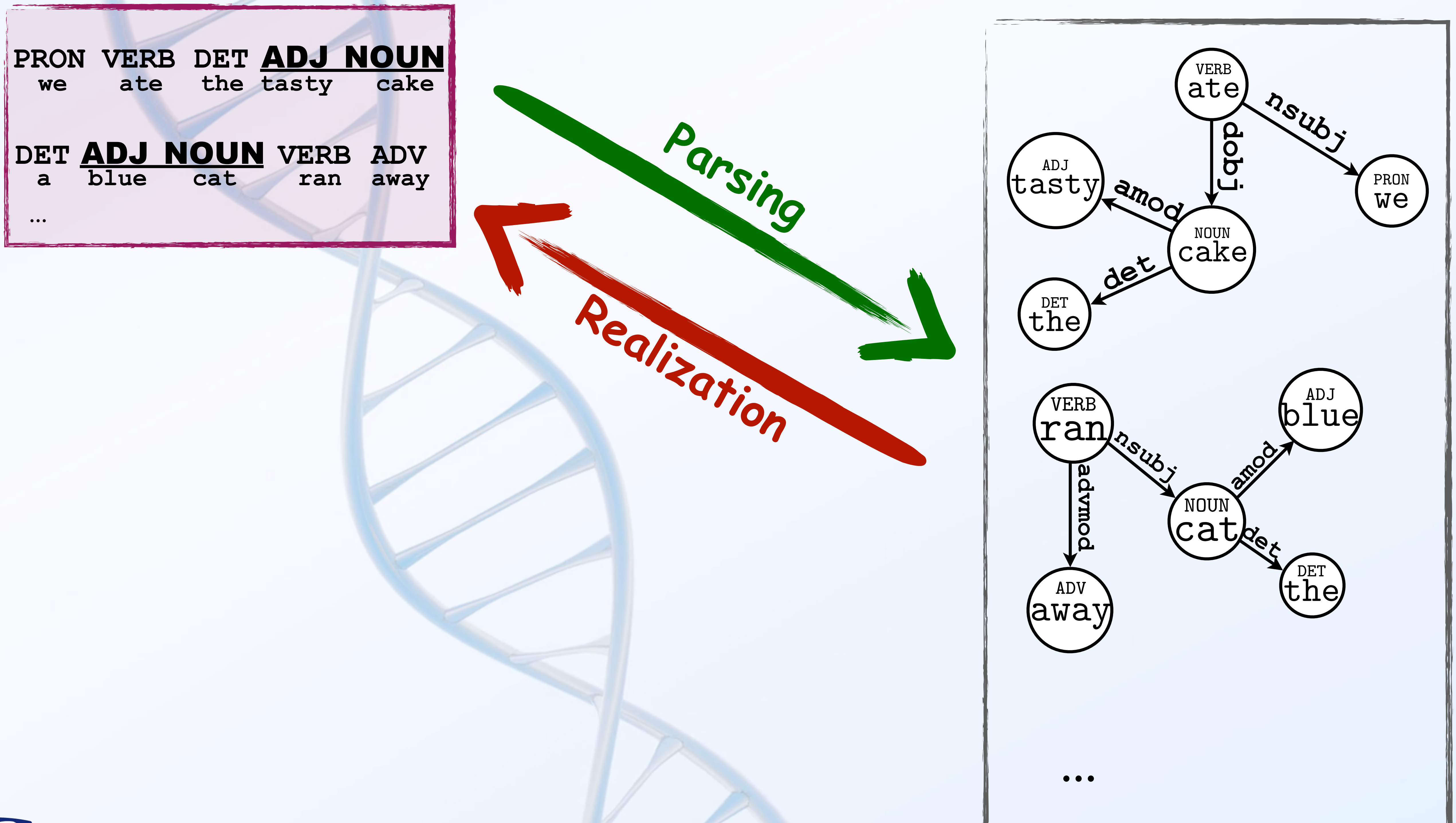


JOHNS HOPKINS UNIVERSITY

Synthetic Data

Surface order
(Lang. specific)

Unordered dep.
(Lang. universal)



Synthetic Data

Surface order
(Lang. specific)

PRON	VERB	DET	ADJ NOUN	
we	ate	the	tasty	cake

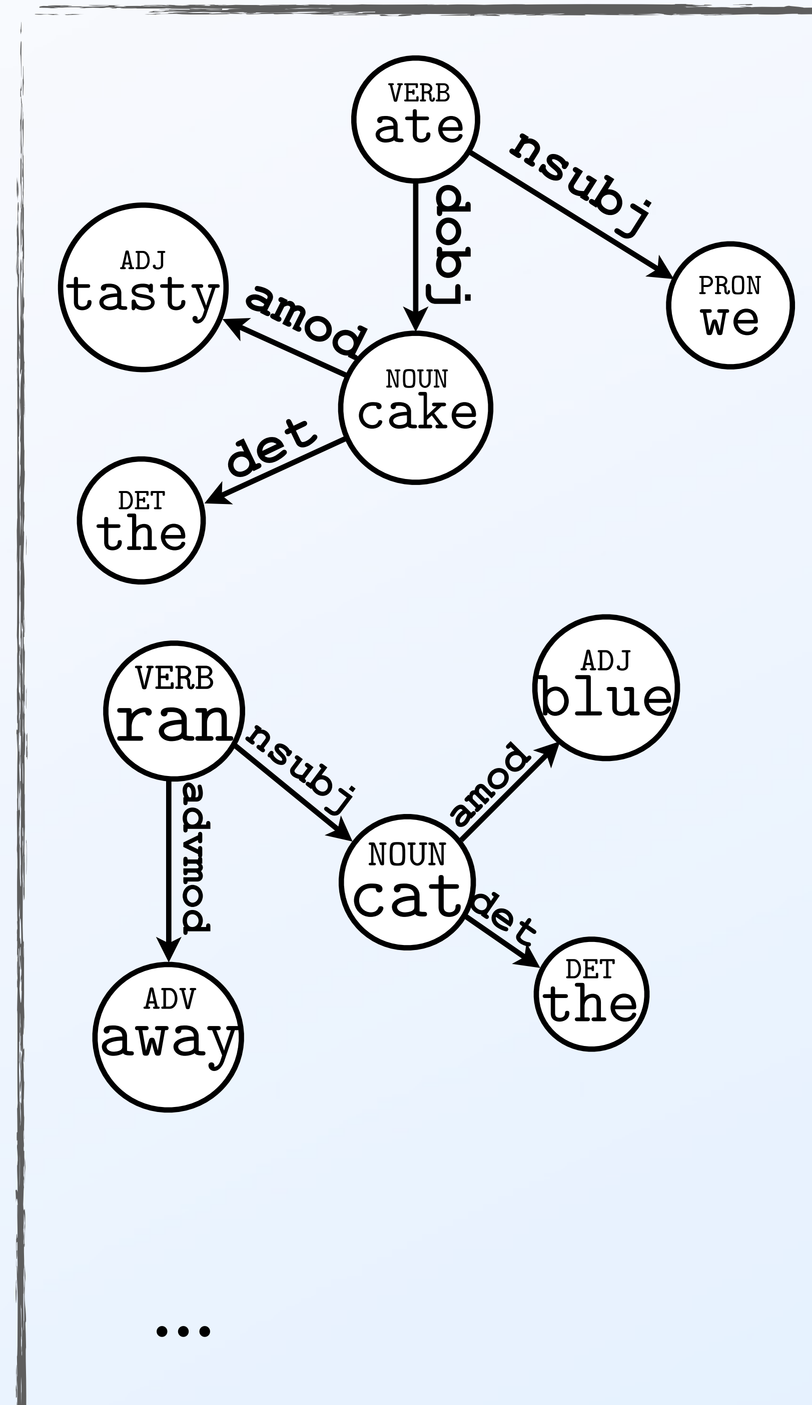
DET	ADJ NOUN		VERB	ADV
a	blue	cat	ran	away

…

Parsing

(English) Realization

Unordered dep.
(Lang. universal)



Synthetic Data

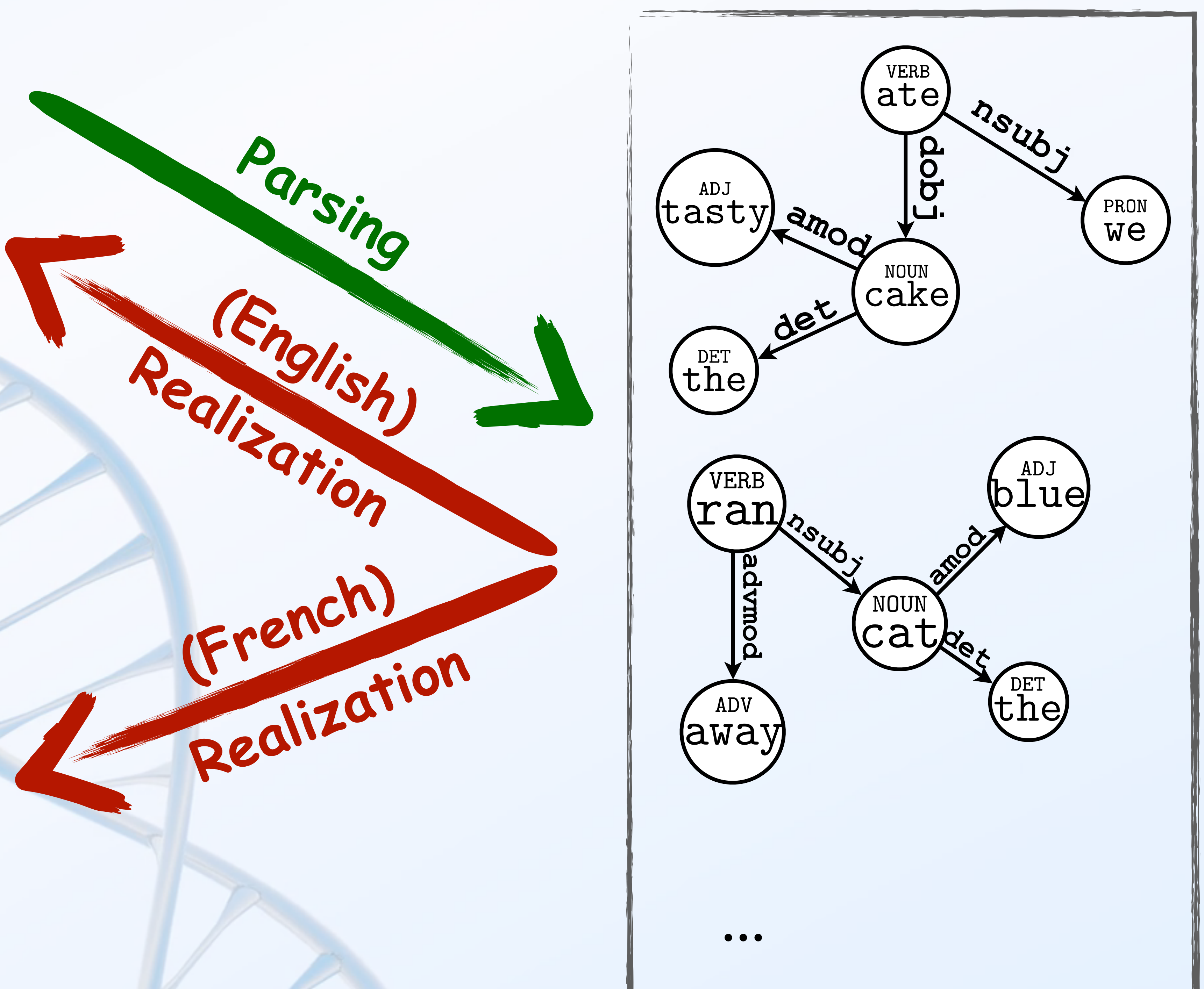
Surface order
(Lang. specific)

PRON VERB DET **ADJ NOUN**
we ate the tasty cake

DET **ADJ NOUN** VERB ADV
a blue cat ran away

…

Unordered dep.
(Lang. universal)

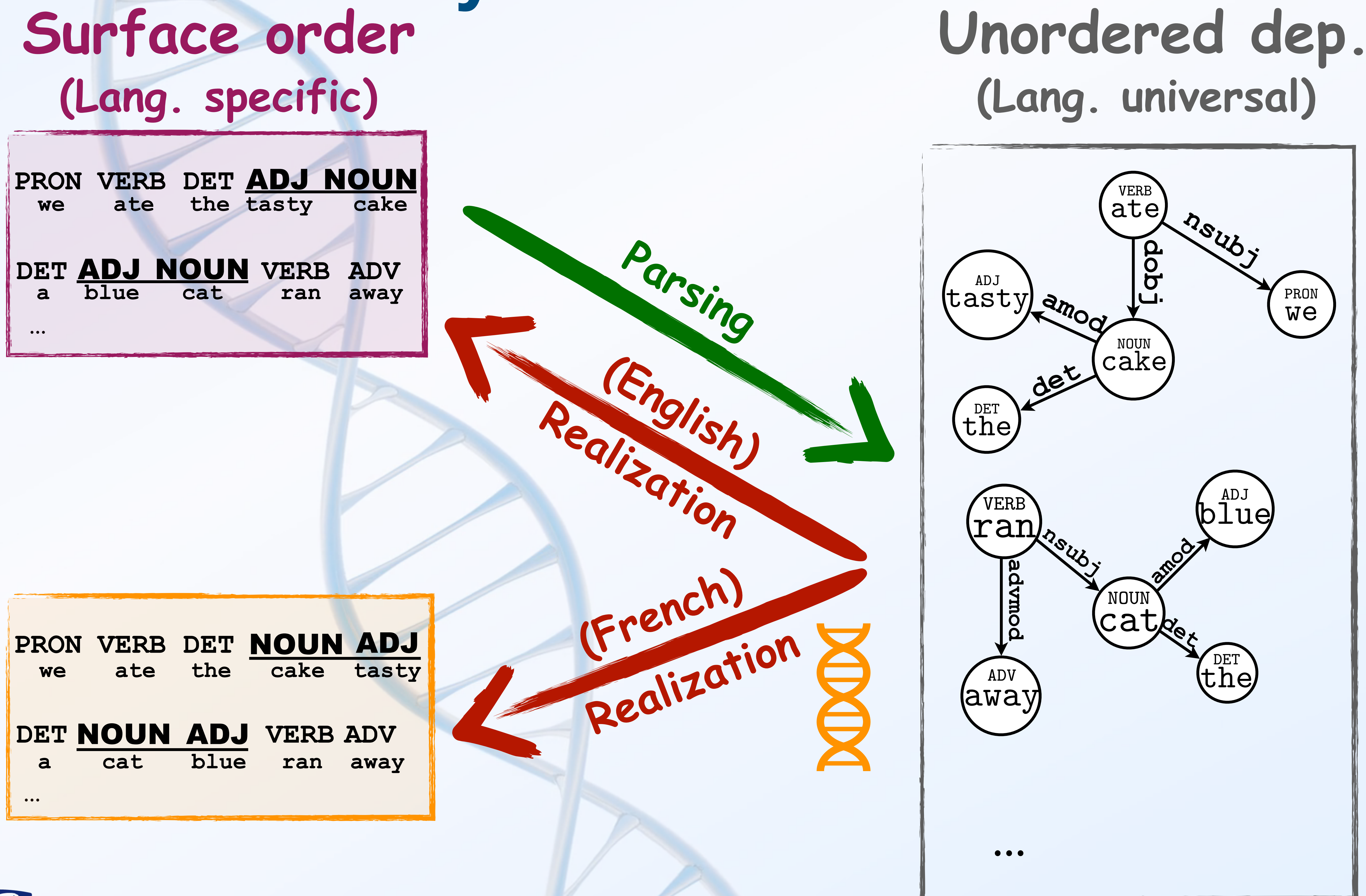


PRON VERB DET **NOUN ADJ**
we ate the cake tasty

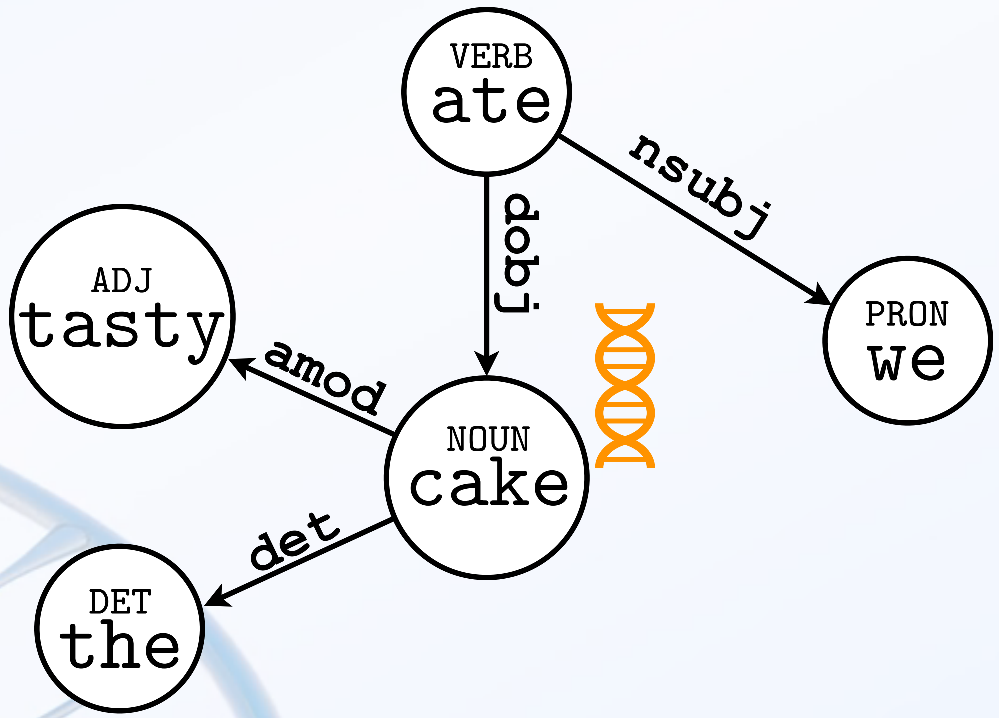
DET **NOUN ADJ** VERB ADV
a cat blue ran away

…

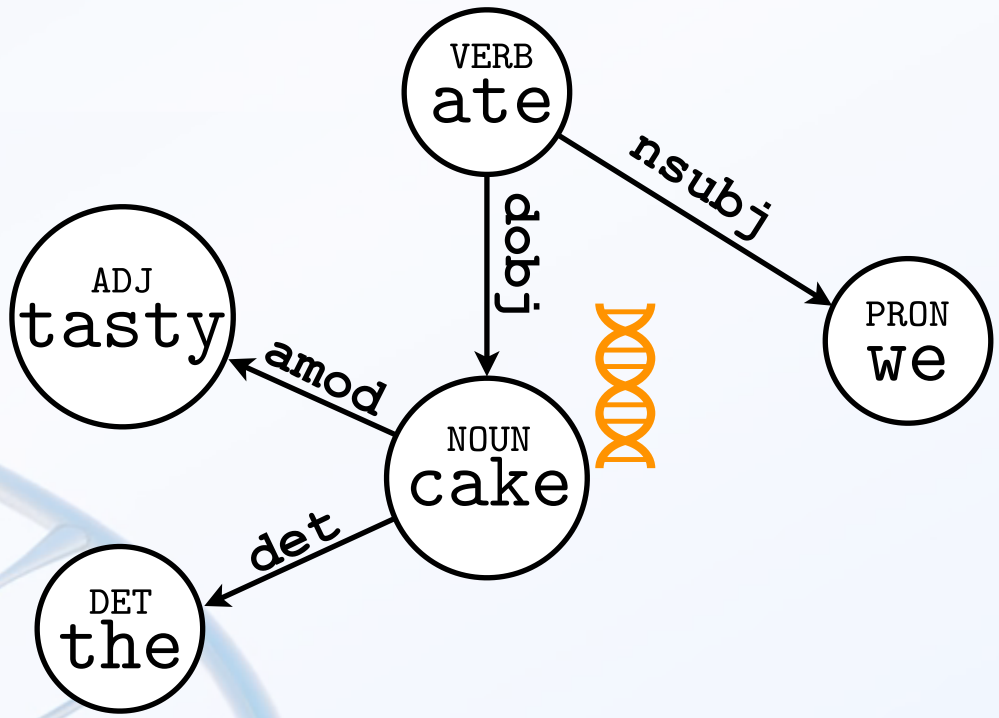
Synthetic Data



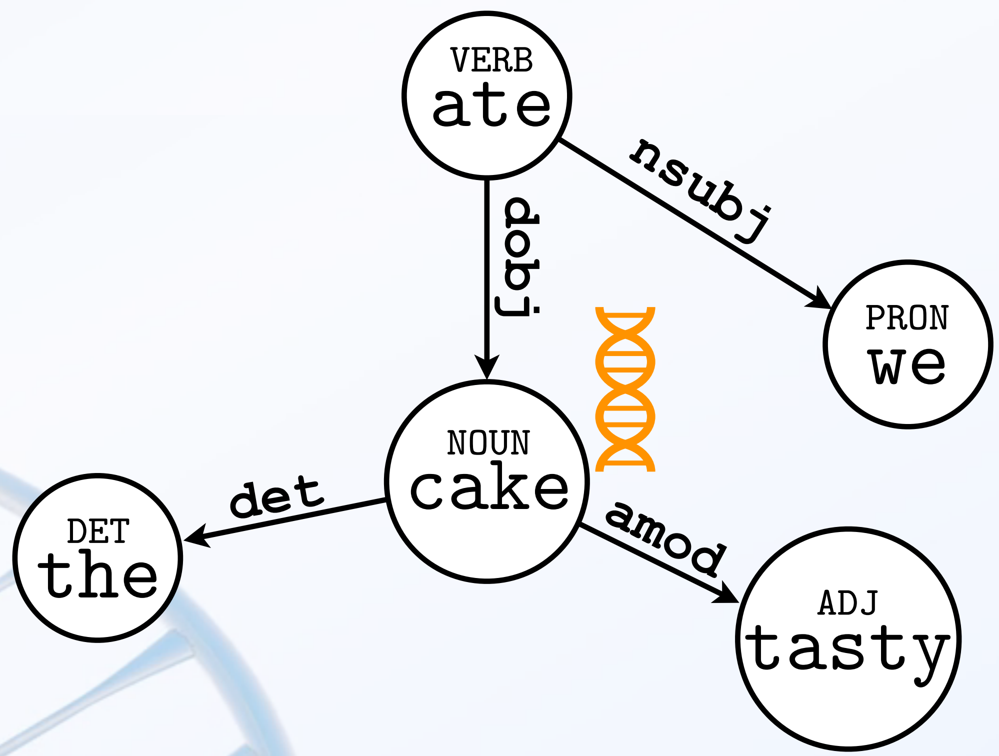
Surface Realization



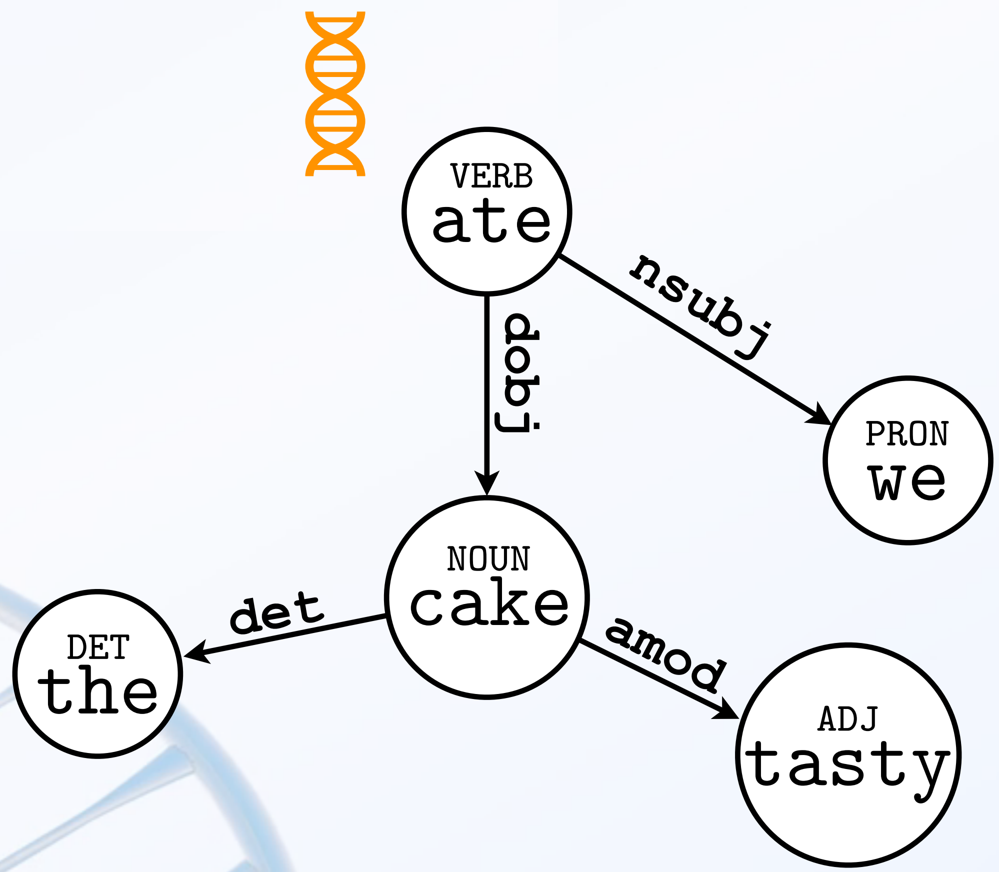
Surface Realization



Surface Realization

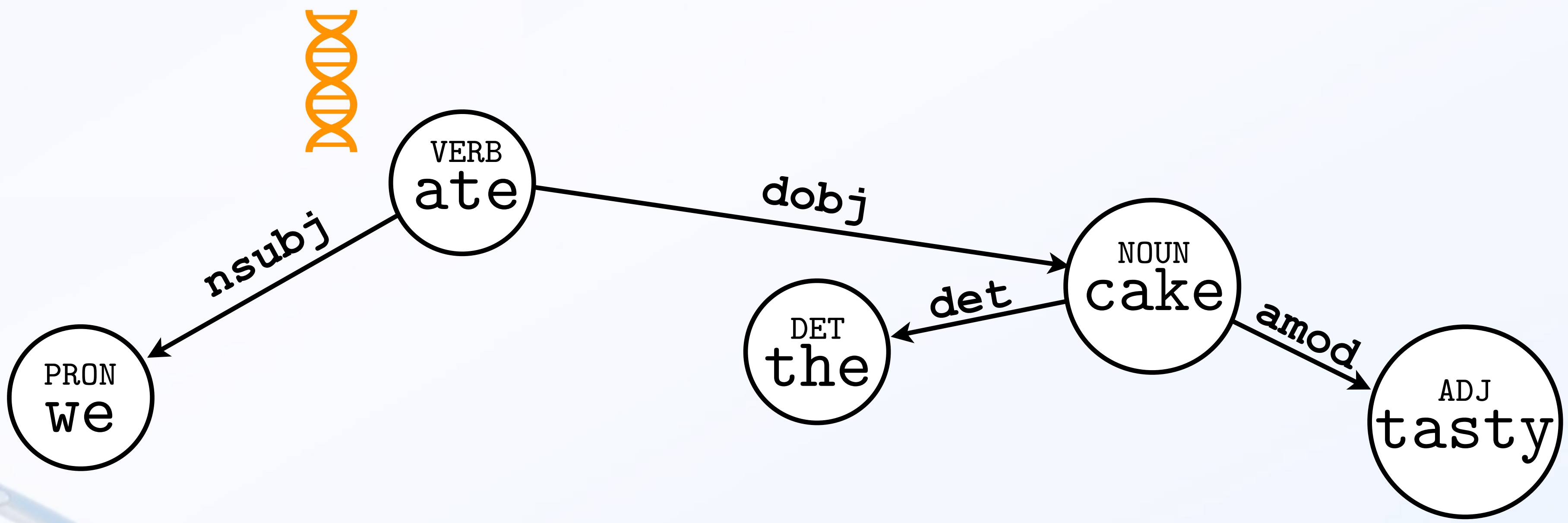


Surface Realization

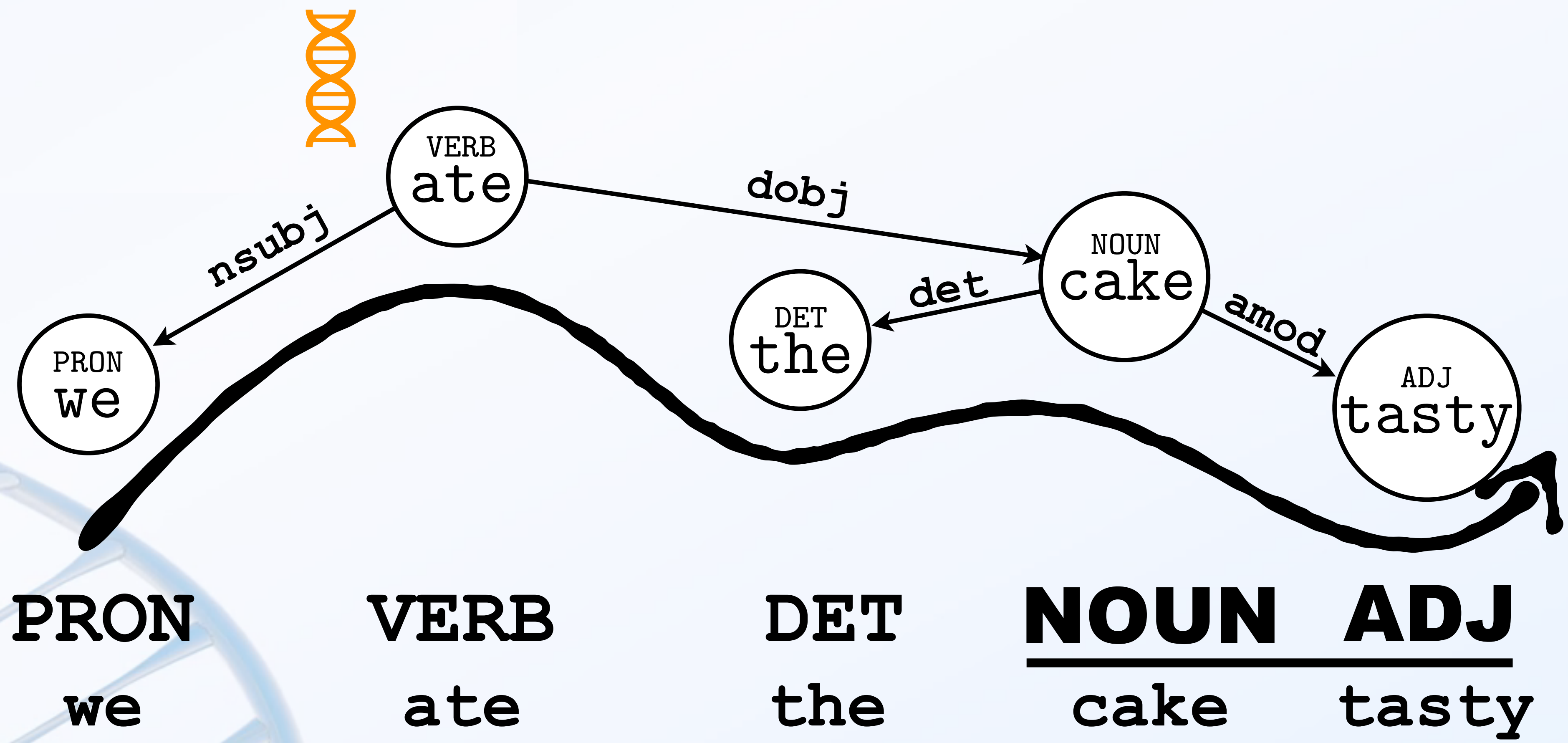


Johns Hopkins University

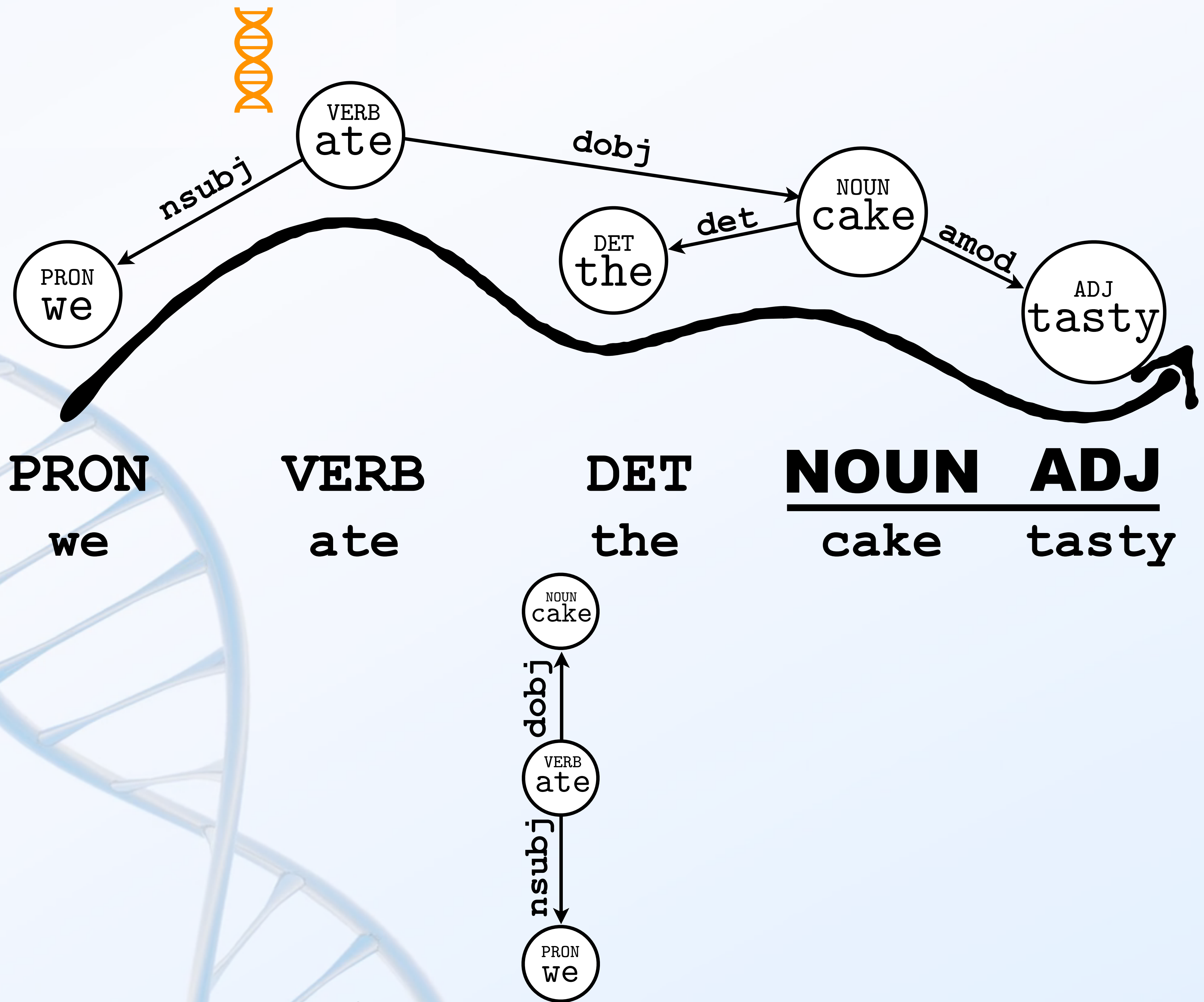
Surface Realization



Surface Realization

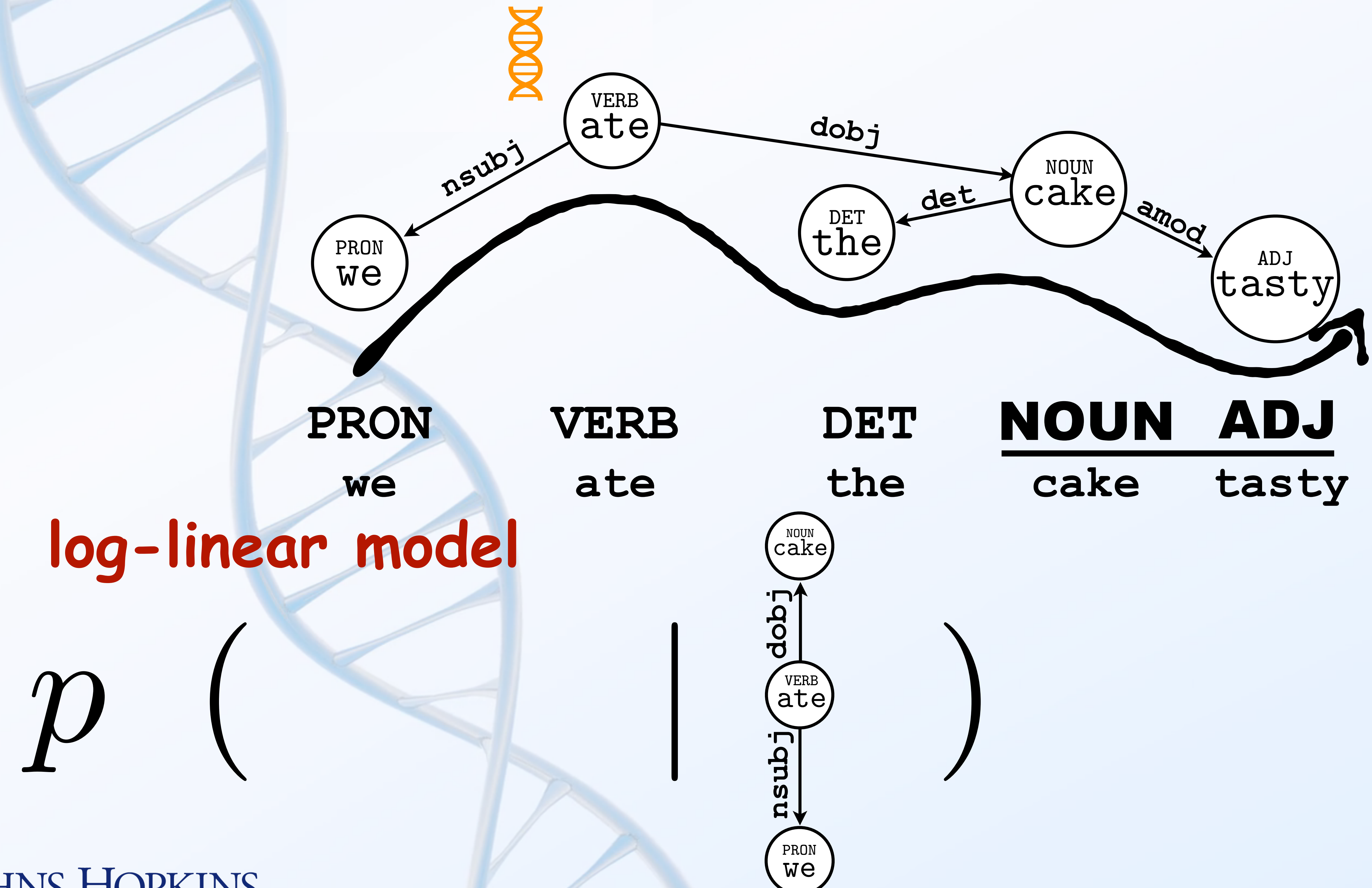


Surface Realization

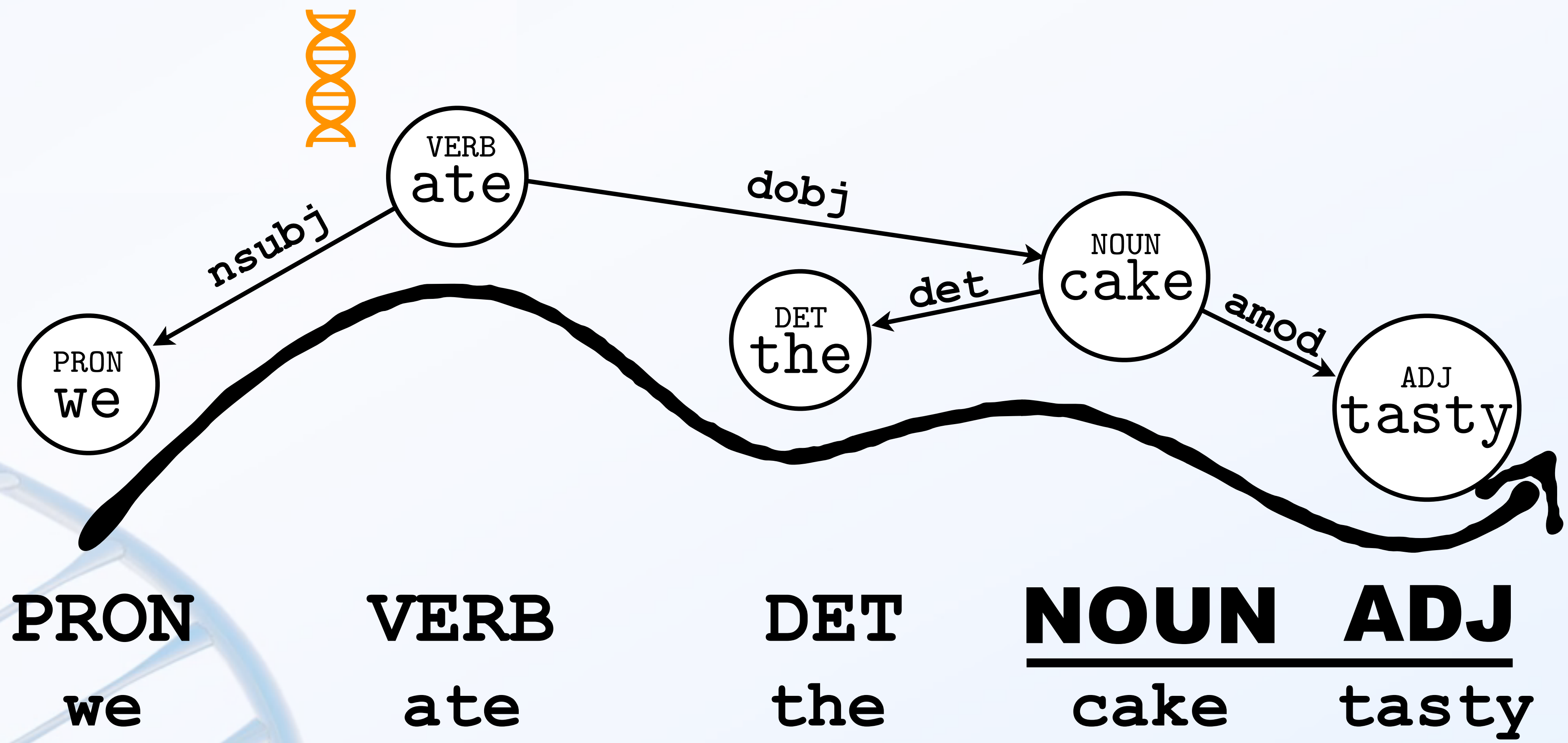


JOHNS HOPKINS
UNIVERSITY

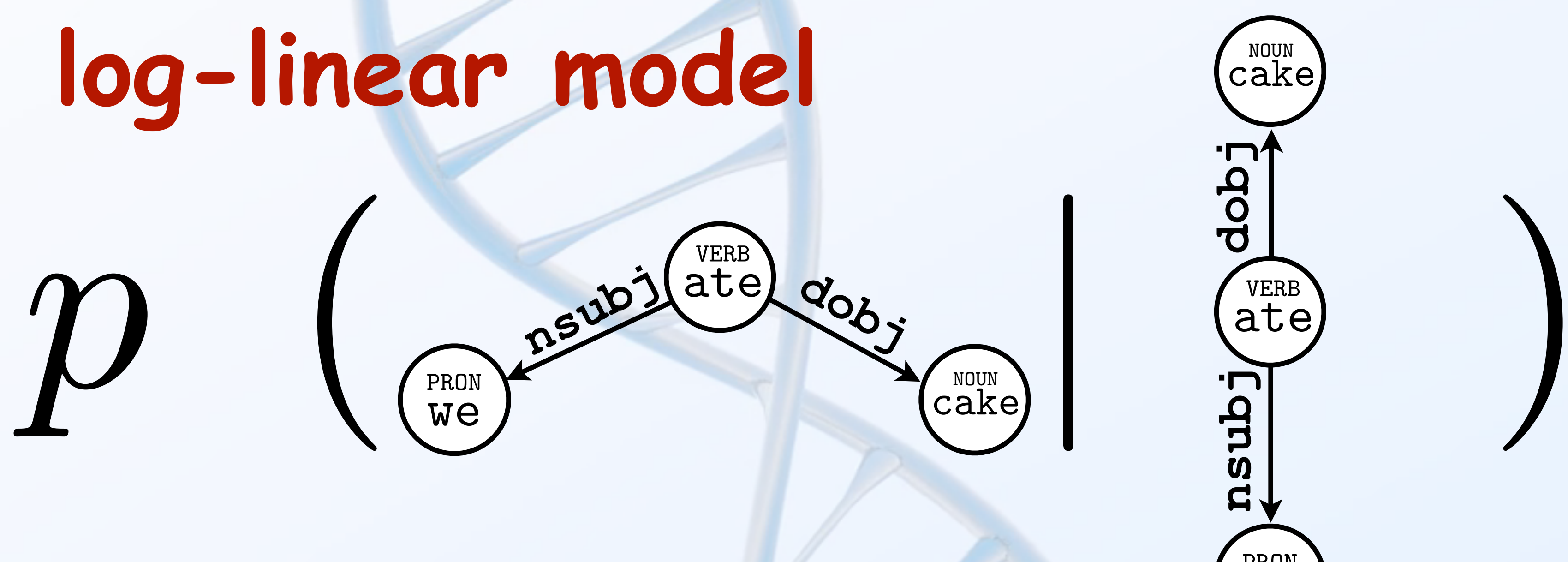
Surface Realization



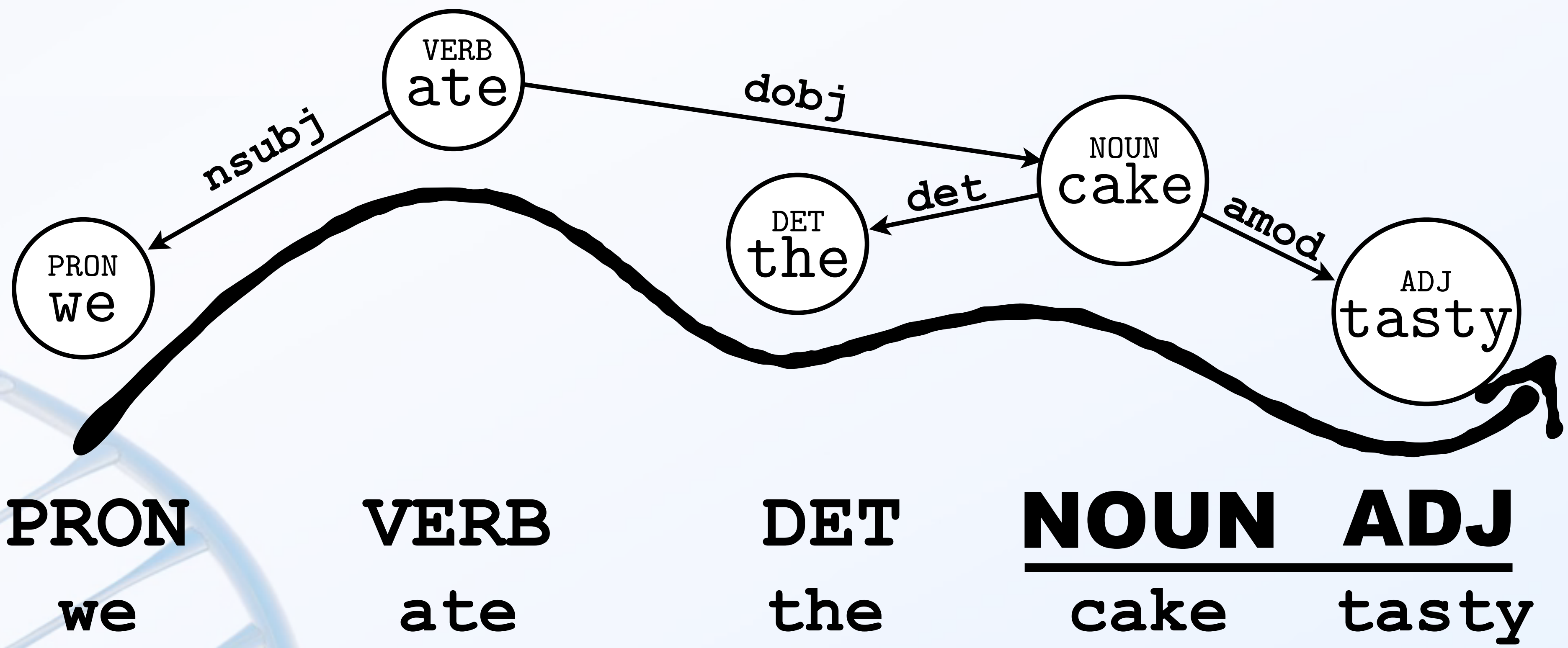
Surface Realization



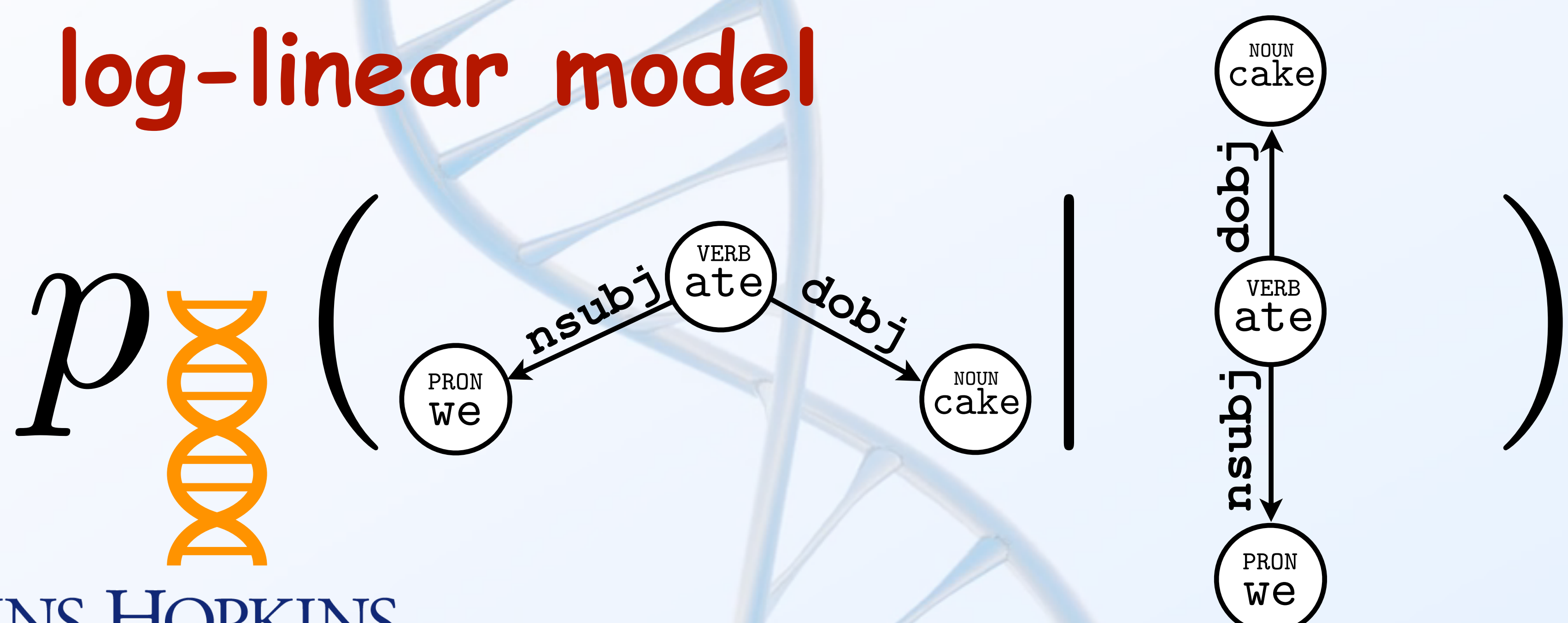
log-linear model



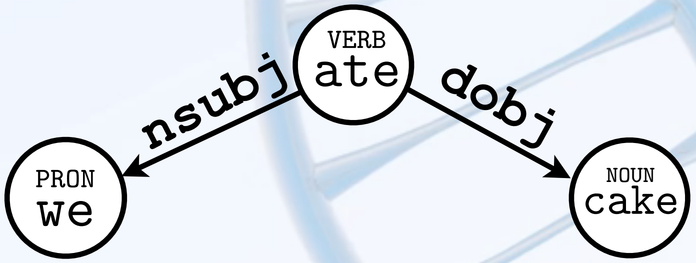
Surface Realization



log-linear model



Modeling Surface Realization



Modeling Surface Realization

S V O

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O					

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1			

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1		

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1
S O V					

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1
S O V		1	1		

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1
S O V		1	1	0	

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1
S O V		1	1	0	0

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1
S O V		1	1	0	0
O S V		0	1	0	1
O V S		0	0	0	1
V S O		0	0	1	1
V O S		0	0	1	0

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O		1	1	1	1
S O V		1	1	0	0
O S V		0	1	0	1
O V S		0	0	0	1
V S O		0	0	1	1
V O S		0	0	1	0

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O	0.8	1	1	1	1
S O V	0.03	1	1	0	0
O S V	0.1	0	1	0	1
O V S	0.02	0	0	0	1
V S O	0.03	0	0	1	1
V O S	0.02	0	0	1	0

Modeling Surface Realization

Order	Prob.	BOS S adj.	S<V	V<O	SV adj.
S V O	0.8	1	1	1	1
S O V	0.03	1	1	0	0
O S V	0.1	0	1	0	1
O V S	0.02	0	0	0	1
V S O	0.03	0	0	1	1
V O S	0.02	0	0	1	0

How to find 🧬 ?

Source

Target

POS corpus

JOHNS HOPKINS
UNIVERSITY

How to find ?

Source

Target

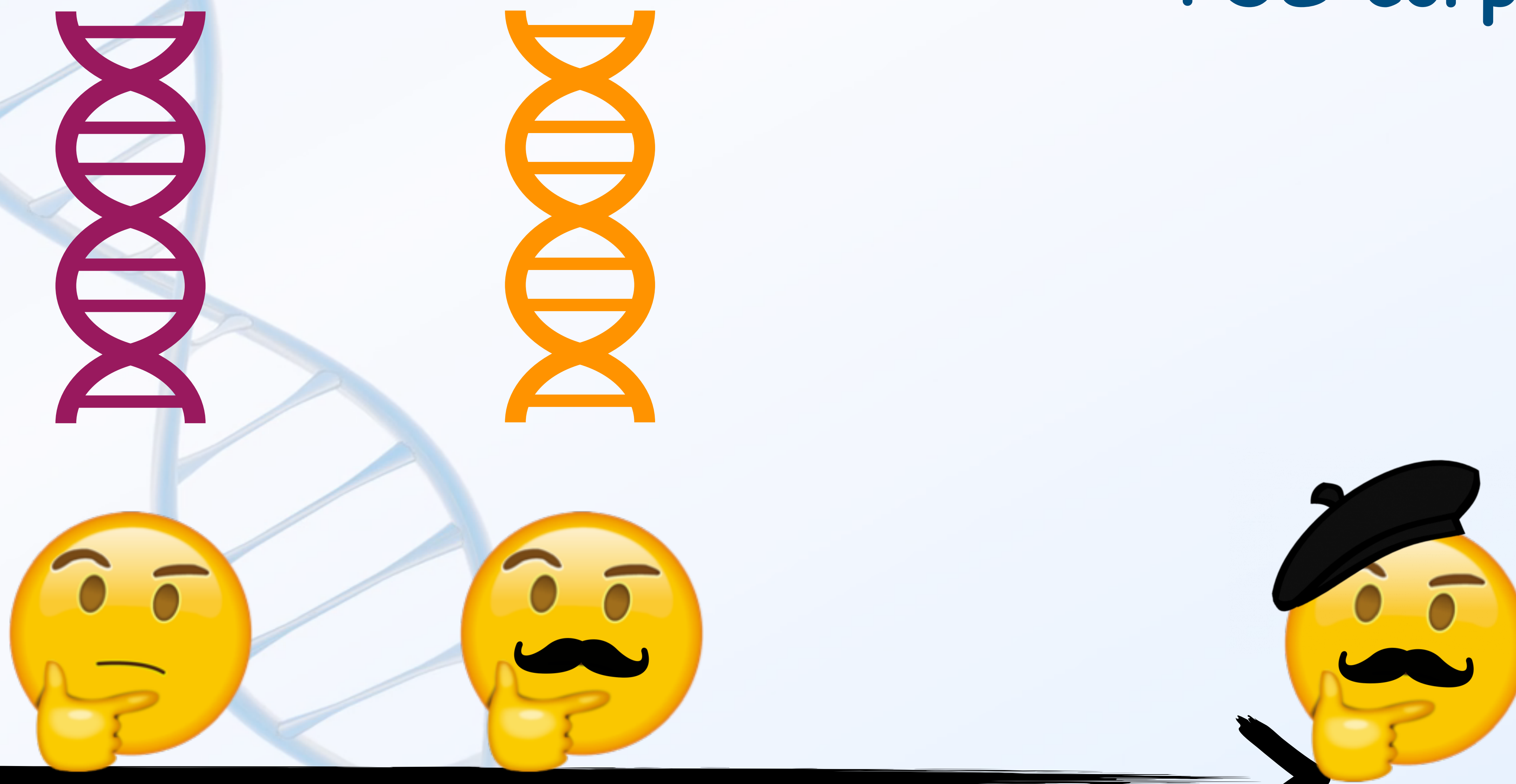
POS corpus

How to find ?

Source

Source'

Target
POS corpus



Surface similarity

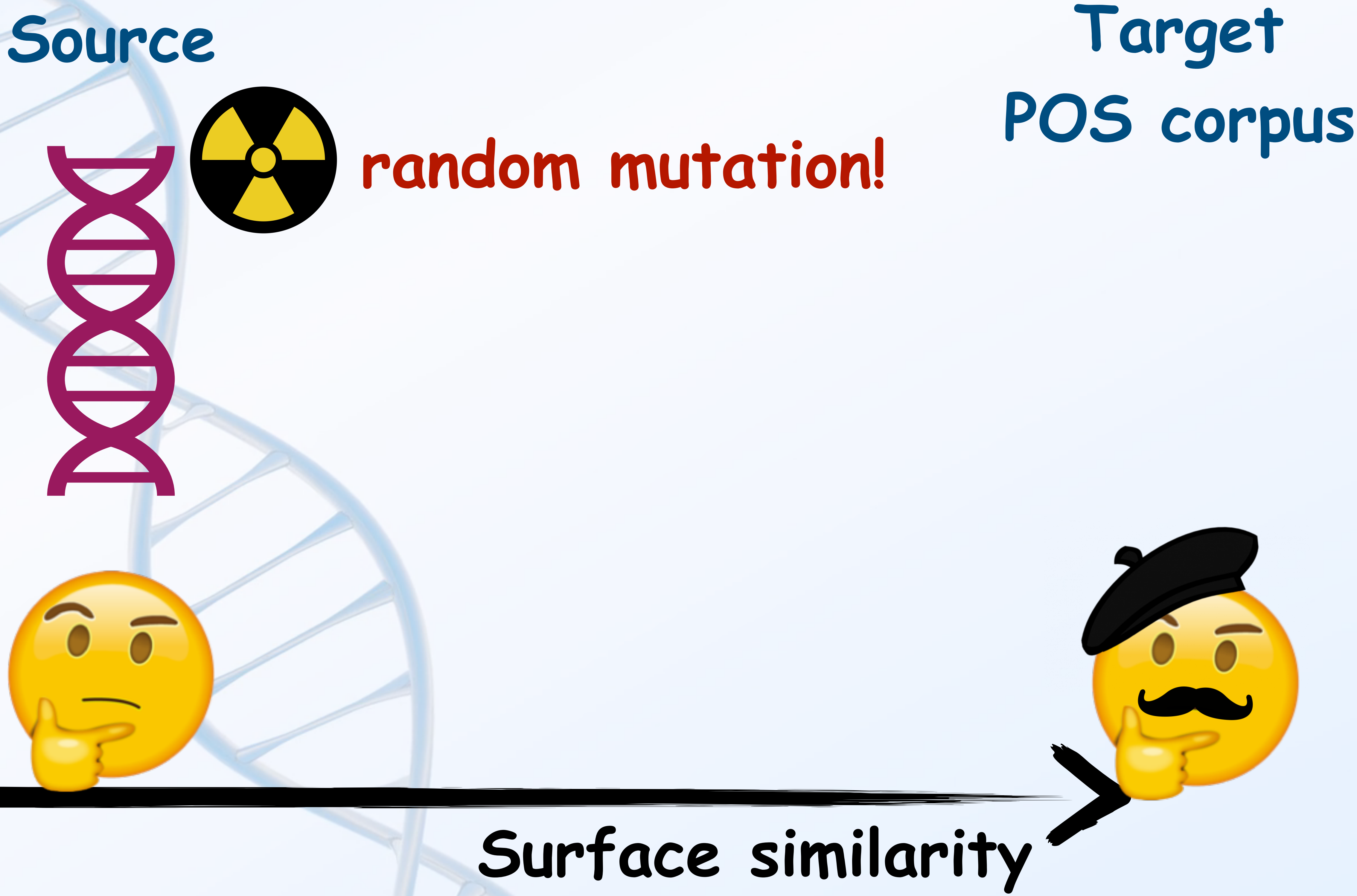
Scattershot

Source

Target
POS corpus

Surface similarity

Scattershot



Scattershot

Source

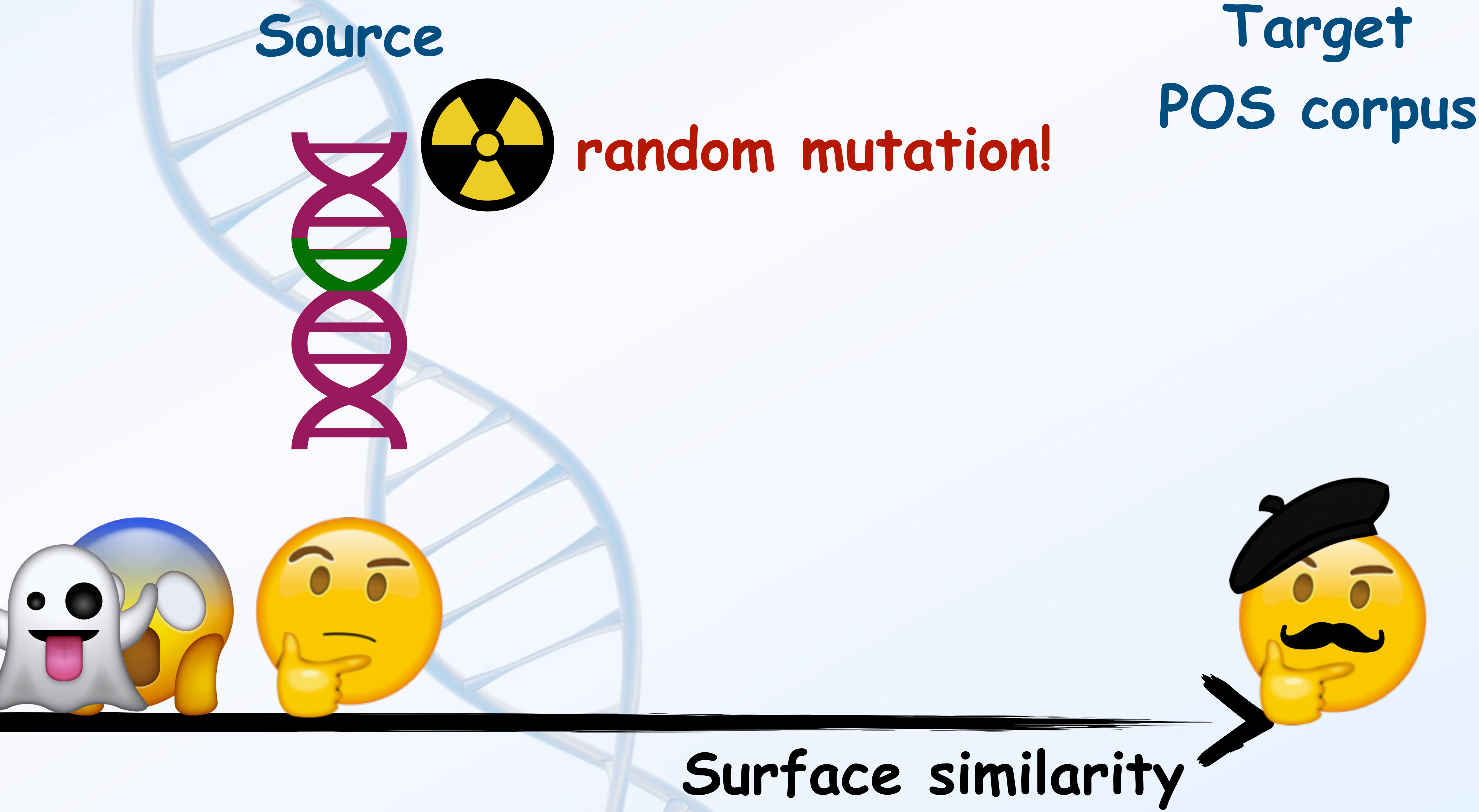
random mutation!

Target

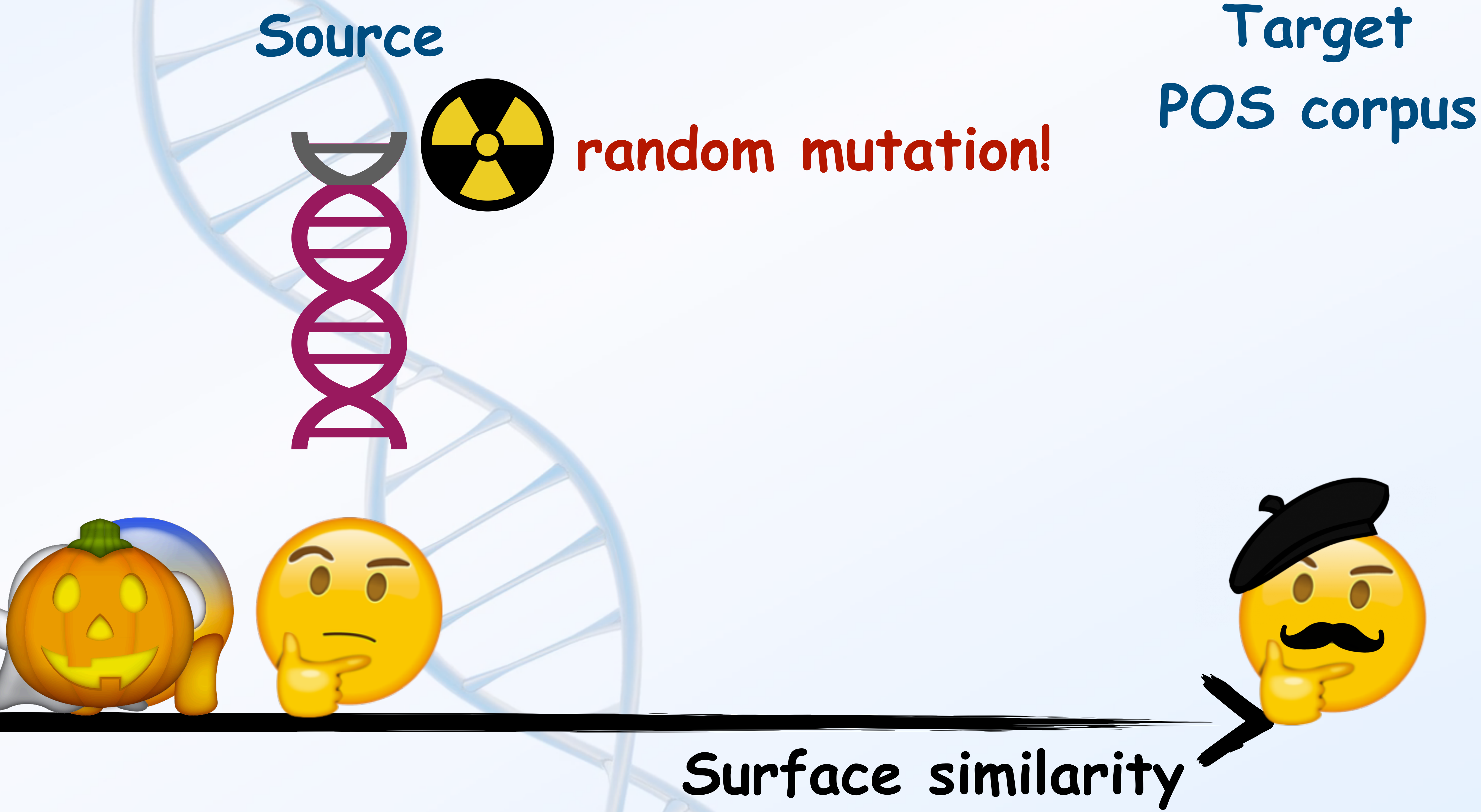
POS corpus

Surface similarity

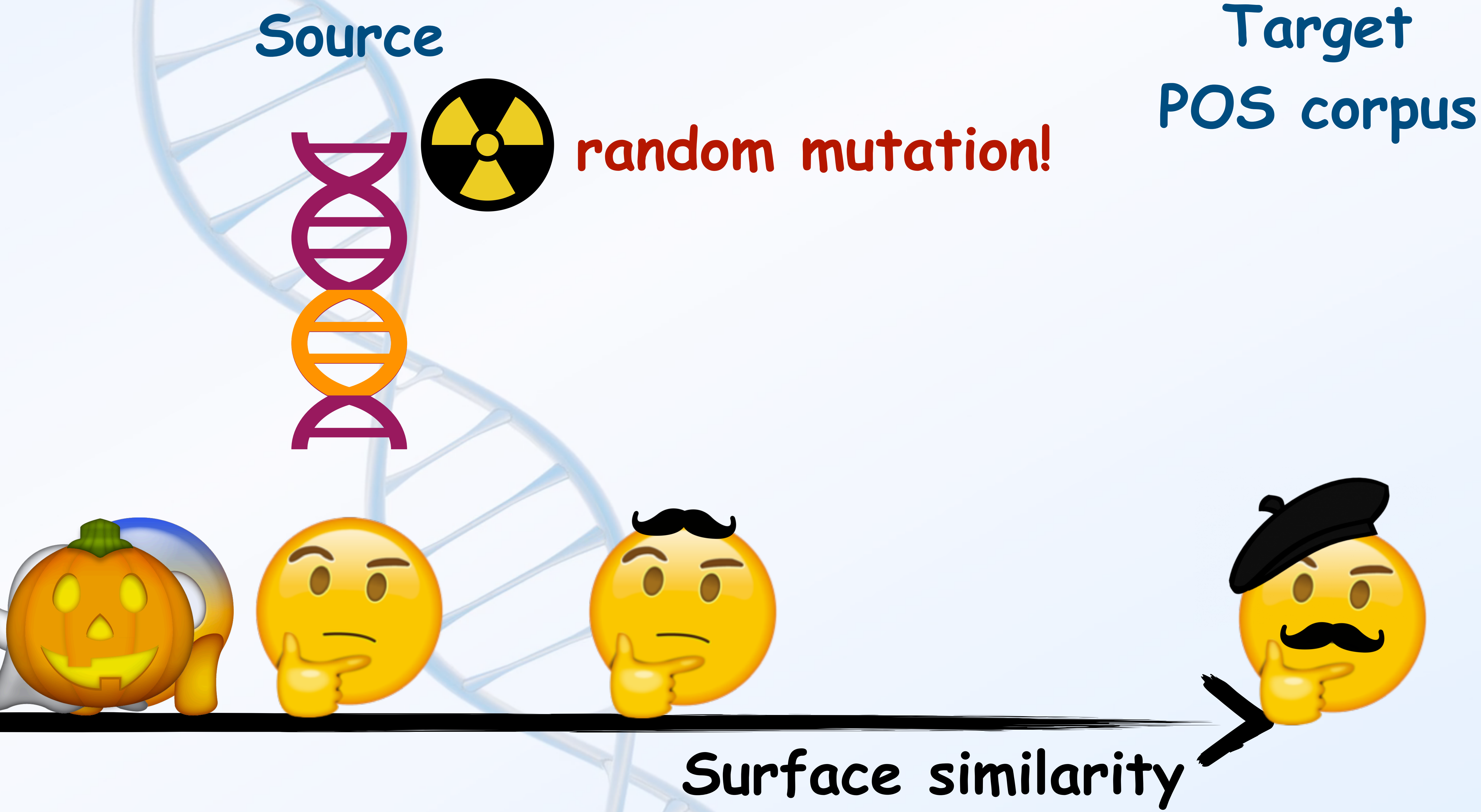
Scattershot



Scattershot



Scattershot

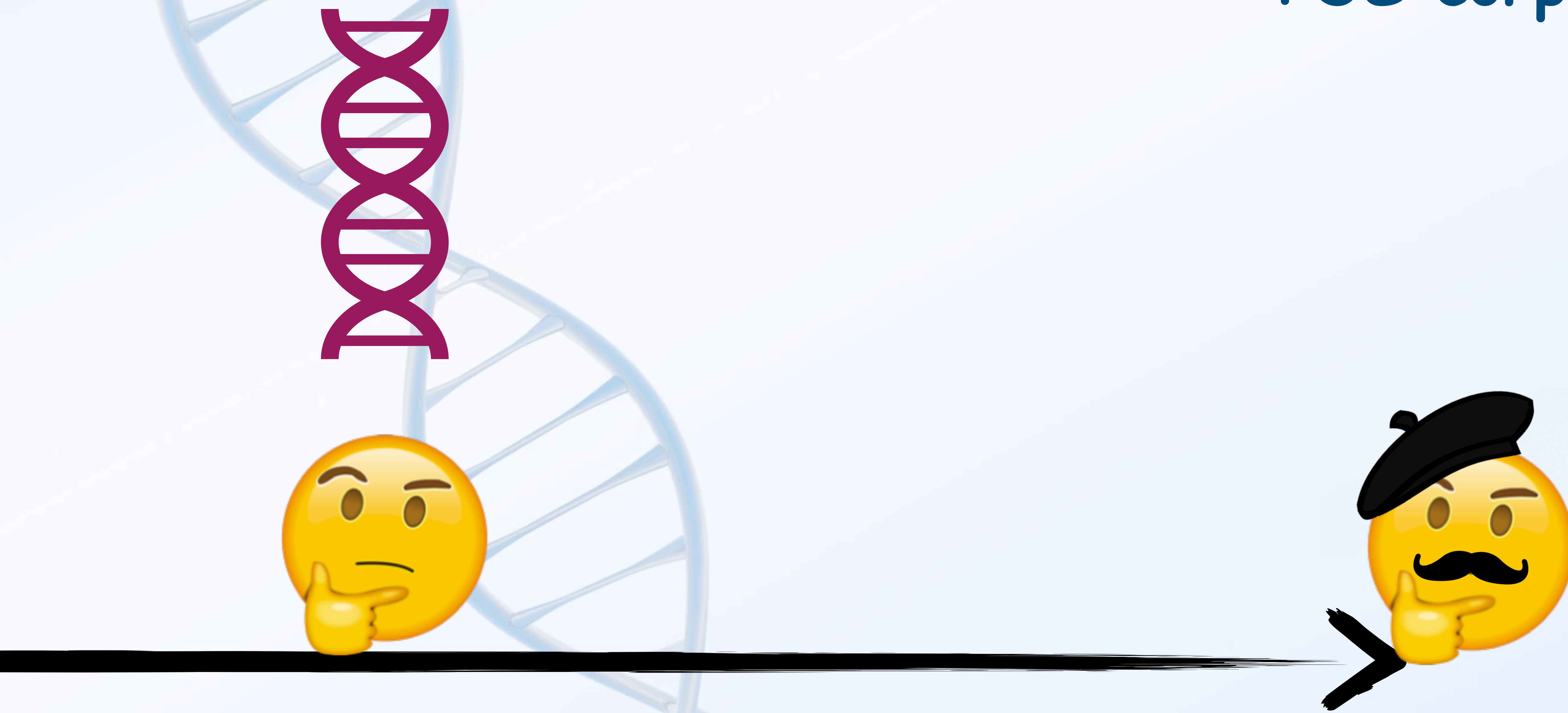


This Work: “Intelligent Design”

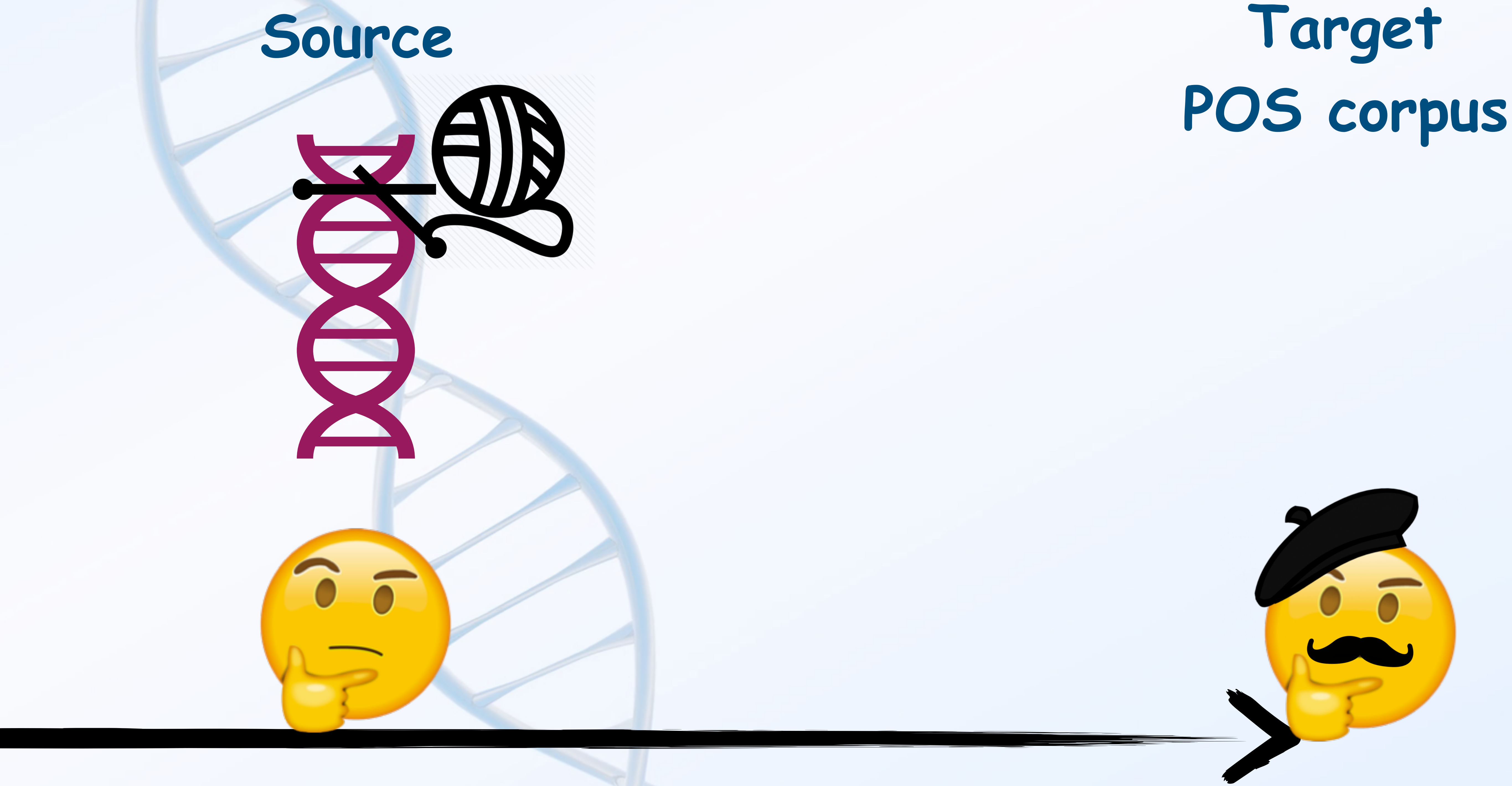
Source

Target

POS corpus



This Work: “Intelligent Design”



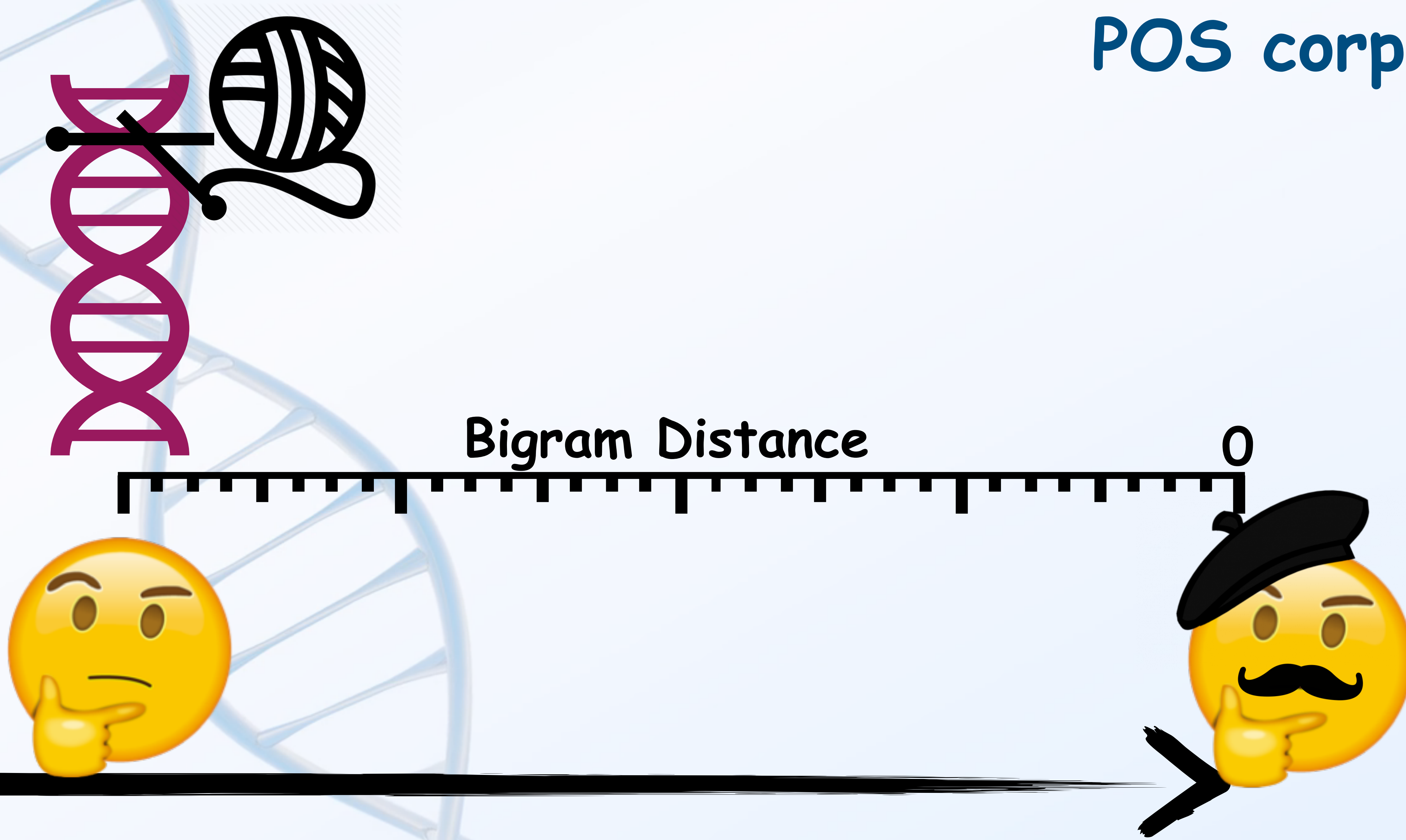
JOHNS HOPKINS
UNIVERSITY

This Work: “Intelligent Design”

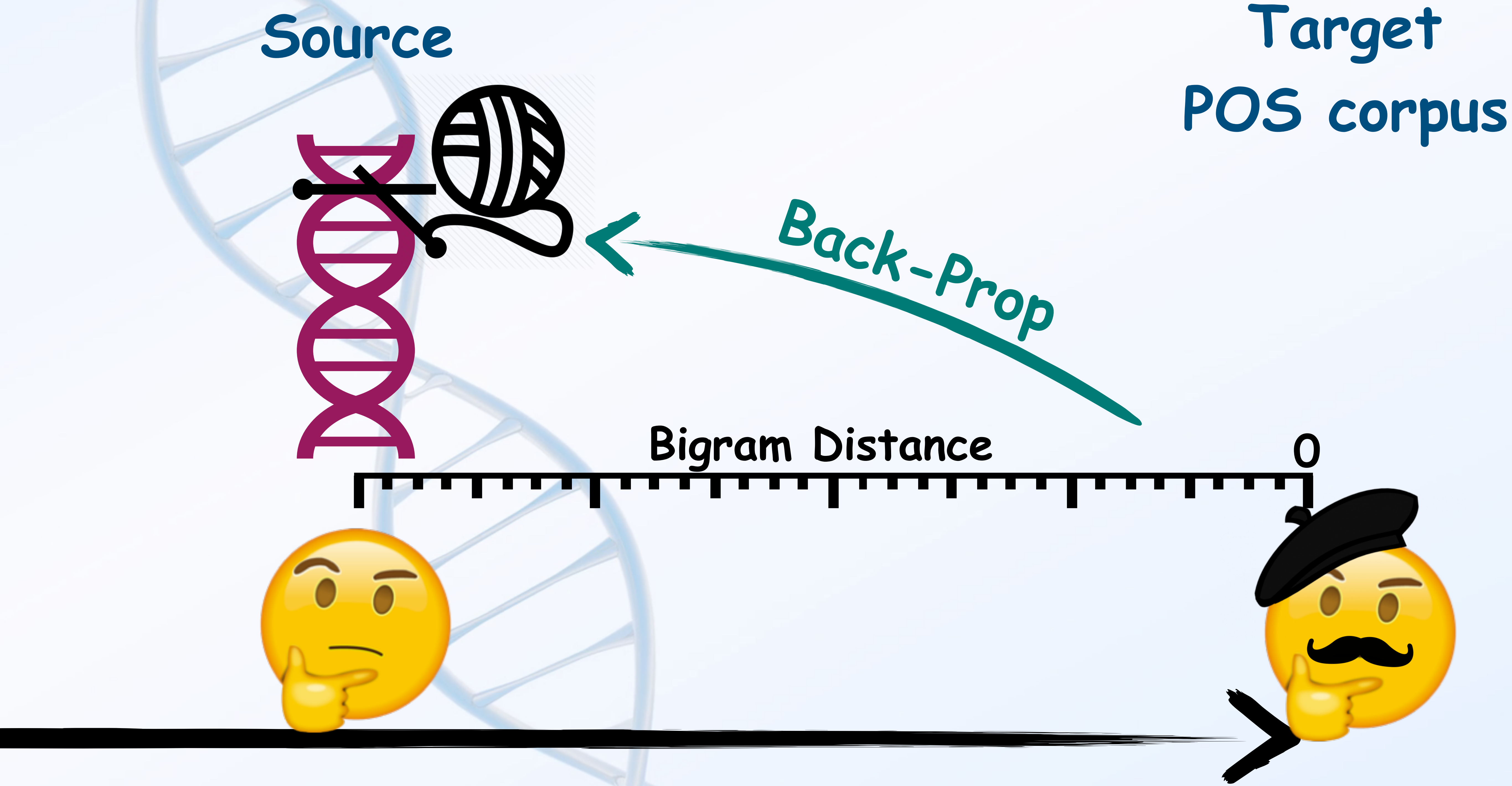
Source

Target

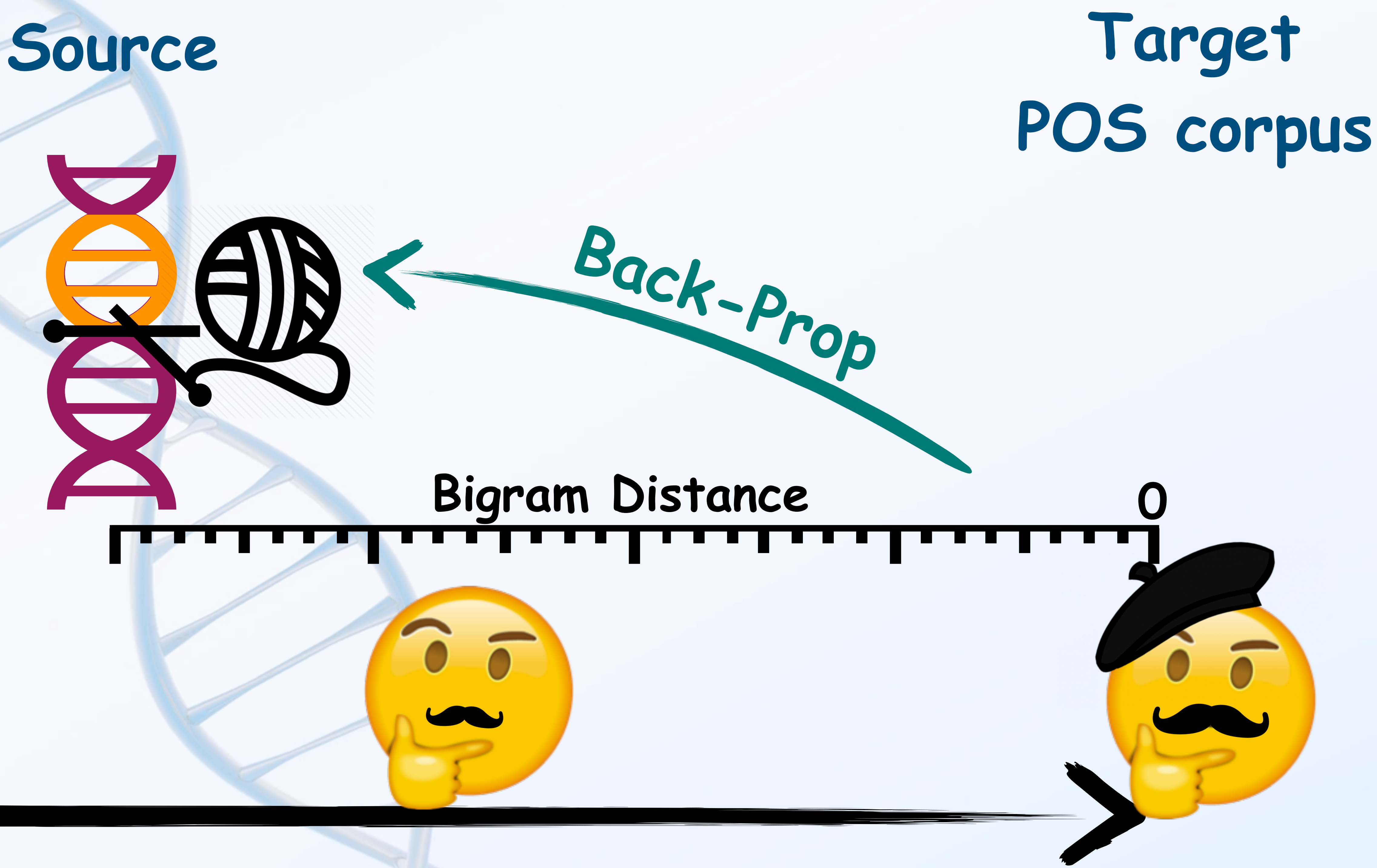
POS corpus



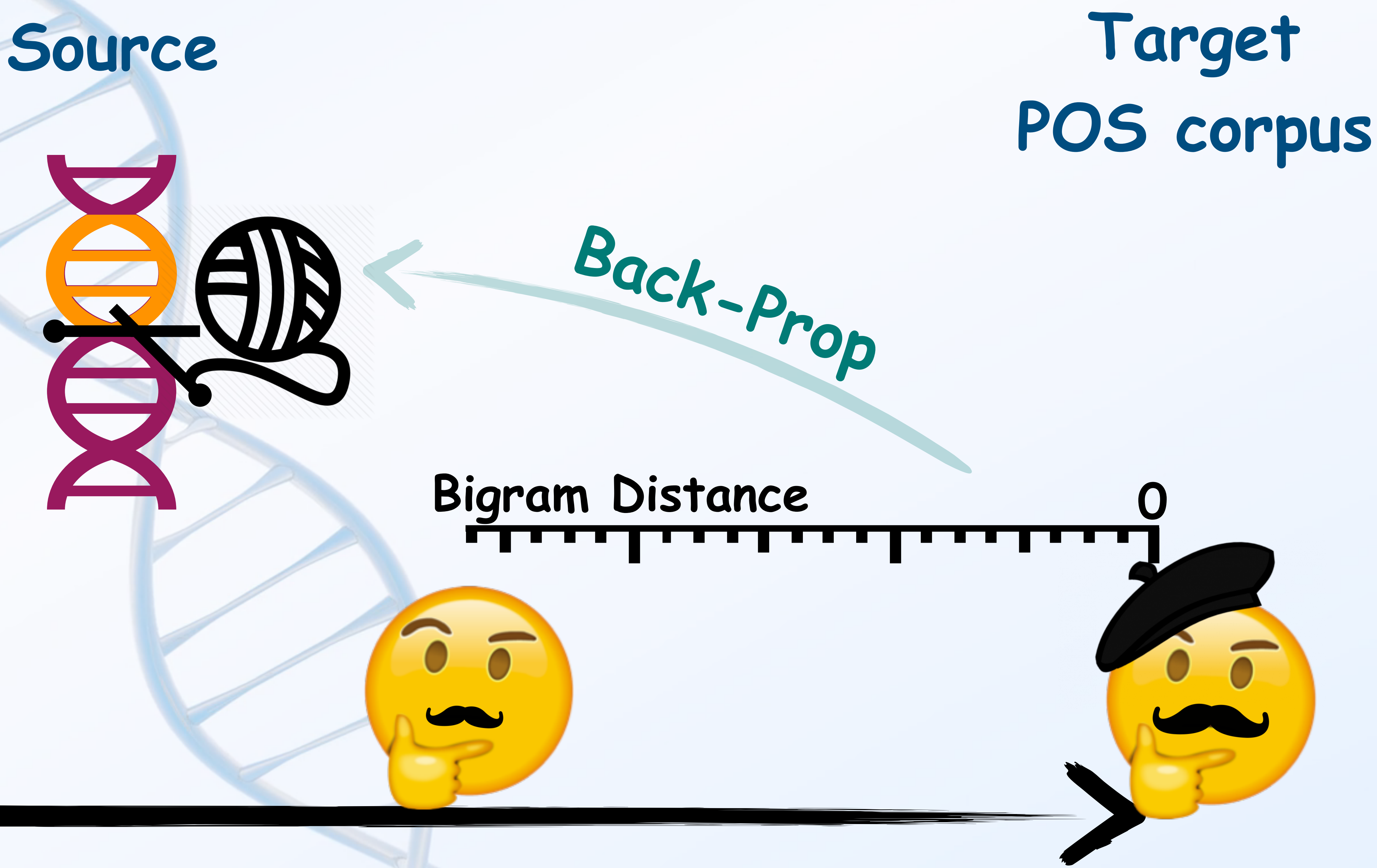
This Work: “Intelligent Design”



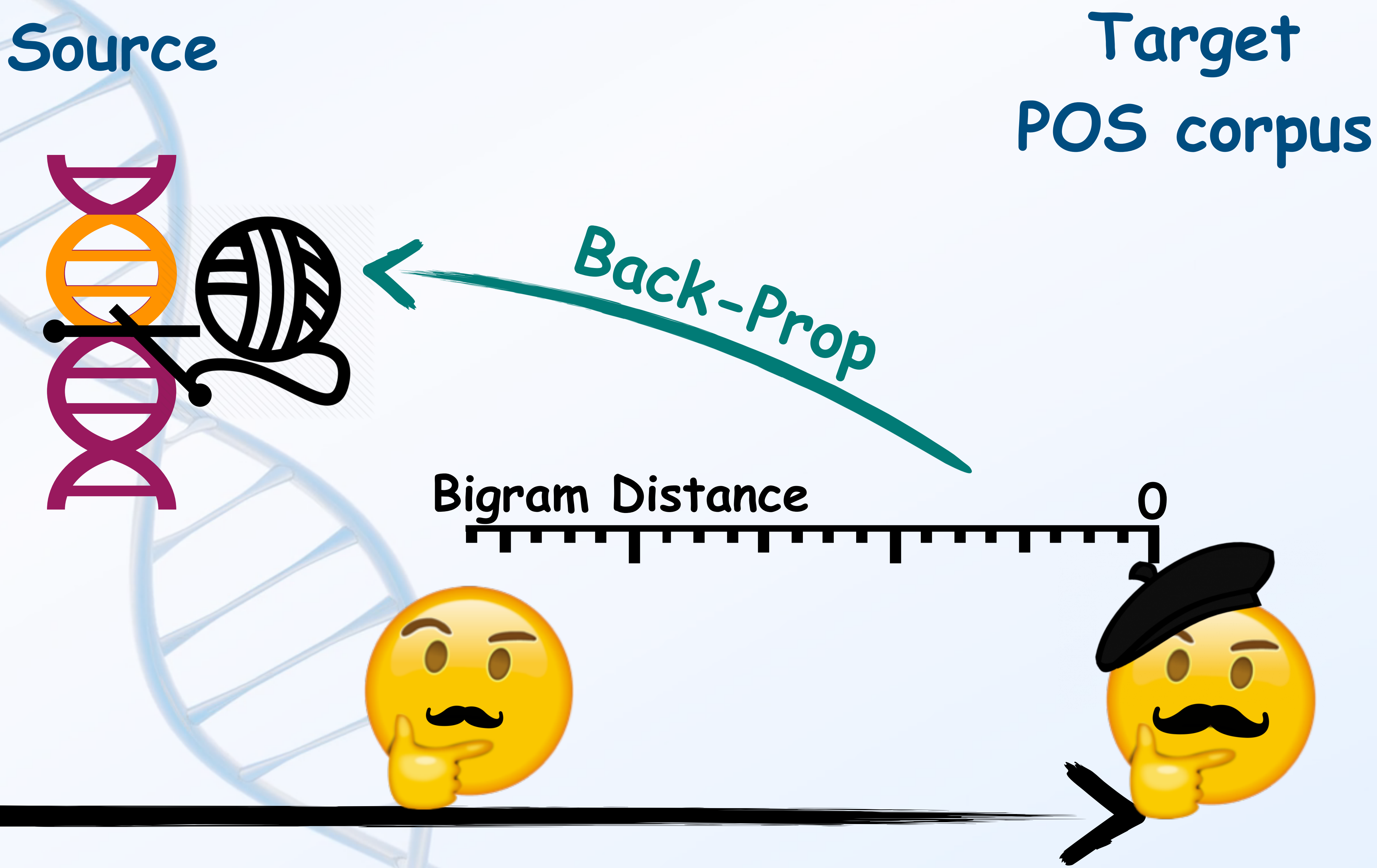
This Work: “Intelligent Design”



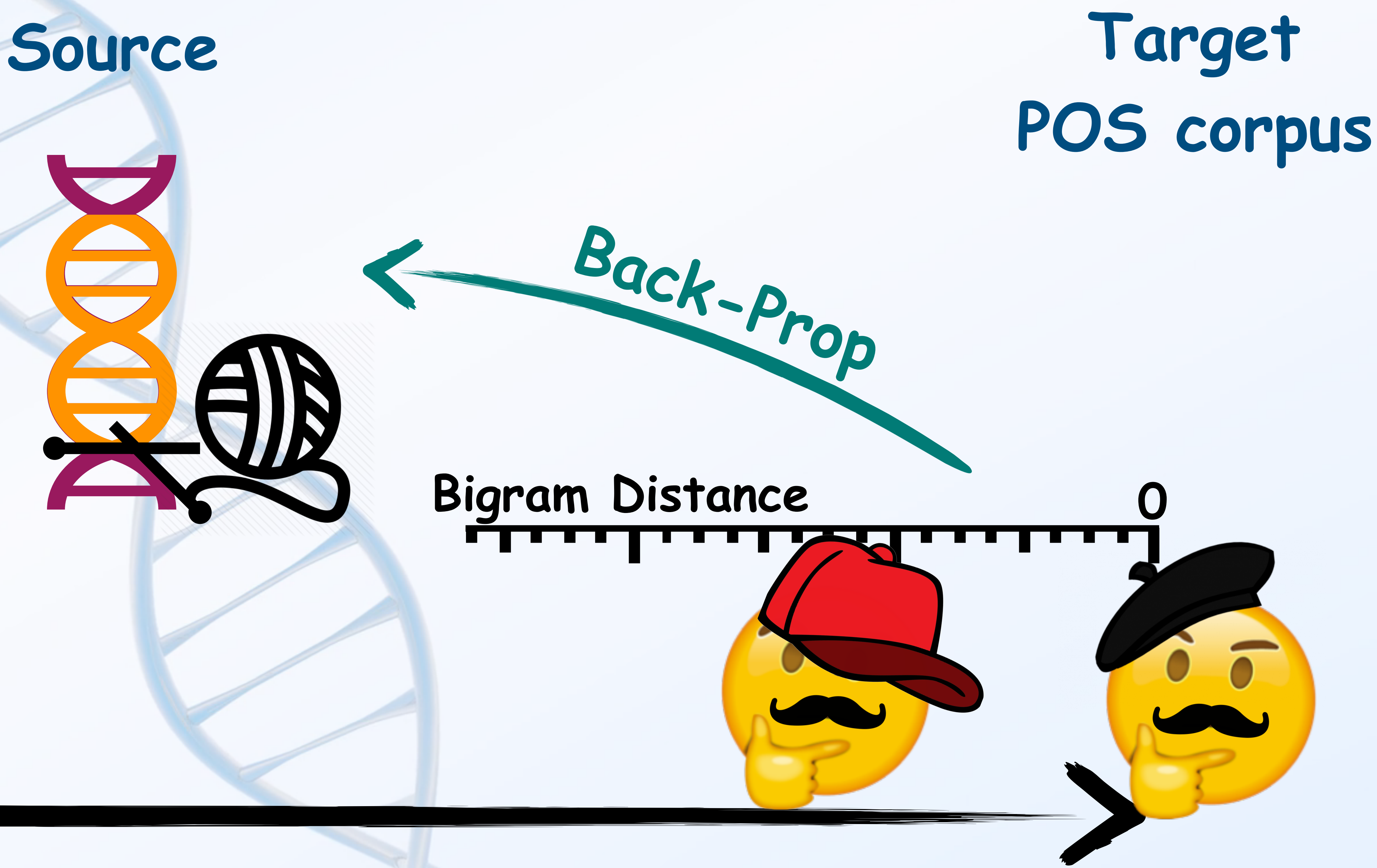
This Work: “Intelligent Design”



This Work: "Intelligent Design"



This Work: “Intelligent Design”



Bigram Distance

Bigram Distance

- Whether a bigram language model trained on the source' could give a high likelihood on the target POS-corpus?

$$\texttt{p(ADJ | NOUN)} = \frac{\texttt{\#NOUN ADJ}}{\texttt{\#NOUN}}$$

Bigram Distance

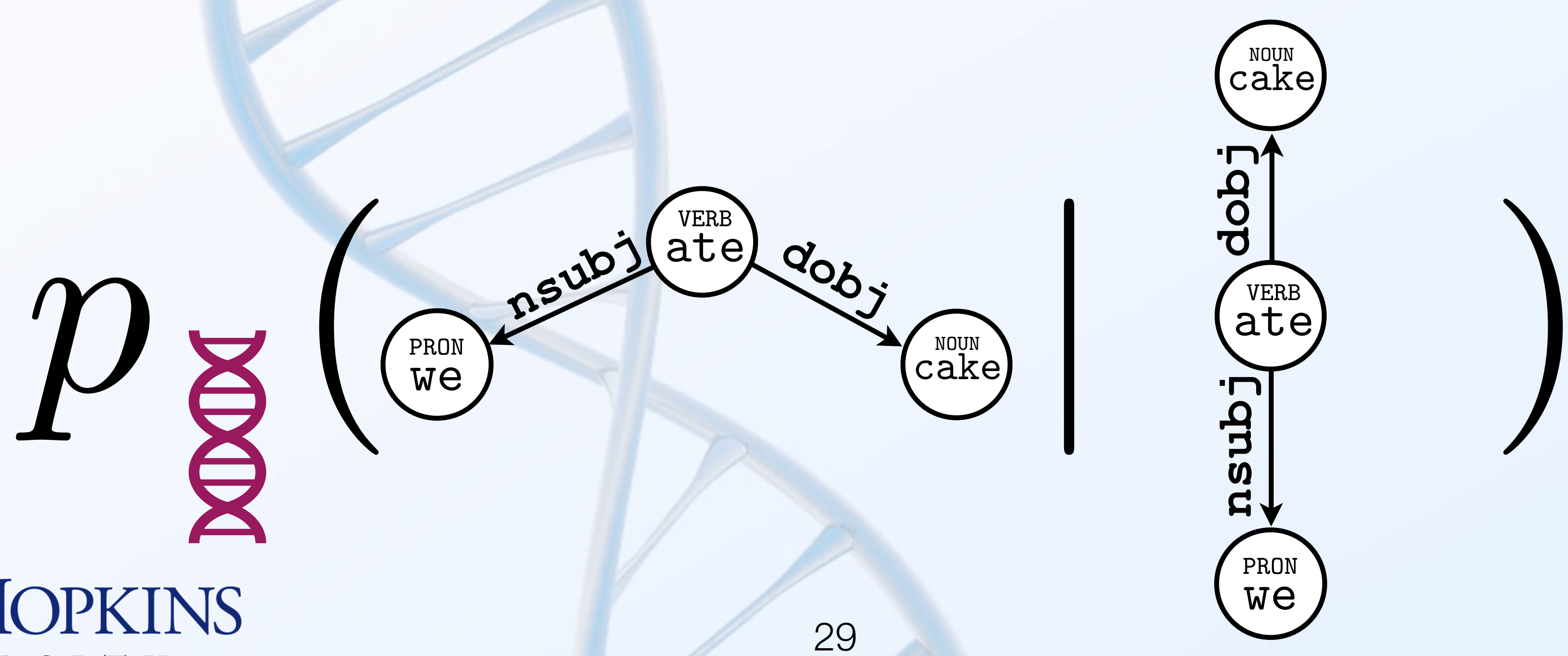
- Whether a bigram language model trained on the source' could give a high likelihood on the target POS-corpus?

$$\texttt{p(ADJ | NOUN)} = \frac{\texttt{\#NOUN ADJ}}{\texttt{\#NOUN}}$$

Bigram Distance

- Whether a bigram language model trained on the source' could give a high likelihood on the target POS-corpus?

$$\texttt{p(ADJ | NOUN)} = \frac{\texttt{\#NOUN ADJ}}{\texttt{\#NOUN}}$$



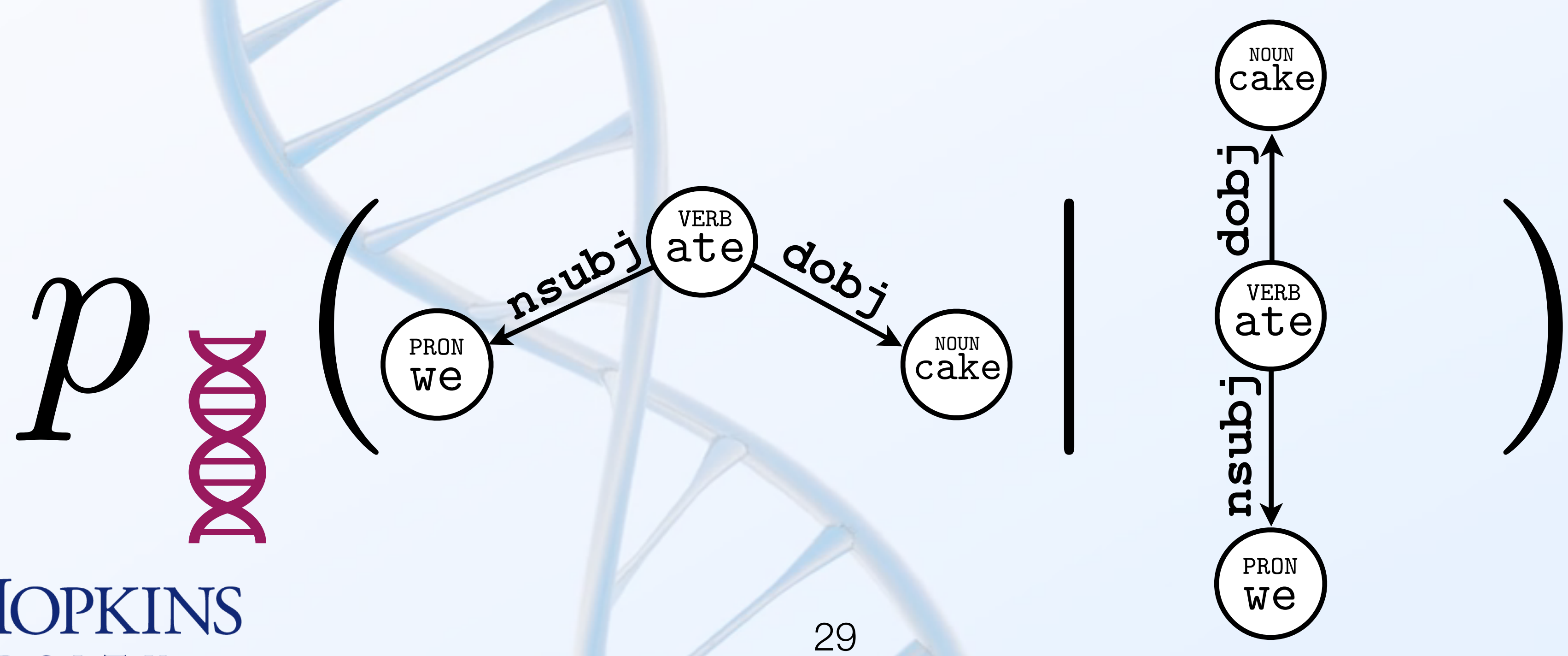
Johns Hopkins University

Bigram Distance

- Whether a bigram language model trained on the source' could give a high likelihood on the target POS-corpus?

$$\texttt{p(ADJ | NOUN)} = \frac{\texttt{\#NOUN ADJ}}{\texttt{\#NOUN}}$$

- How to compute those POS-bigrams counts?

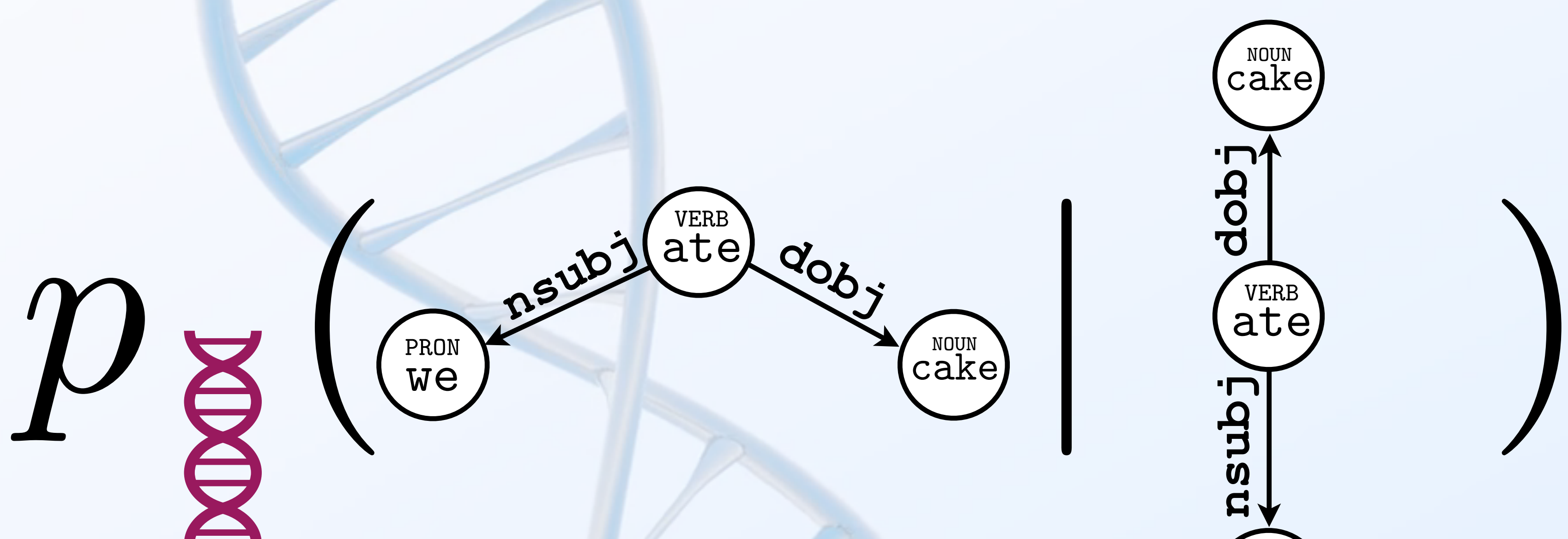


Bigram Distance

- Whether a bigram language model trained on the source' could give a high likelihood on the target POS-corpus?

$$\texttt{p(ADJ | NOUN)} = \frac{\mathbb{E}[\#\text{NOUN ADJ}]}{\#\text{NOUN}}$$

- How to compute those POS-bigrams counts?
  - **Expected Counts** from !



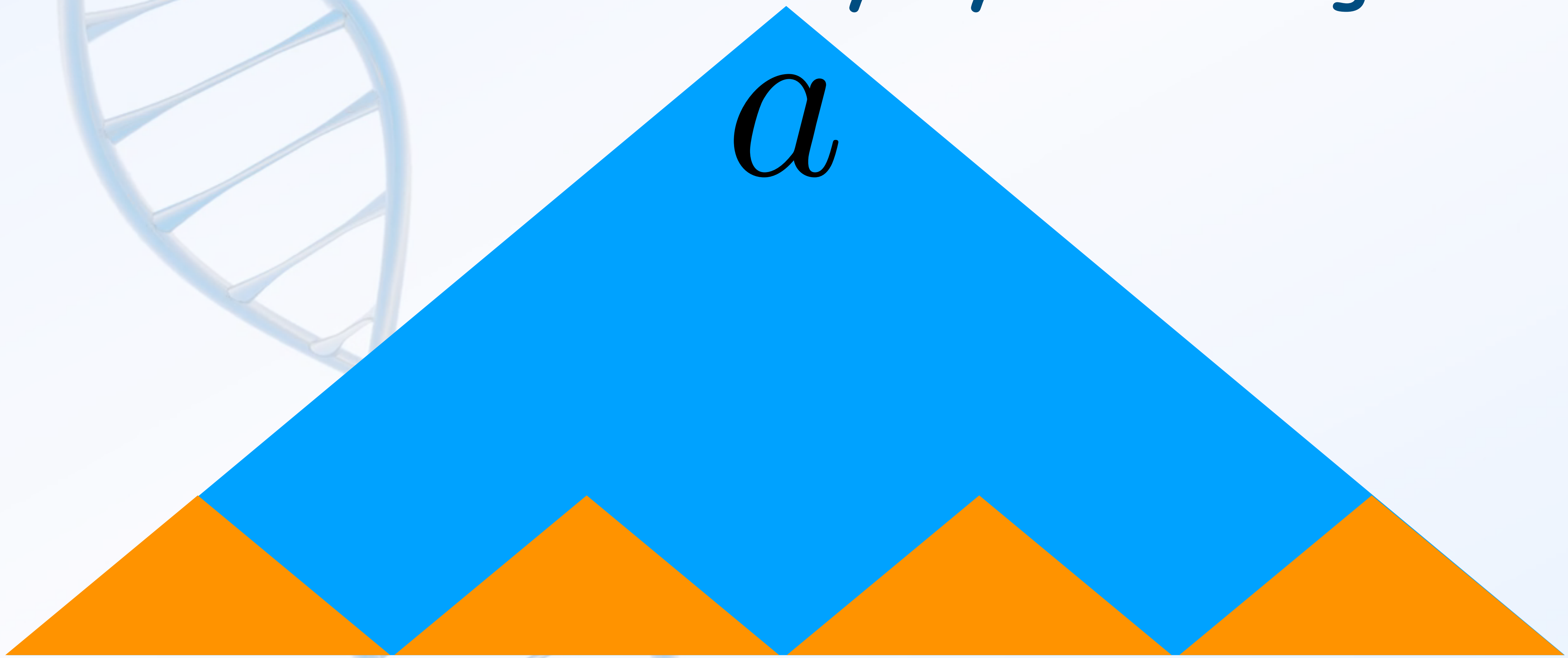
Computing Expected Counts

Computing Expected Counts

by Dynamic Programming

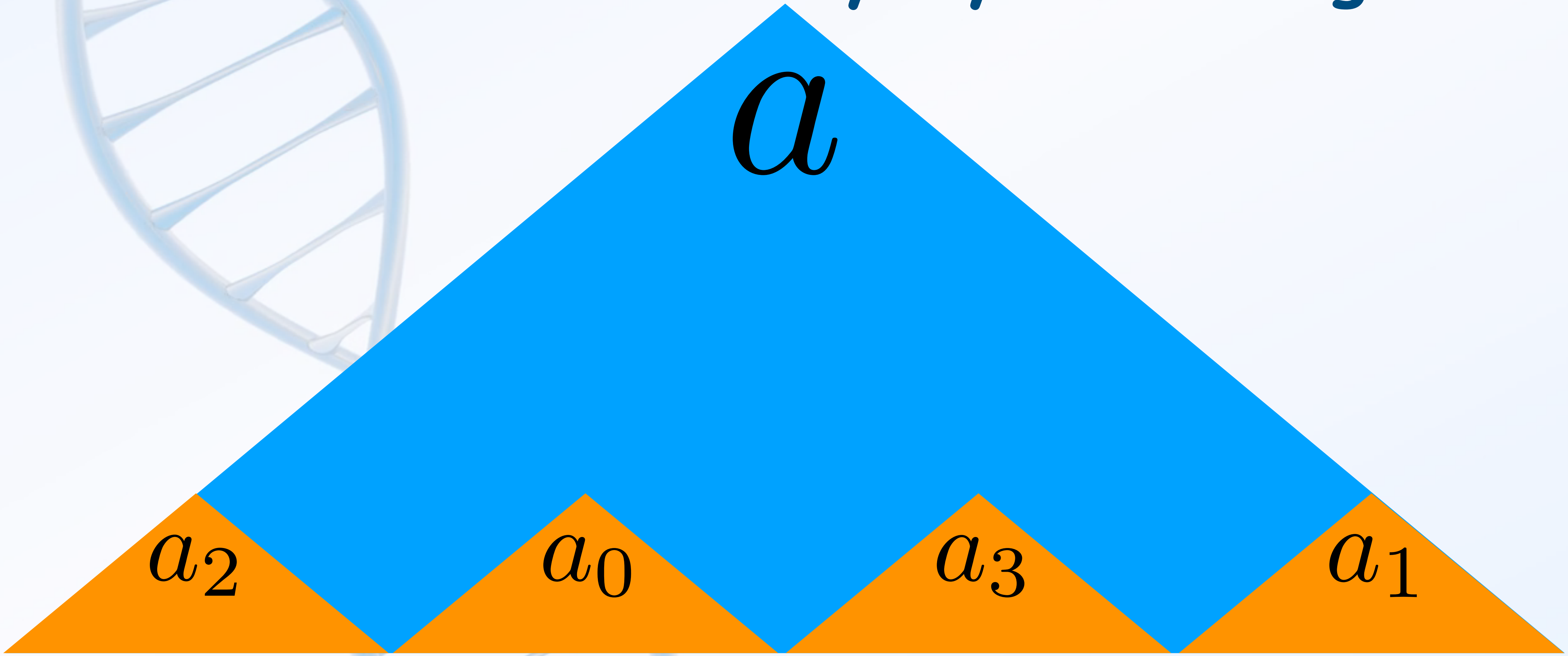
Computing Expected Counts

by Dynamic Programming



Computing Expected Counts

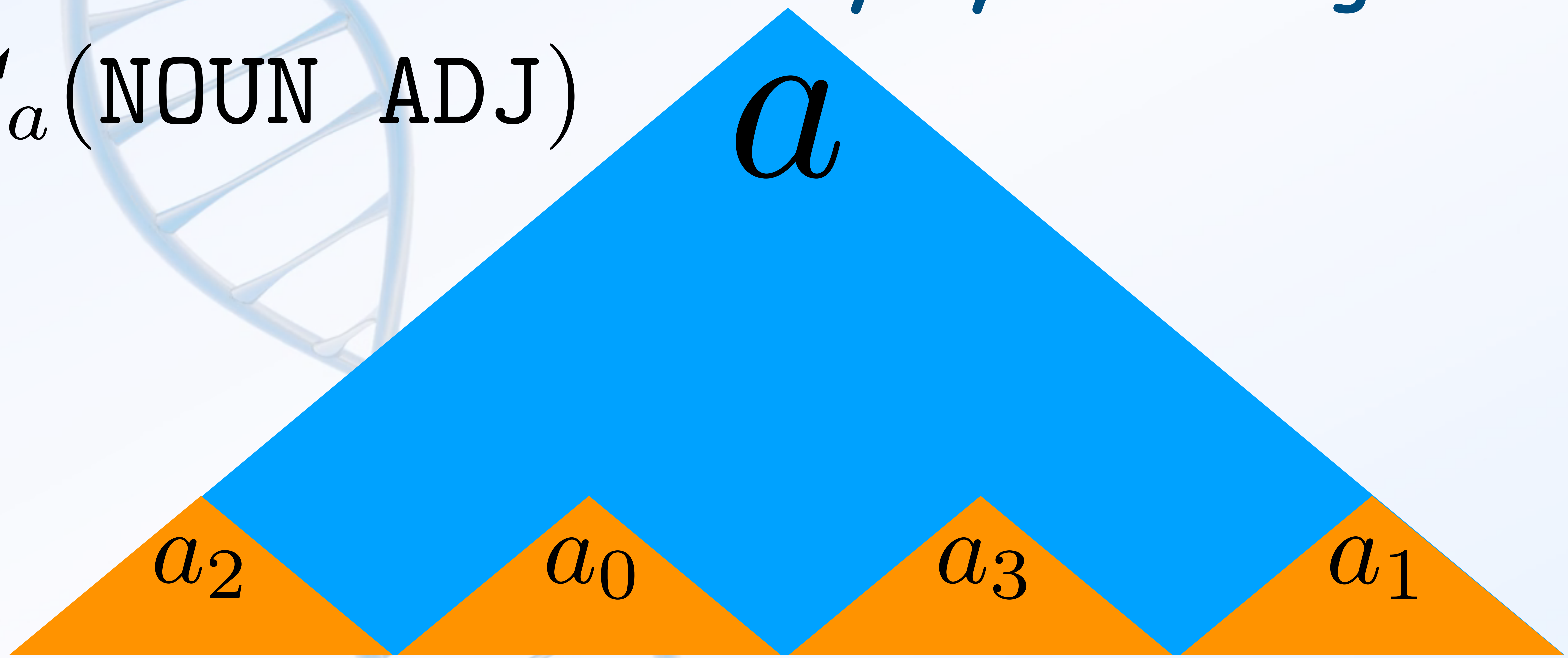
by Dynamic Programming



Computing Expected Counts

by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$



Computing Expected Counts
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$

Computing Expected Counts

by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$

Johns Hopkins University

Computing Expected Counts

by Dynamic Programming

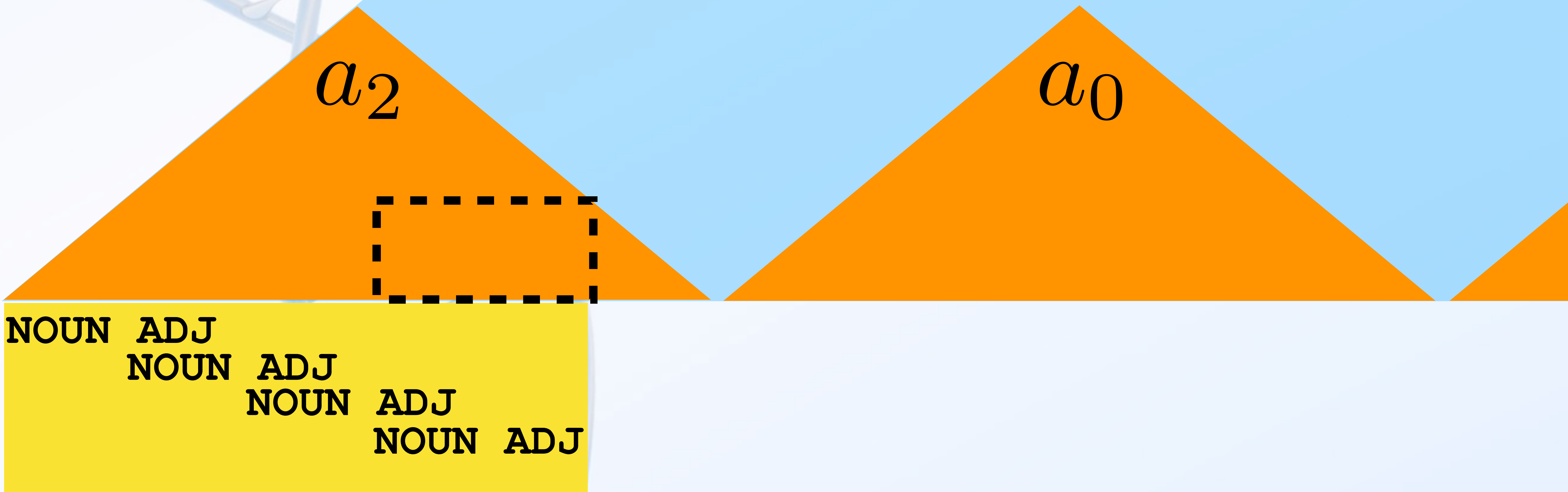
$$C_a(\texttt{NOUN ADJ})$$

JOHNS HOPKINS UNIVERSITY

Computing Expected Counts

by Dynamic Programming

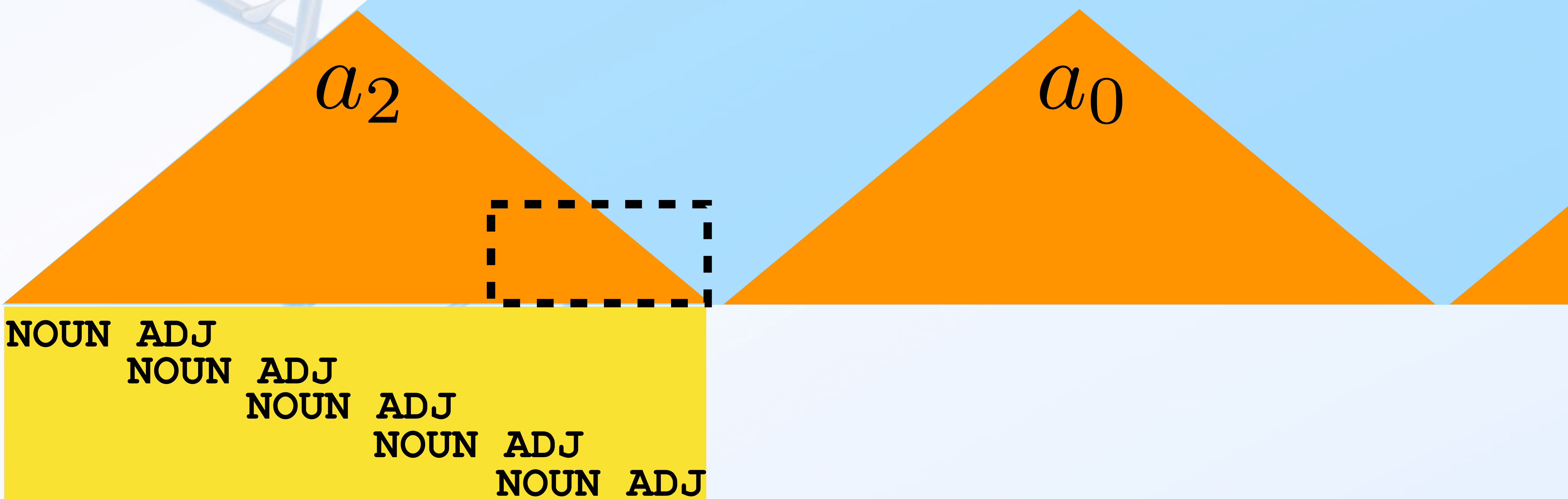
$$C_a(\texttt{NOUN ADJ})$$



Computing Expected Counts

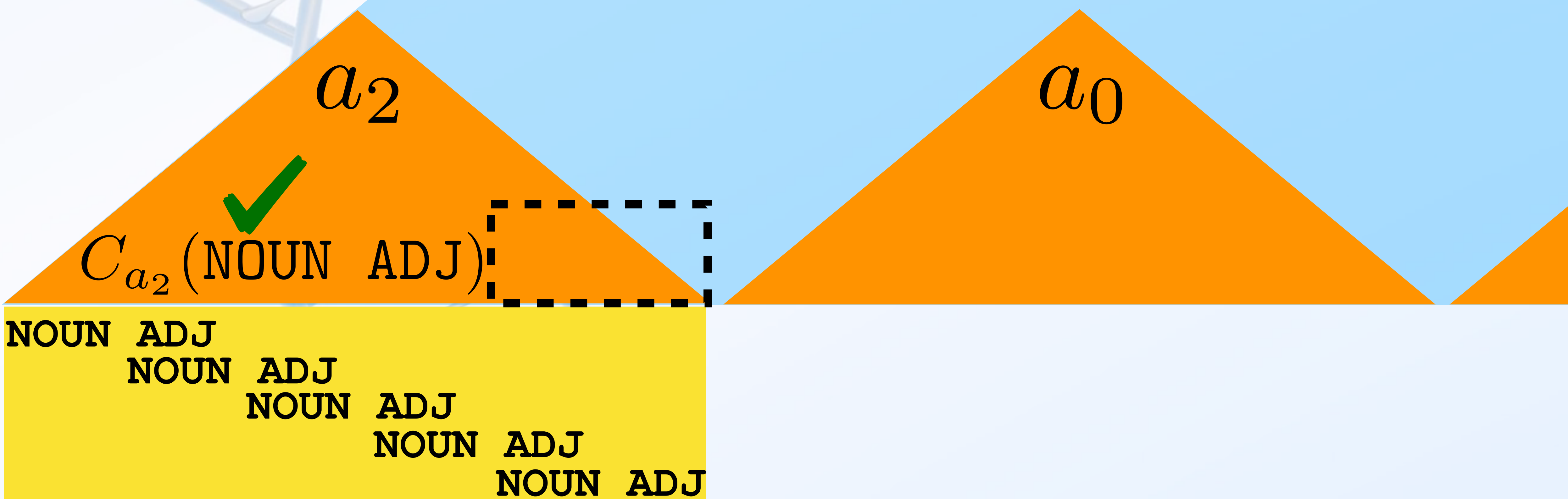
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$



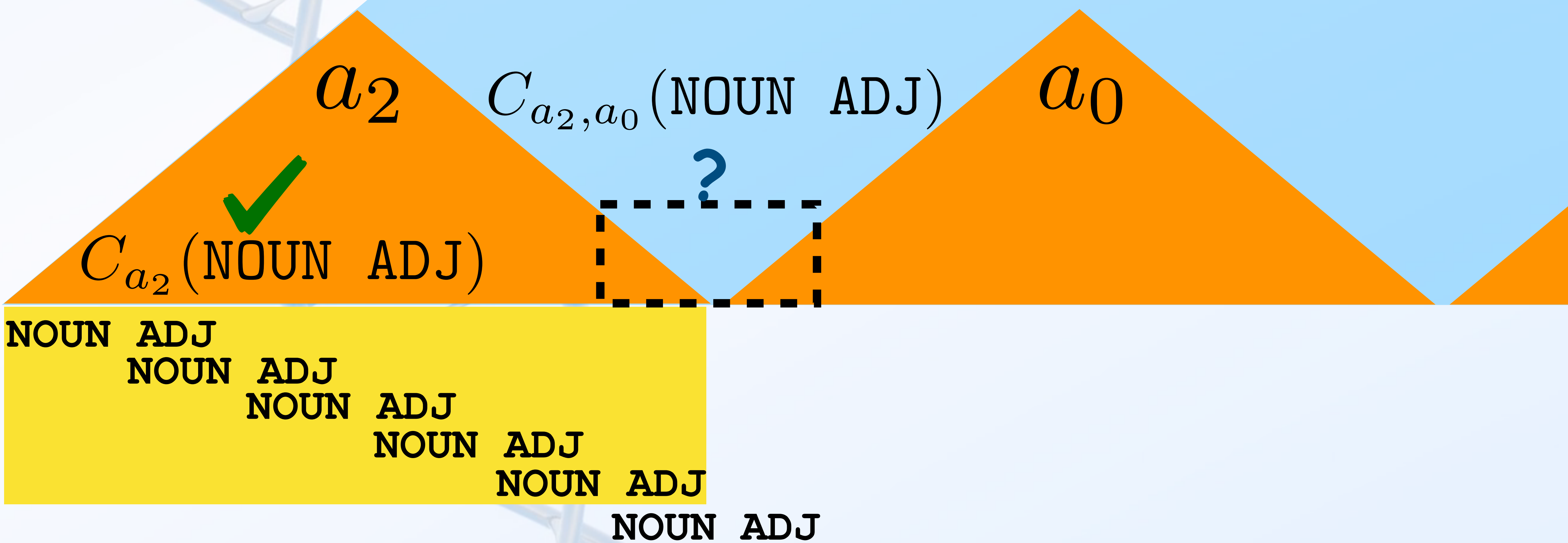
Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$



Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$



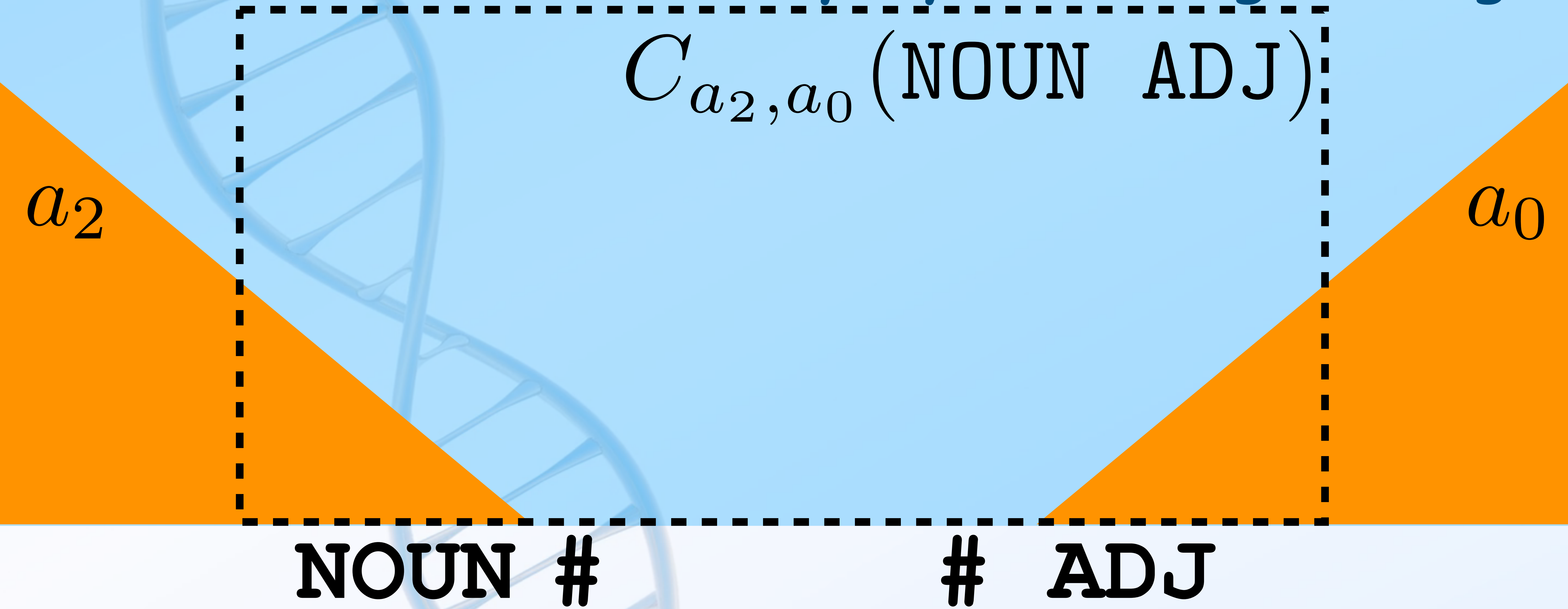
Computing Expected Counts

by Dynamic Programming

NOUN ADJ

Computing Expected Counts

by Dynamic Programming



Computing Expected Counts

by Dynamic Programming

$$C_{a_2,a_0}(\texttt{NOUN ADJ})$$

a_2

a_0

NOUN #

 # ADJ

$$C_{a_2}(\texttt{NOUN \#}) \qquad C_{a_0}(\texttt{\# ADJ})$$

JOHNS HOPKINS UNIVERSITY

Computing Expected Counts

by Dynamic Programming

$C_{a_2,a_0}(\texttt{NOUN ADJ})$

a_2 a_0

NOUN # **# ADJ**

$C_{a_2}(\texttt{NOUN \#})$ ✓

$C_{a_0}(\texttt{\# ADJ})$ ✓

Computing Expected Counts

by Dynamic Programming

$C_{a_2,a_0}(\texttt{NOUN ADJ})$

a_2

a_0

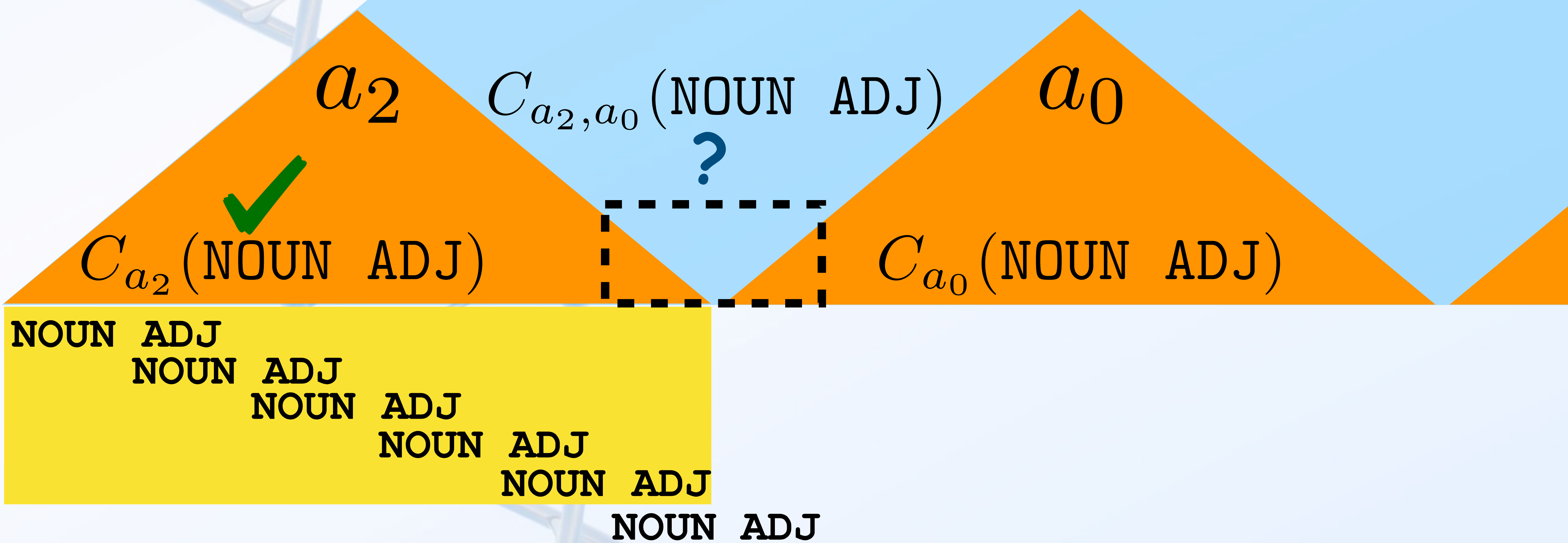
NOUN # **# ADJ**

$$C_{a_2}(\texttt{NOUN \#}) \times C_{a_0}(\texttt{\# ADJ})$$

✔ ✔

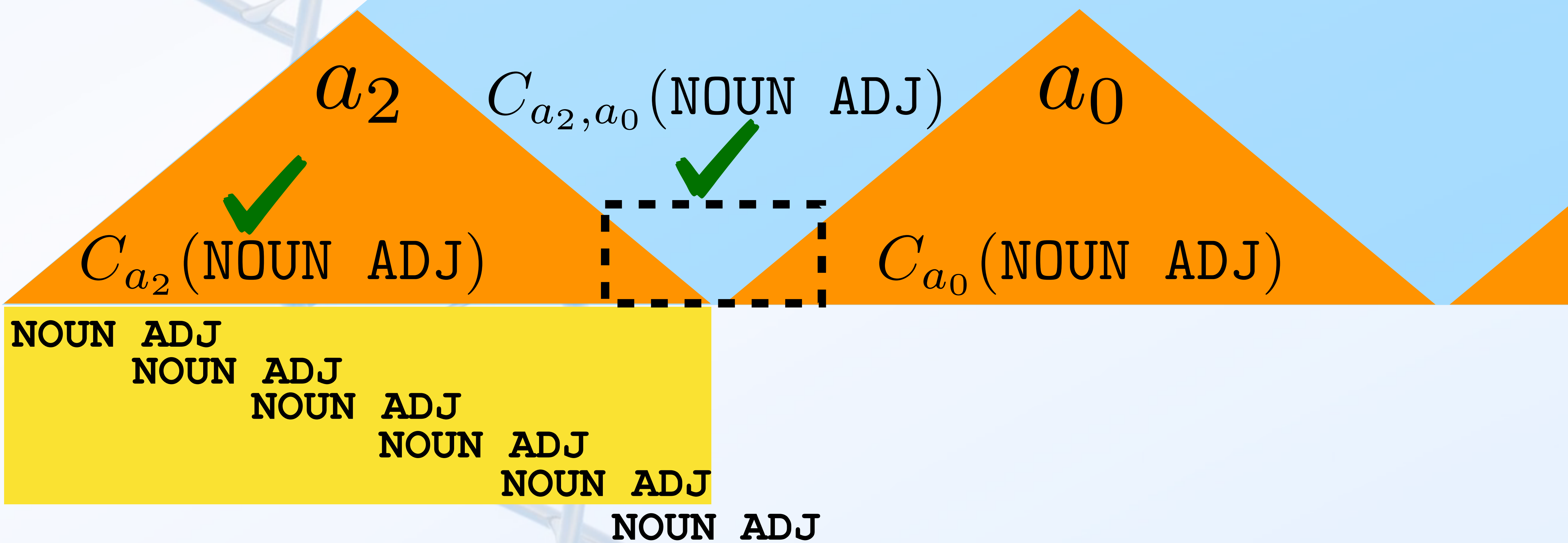
Computing Expected Counts
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$



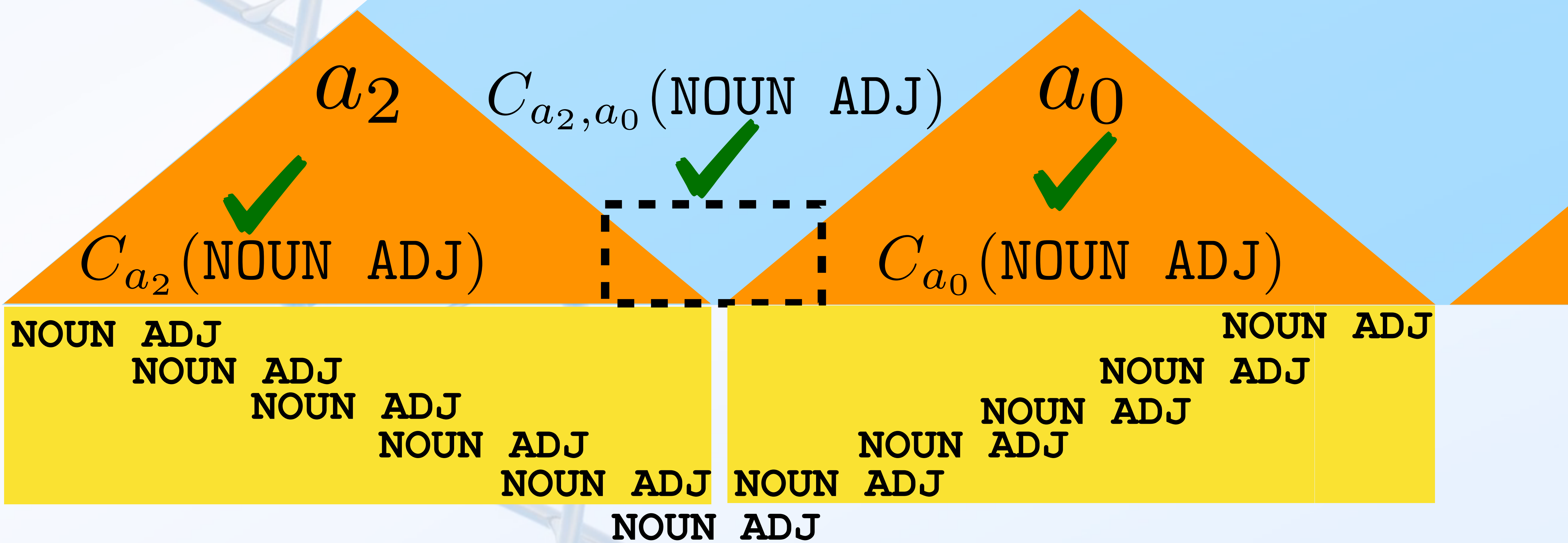
Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$



Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$

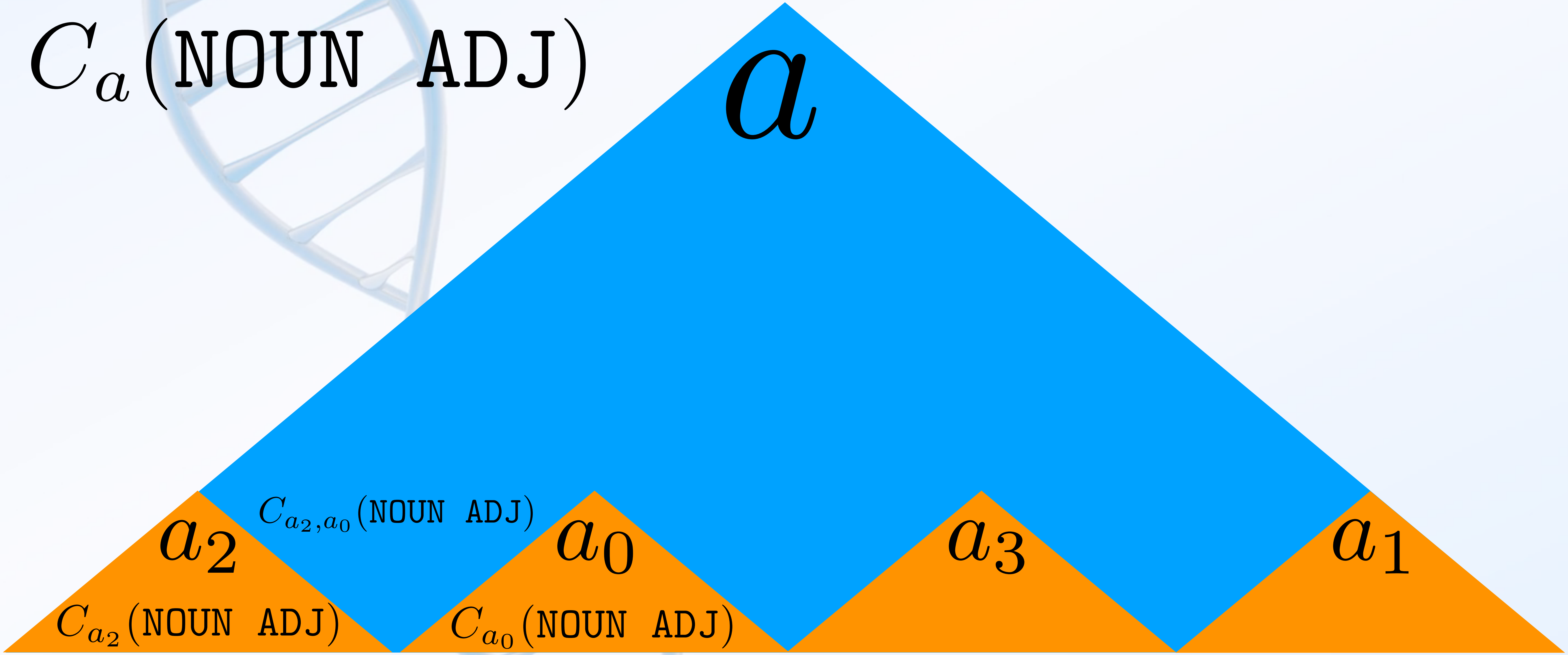


JOHNS HOPKINS UNIVERSITY

Computing Expected Counts

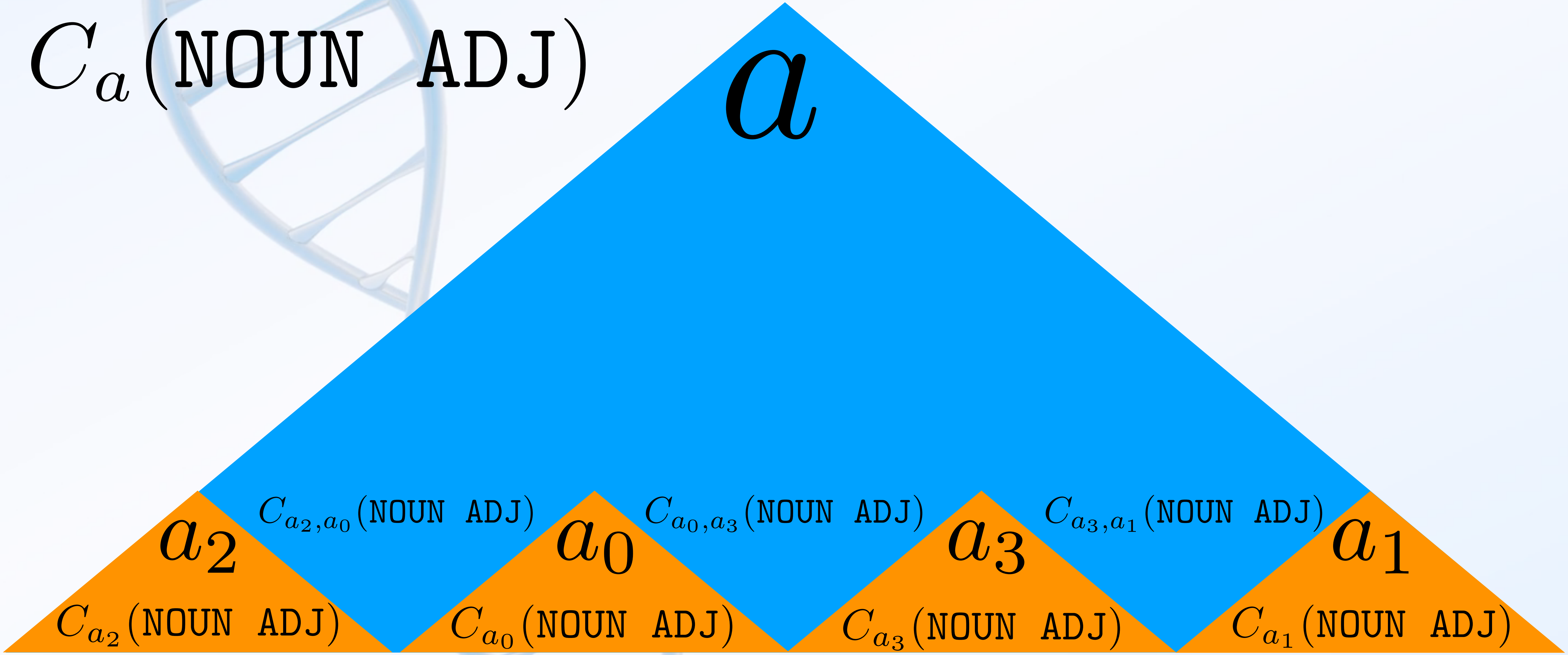
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

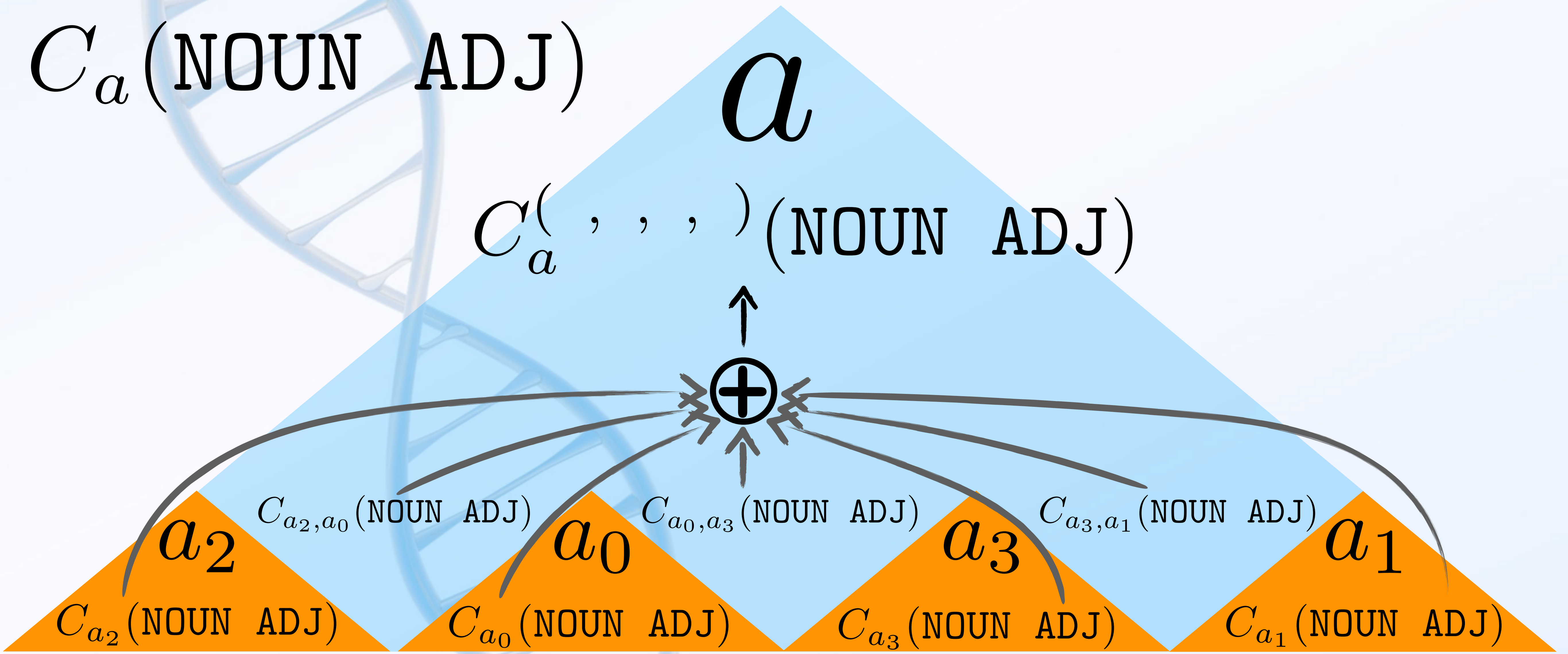


Computing Expected Counts
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

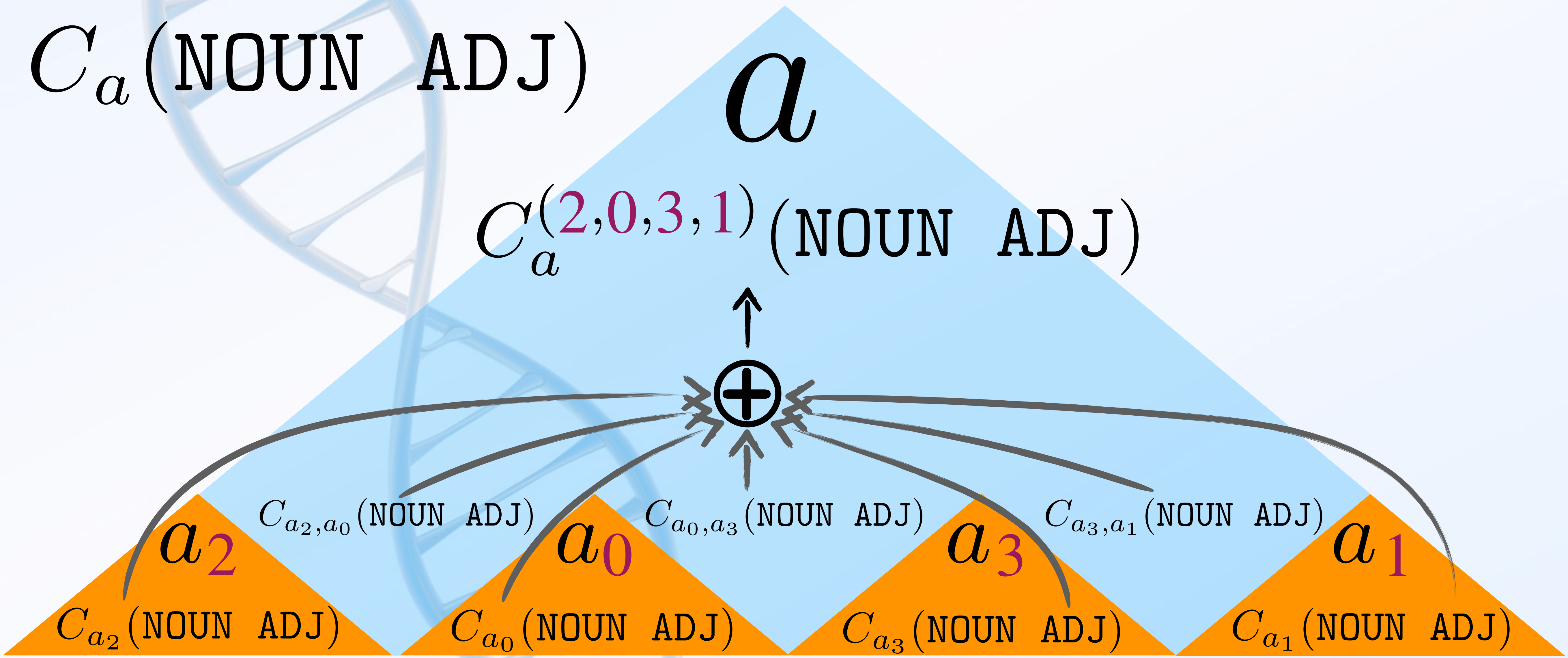


Computing Expected Counts
by Dynamic Programming



Computing Expected Counts
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

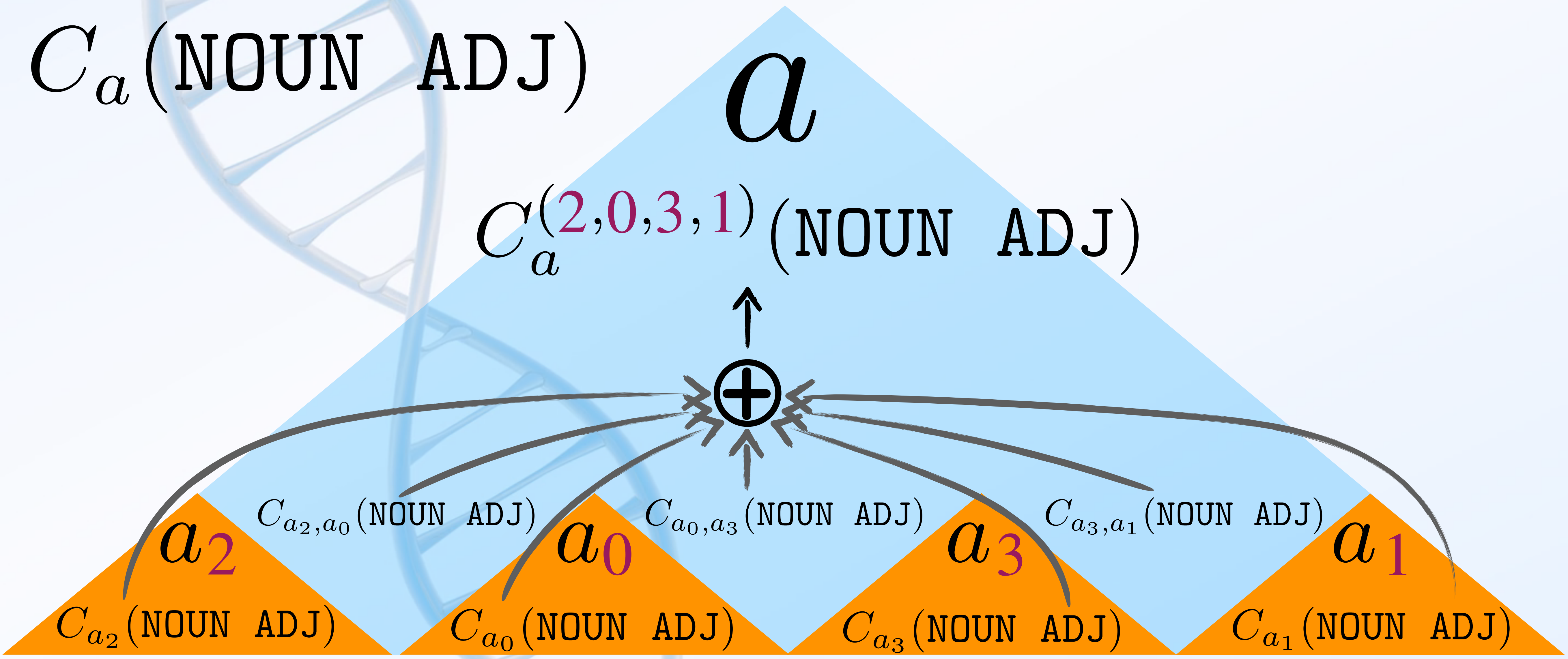


JOHNS HOPKINS UNIVERSITY

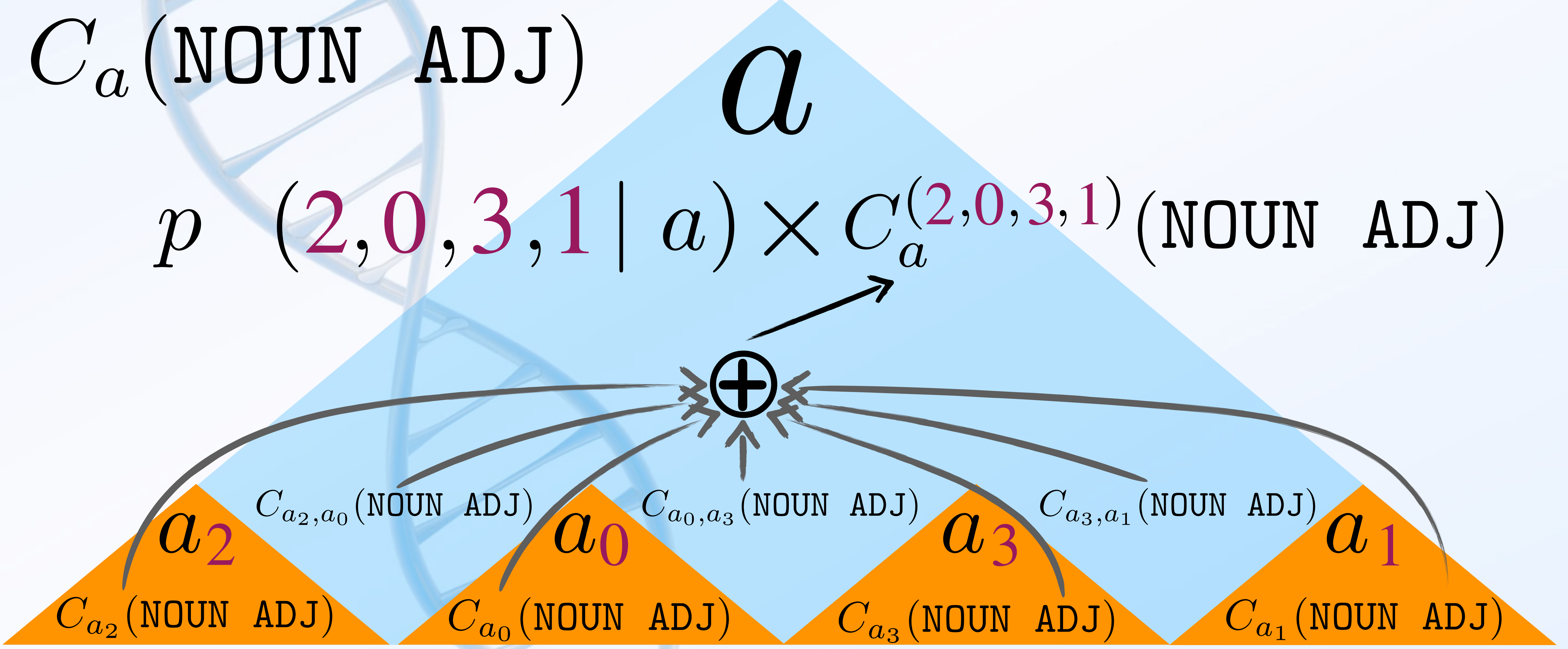
Computing Expected Counts

by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$



Computing Expected Counts
by Dynamic Programming

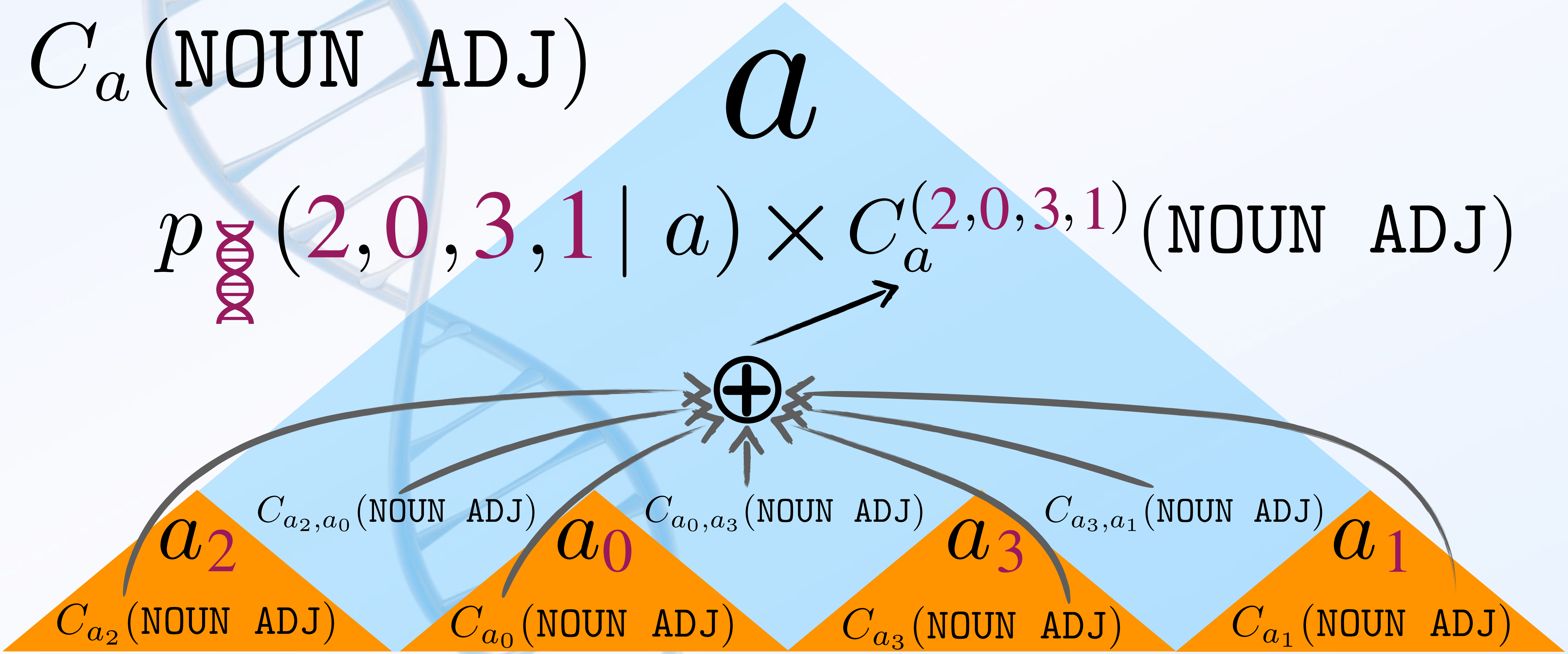


JOHNS HOPKINS UNIVERSITY

Computing Expected Counts
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

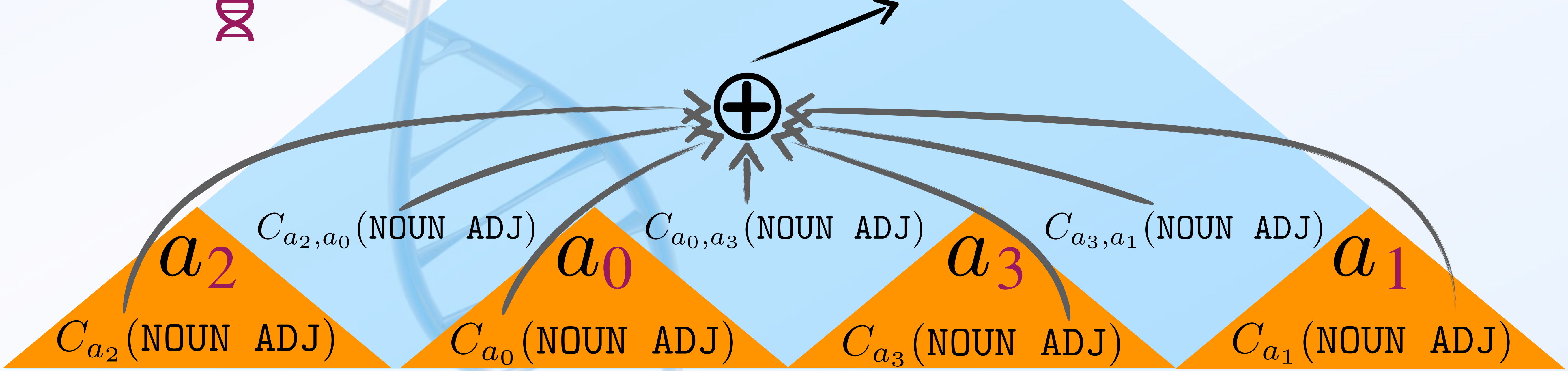
$$p(2,0,3,1 \mid a) \times C_a^{(2,0,3,1)}(\texttt{NOUN ADJ})$$



Computing Expected Counts

by Dynamic Programming

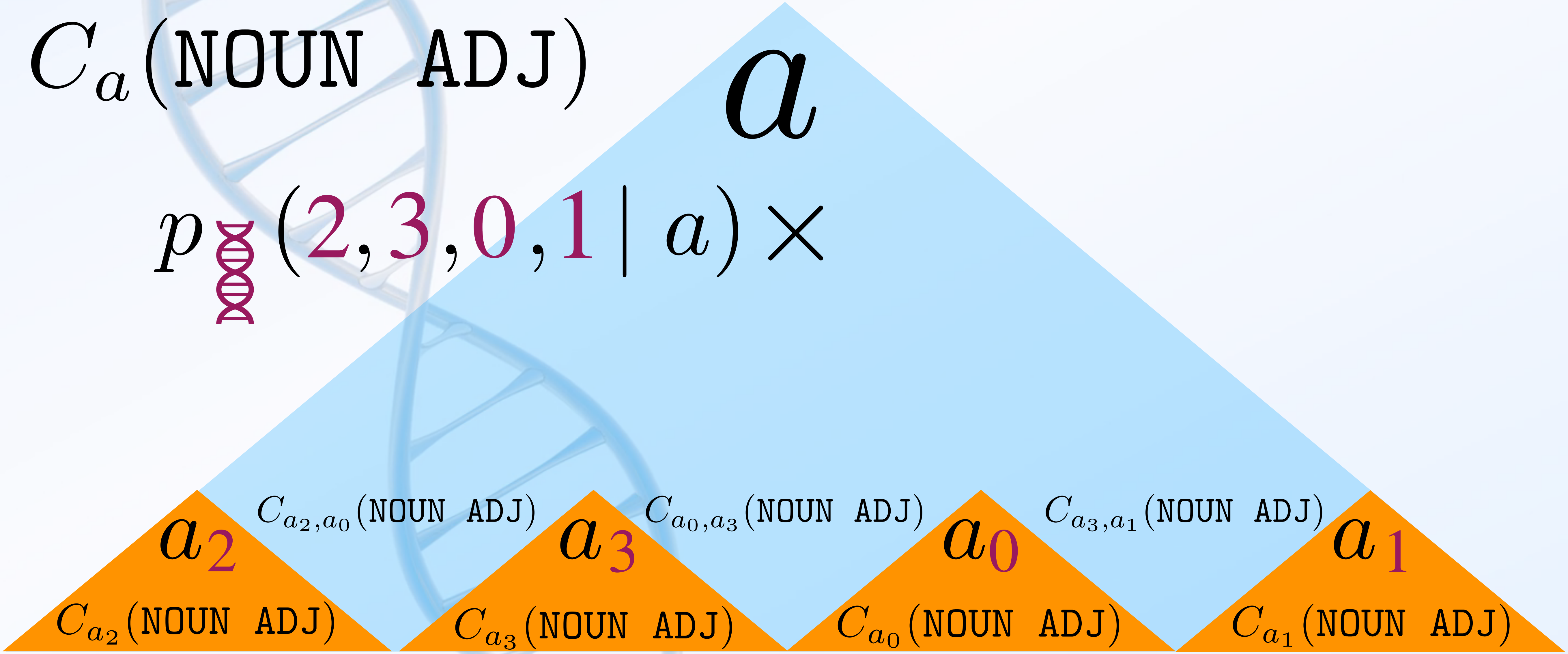
$$C_a(\texttt{NOUN ADJ}) \mathrel{+}= p(2,0,3,1 \mid a) \times C_a^{(2,0,3,1)}(\texttt{NOUN ADJ})$$



Computing Expected Counts
by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

$p_{🧬}(2, 3, 0, 1 \mid a) \times$

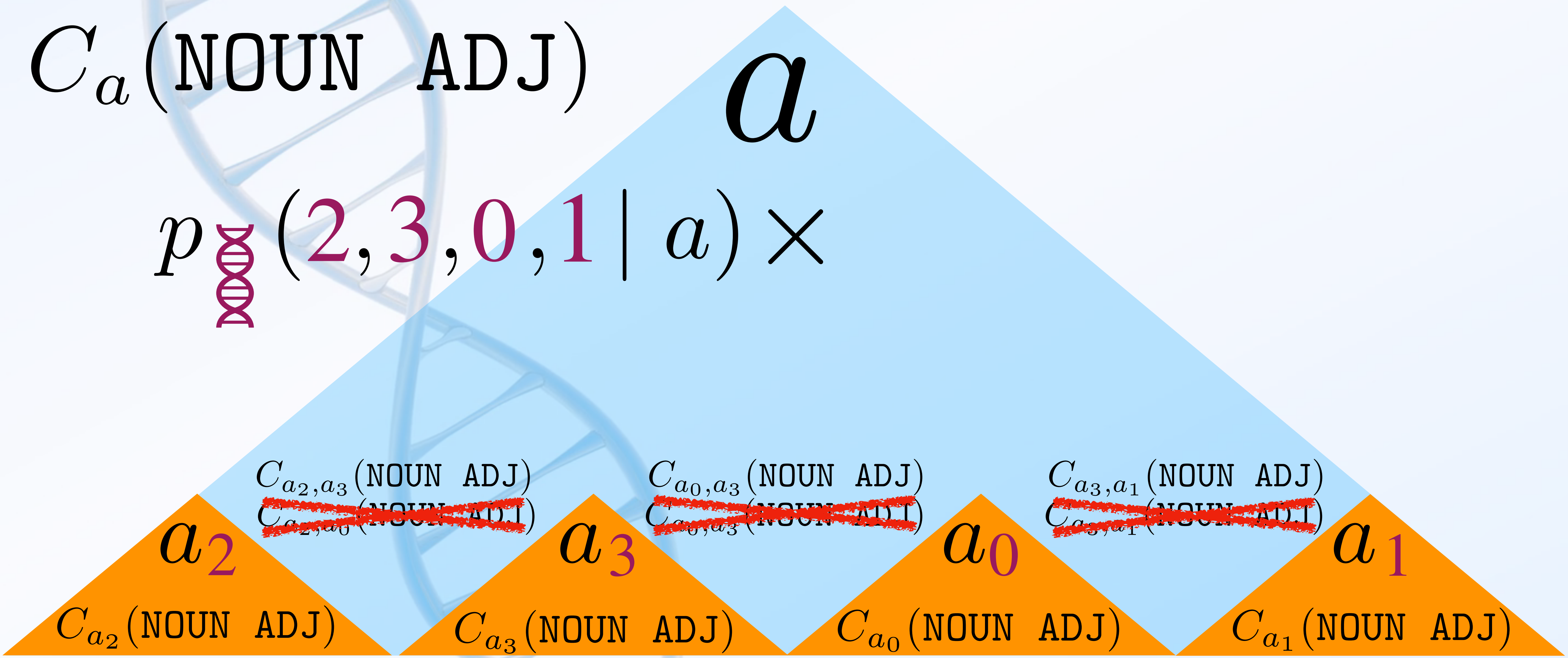


Computing Expected Counts

by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$

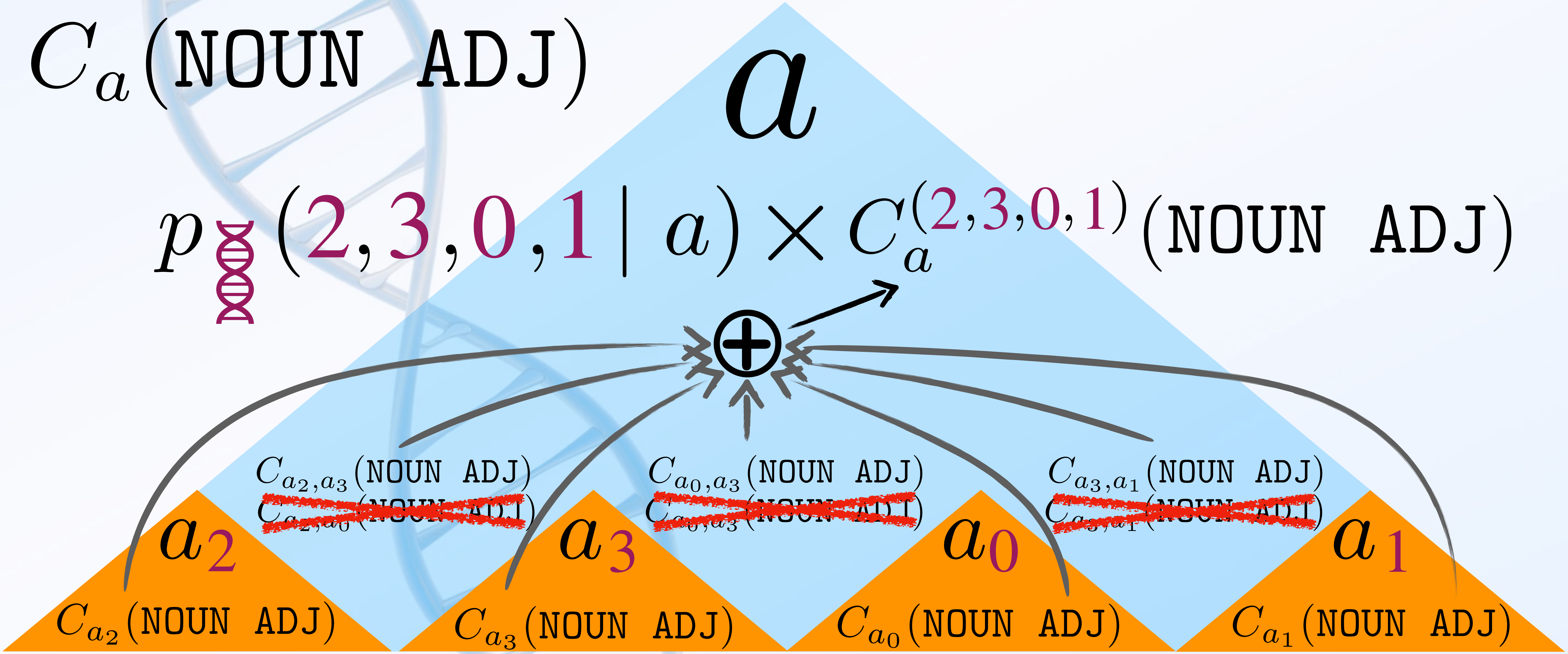
$$p(2,3,0,1 \mid a) \times$$



Computing Expected Counts

by Dynamic Programming

$C_a(\texttt{NOUN ADJ})$

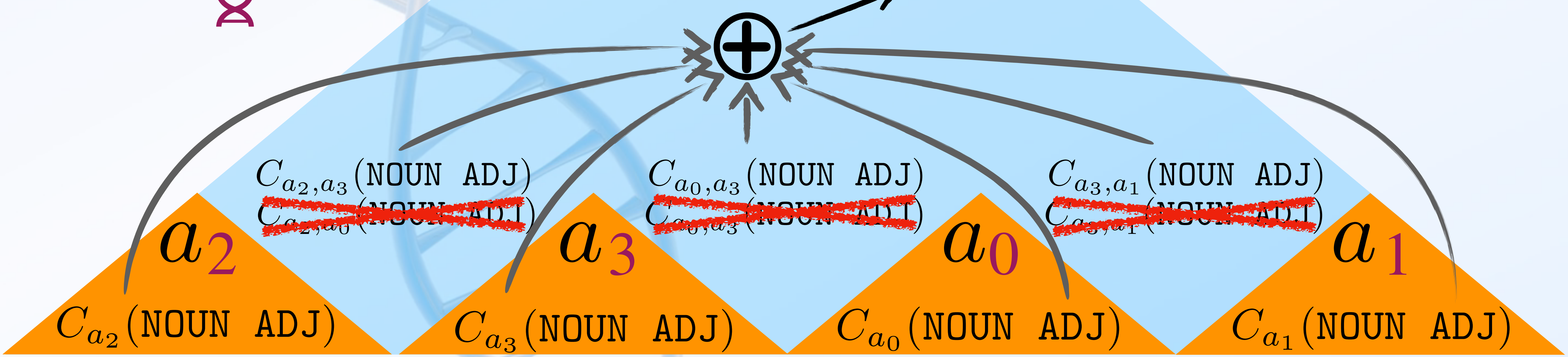


JOHNS HOPKINS UNIVERSITY

Computing Expected Counts

by Dynamic Programming

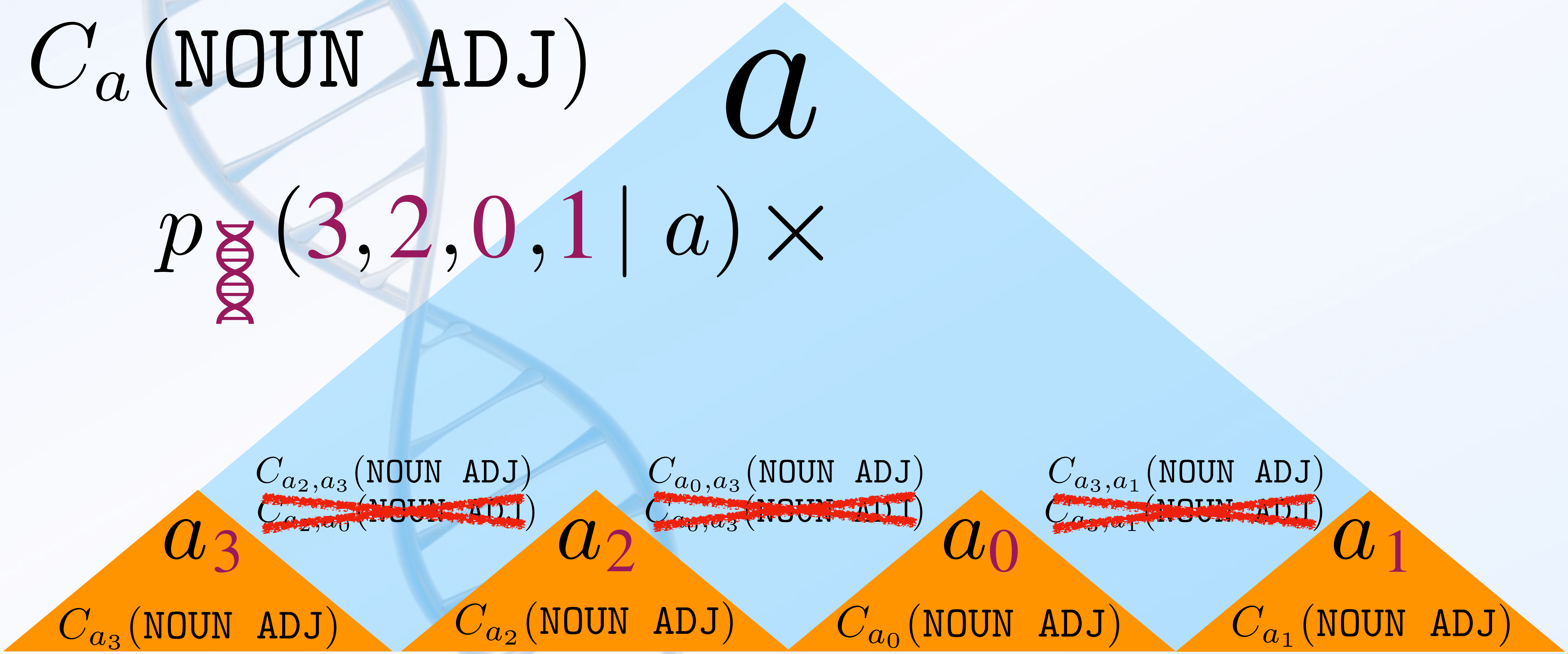
$$C_a(\texttt{NOUN ADJ})$$
$$\mathrel{+}= p(2,3,0,1 \mid a) \times C_a^{(2,3,0,1)}(\texttt{NOUN ADJ})$$



Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ})$$

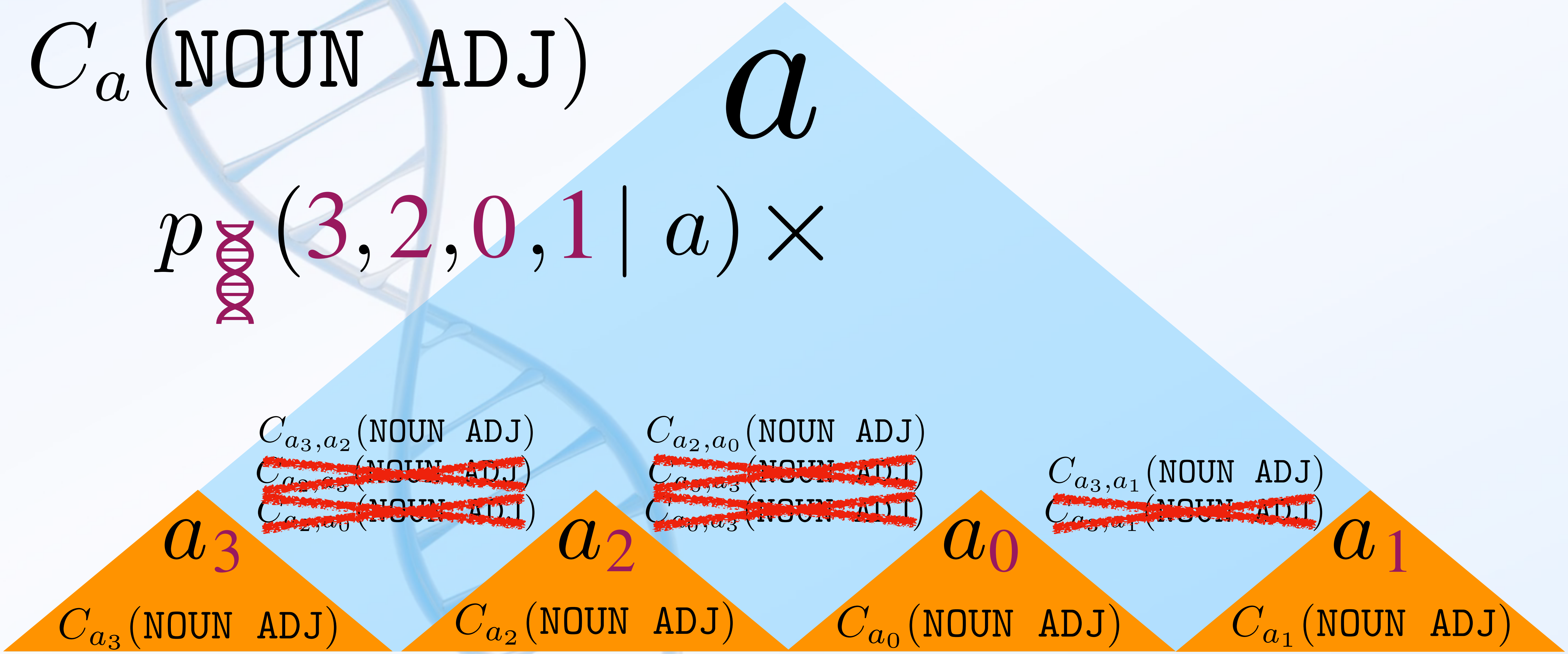
$$p_{🧬}(3, 2, 0, 1 \mid a) \times$$



Computing Expected Counts
by Dynamic Programming

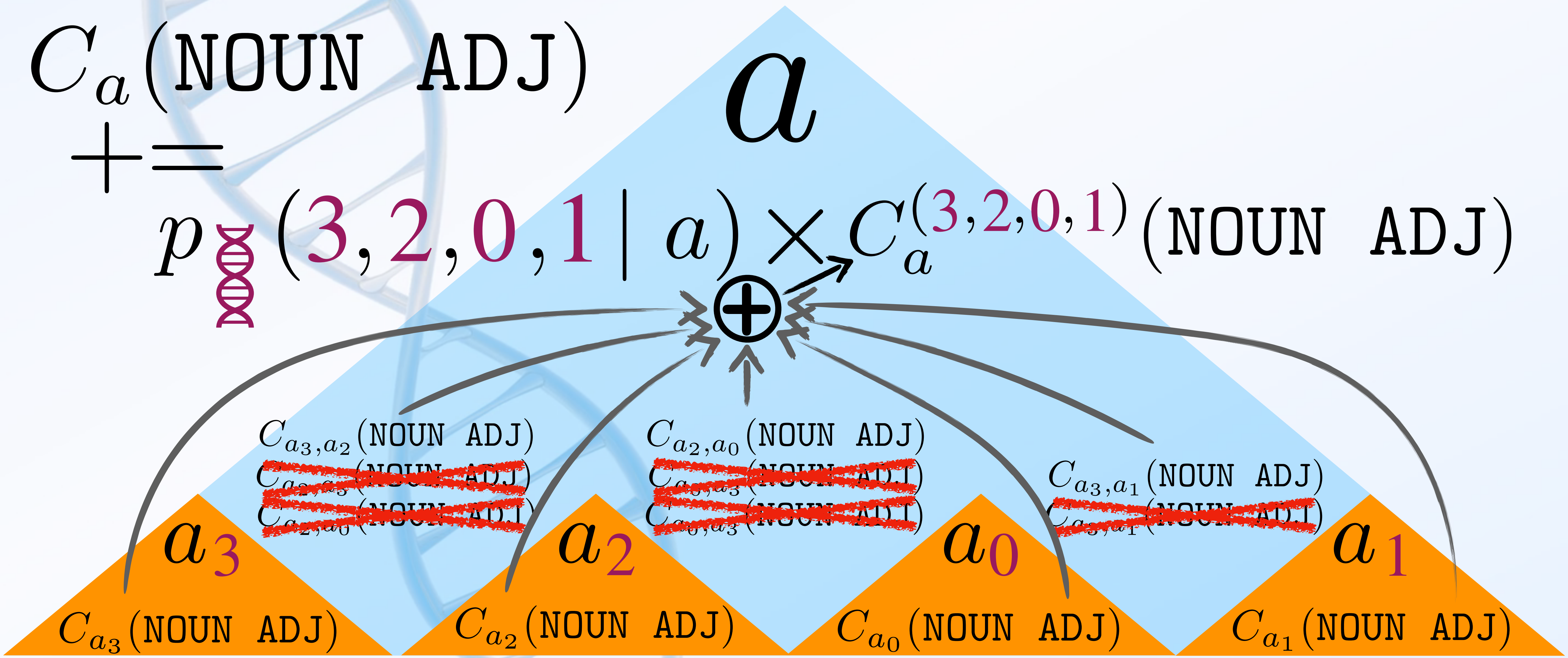
$C_a(\texttt{NOUN ADJ})$

$p(3,2,0,1 \mid a) \times$



Computing Expected Counts
by Dynamic Programming

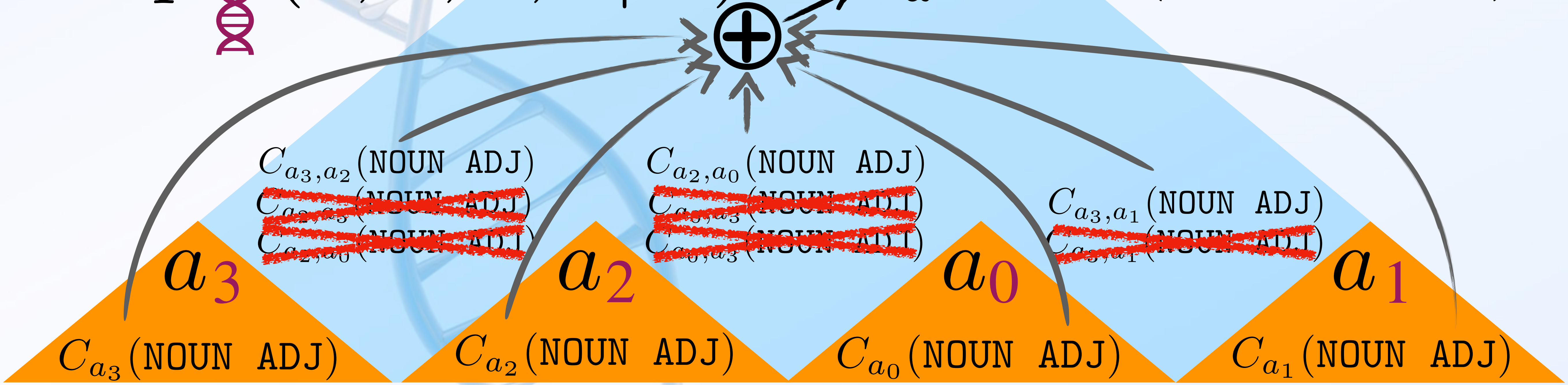
$$C_a(\texttt{NOUN ADJ}) \mathrel{+}= p(3,2,0,1 \mid a) \times C_a^{(3,2,0,1)}(\texttt{NOUN ADJ})$$



JOHNS HOPKINS UNIVERSITY

Computing Expected Counts
by Dynamic Programming

$$C_a(\texttt{NOUN ADJ}) \mathrel{+}= p(3,2,0,1 \mid a) \times C_a^{(3,2,0,1)}(\texttt{NOUN ADJ})$$



4! Permutations

Data

Johns Hopkins University

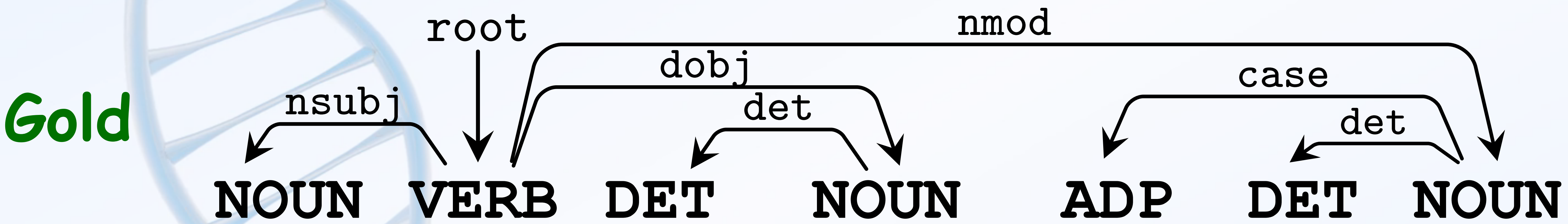
Data

- Universal Dependencies version 1.2
  - A collection of 37 dependency treebanks for 33 languages.

Train	Test
cs, es, fr, hi, de, it, la_itt, no, ar, pt, en, nl, da, fi, got, grc, et, la_proiel, grc_proiel, bg	la, hr, ga, he, hu, fa, ta, cu, el, ro, sl, ja_ktc, sv, fi_ftb, id, eu, pl

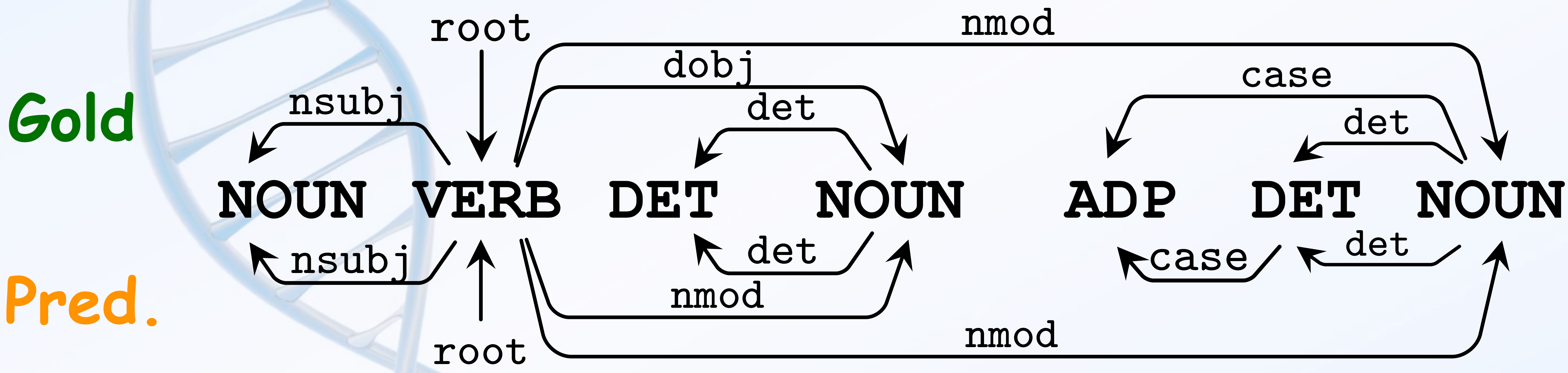
How to evaluate parses?

How to evaluate parses?



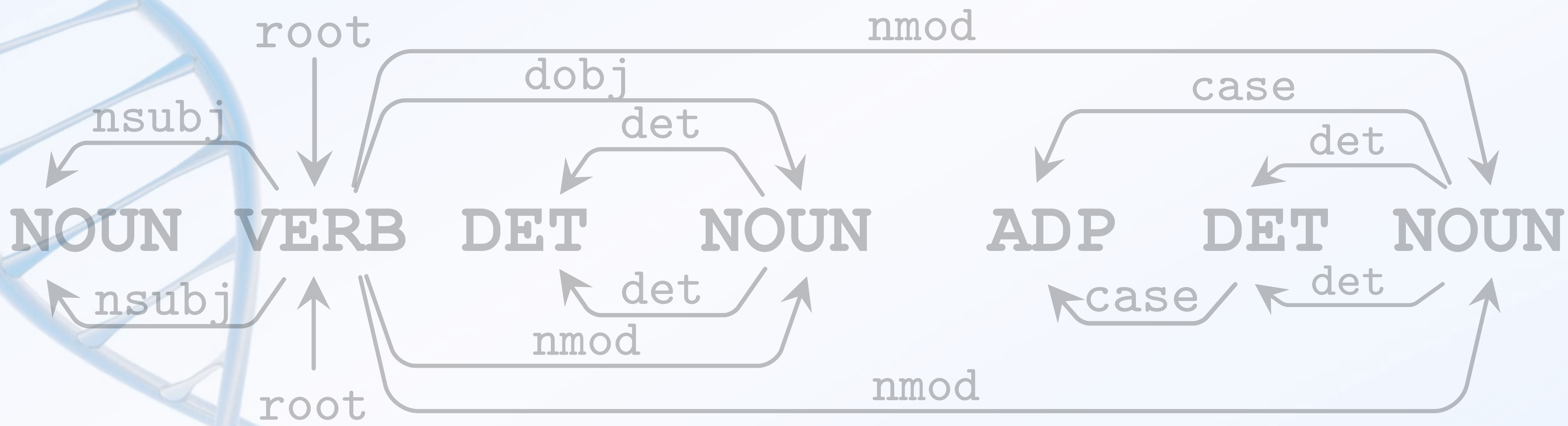
JOHNS HOPKINS UNIVERSITY

How to evaluate parses?



How to evaluate parses?

Gold



Pred.

UAS

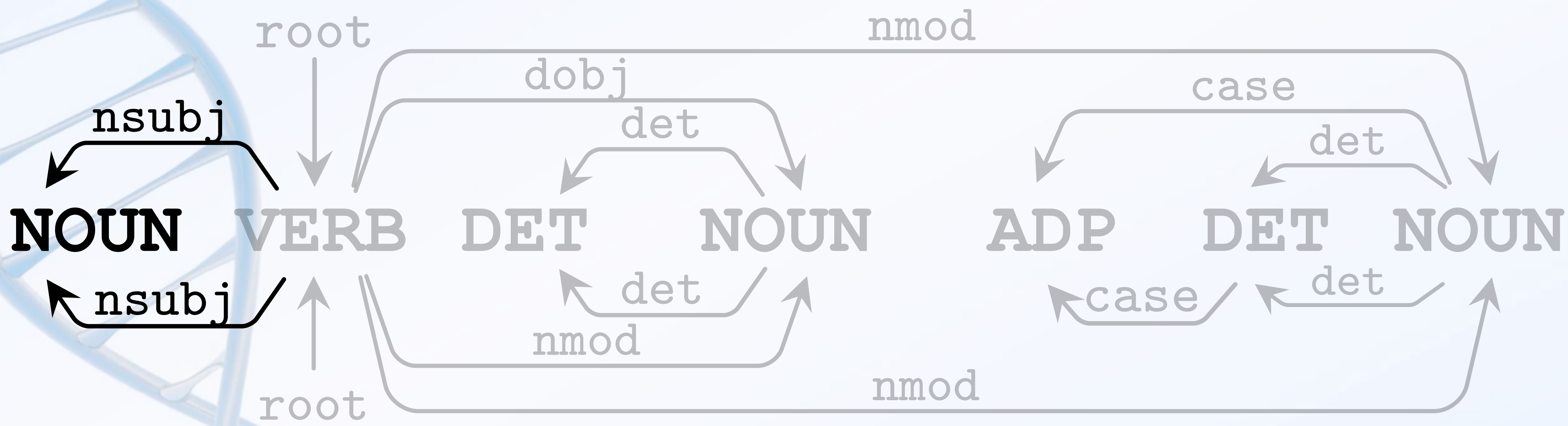
LAS

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

How to evaluate parses?

Gold

Pred.



UAS

LAS

UAS: Unlabelled Attachment Score

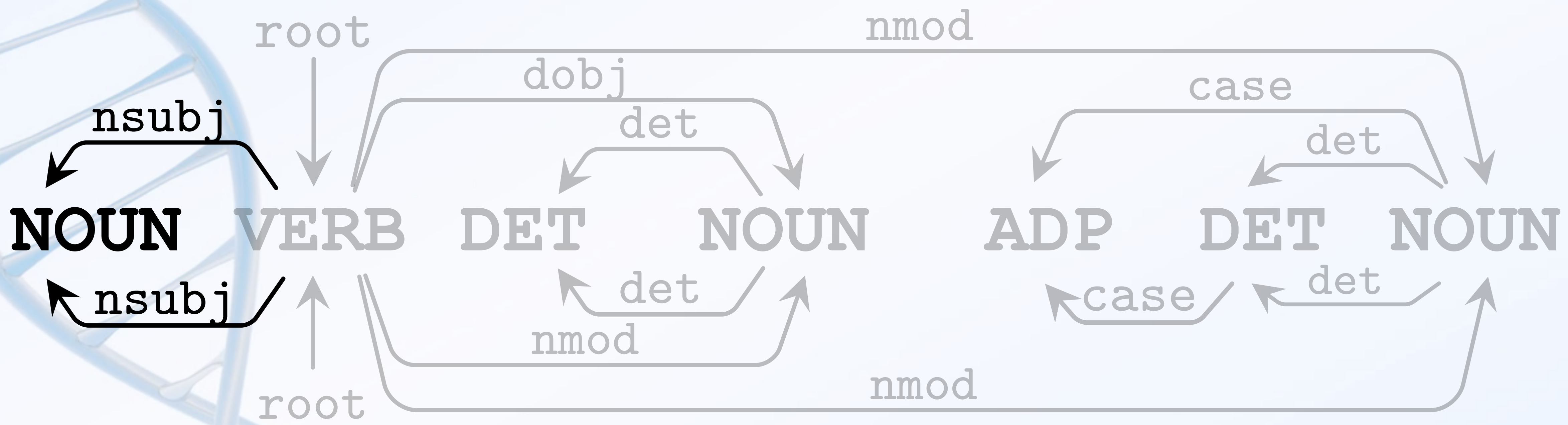
LAS: Labelled Attachment Score

JOHNS HOPKINS UNIVERSITY

How to evaluate parses?

Gold

Pred.



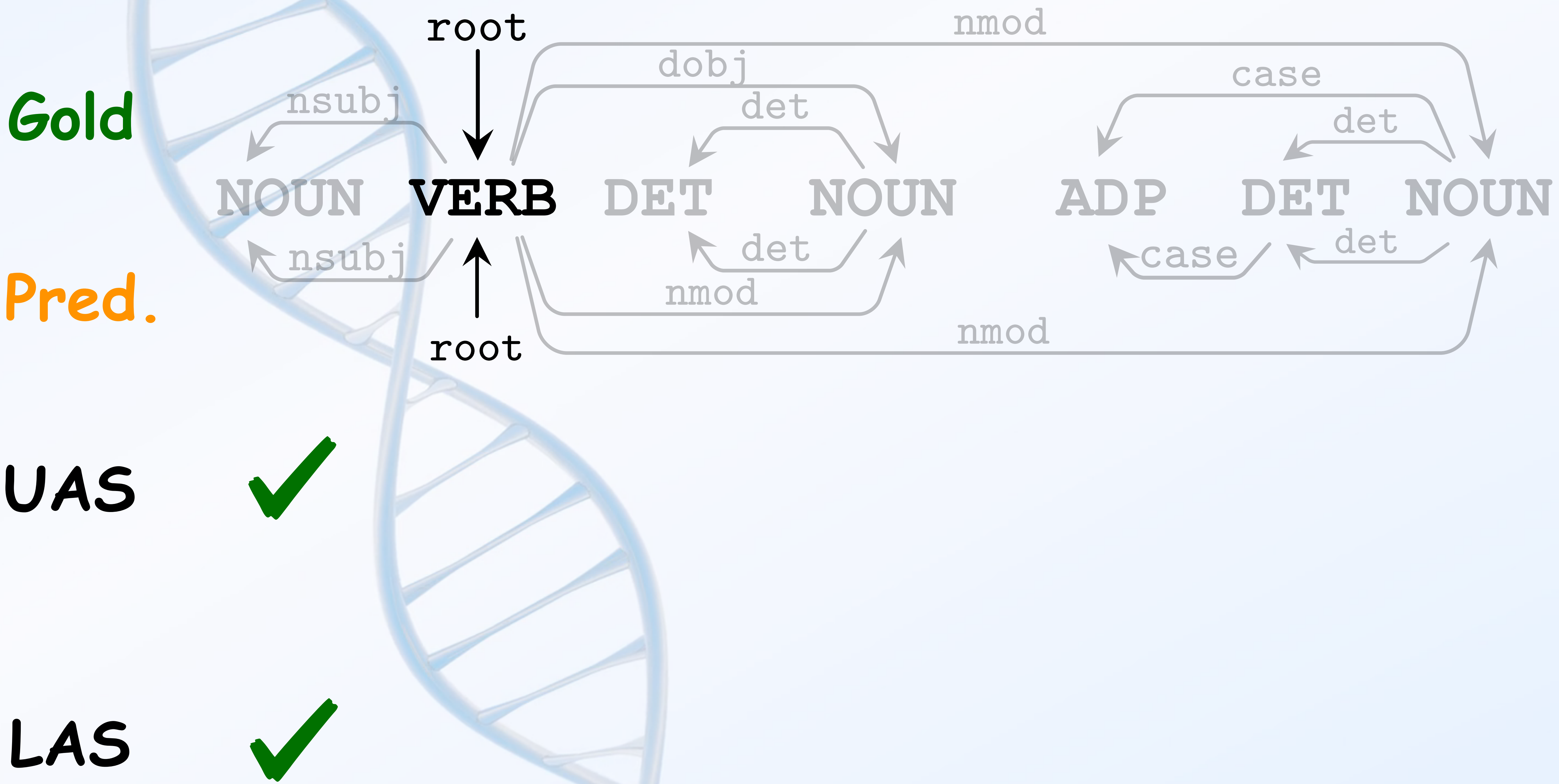
UAS ✓

LAS ✓

UAS: Unlabelled Attachment Score

LAS: Labelled Attachment Score

How to evaluate parses?

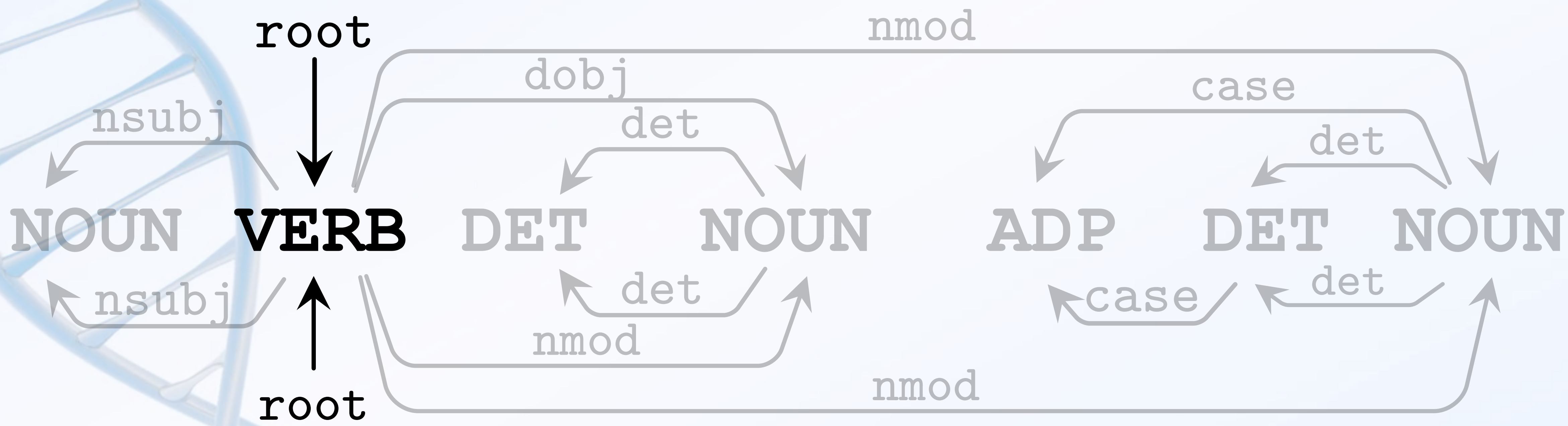


UAS: Unlabelled Attachment Score

LAS: Labelled Attachment Score

How to evaluate parses?

Gold



Pred.

UAS

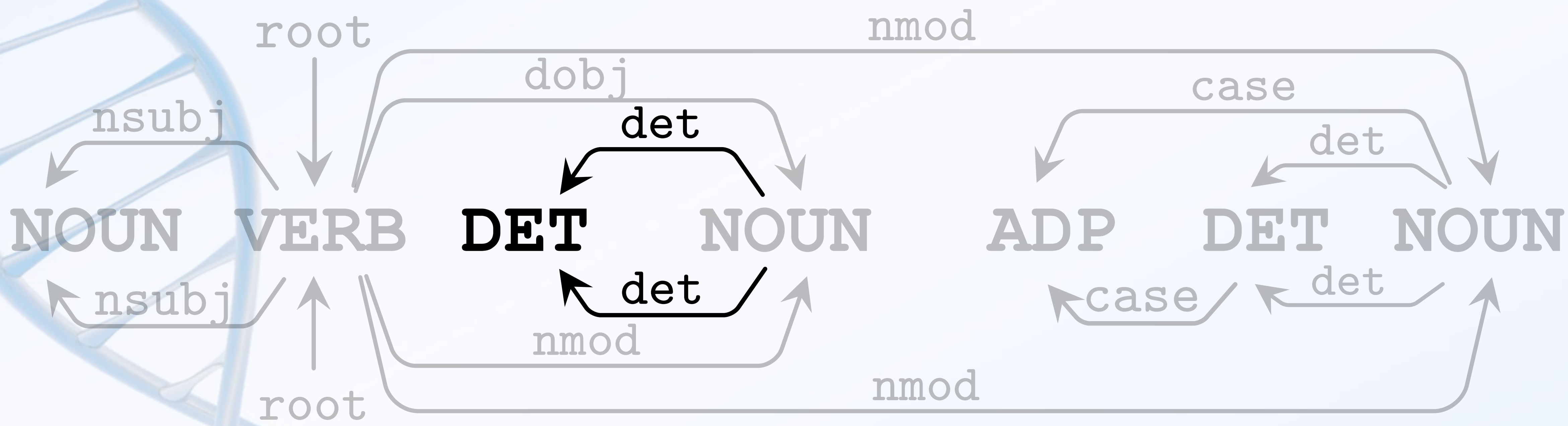
LAS

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

How to evaluate parses?

Gold

Pred.

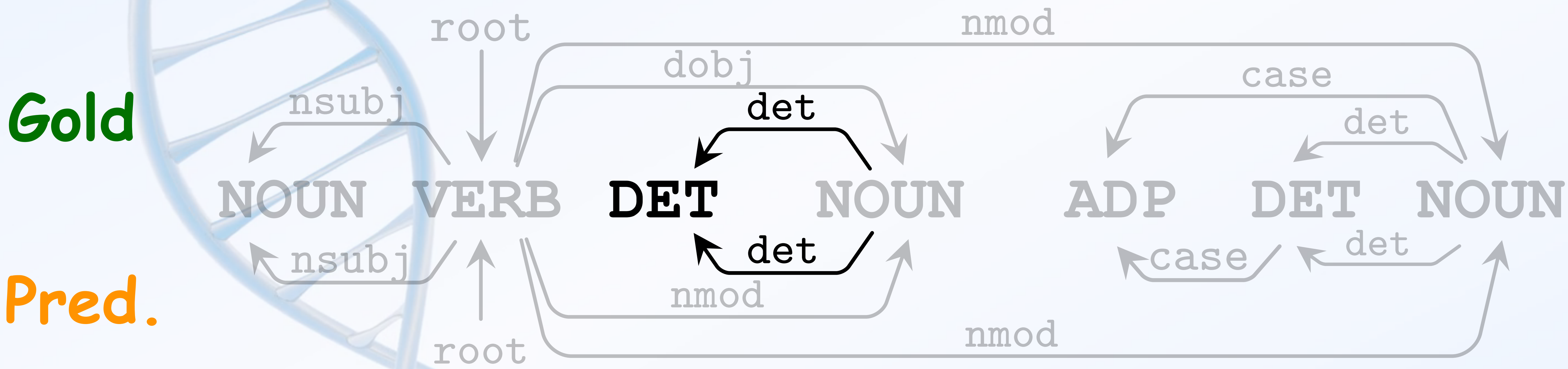


UAS ✓ ✓

LAS ✓ ✓

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

How to evaluate parses?

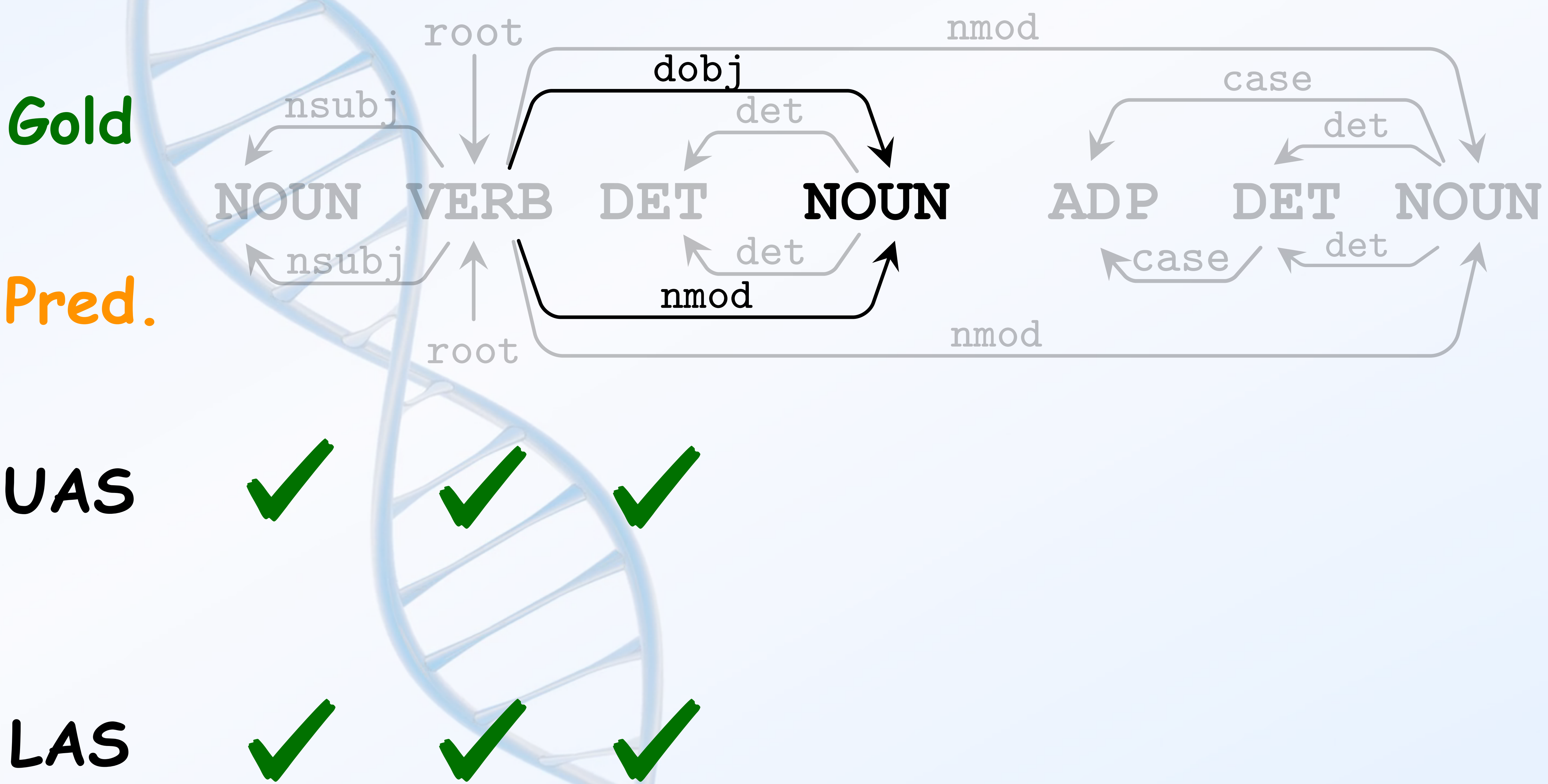


UAS ✓ ✓ ✓

LAS ✓ ✓ ✓

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

How to evaluate parses?

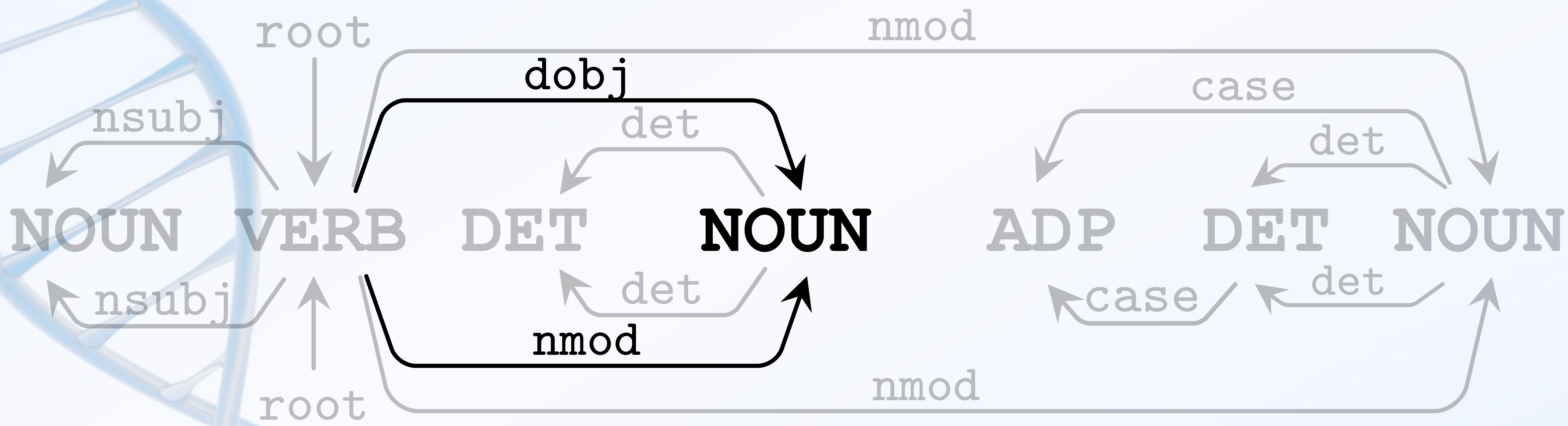


UAS: Unlabelled Attachment Score

LAS: Labelled Attachment Score

How to evaluate parses?

Gold



Pred.

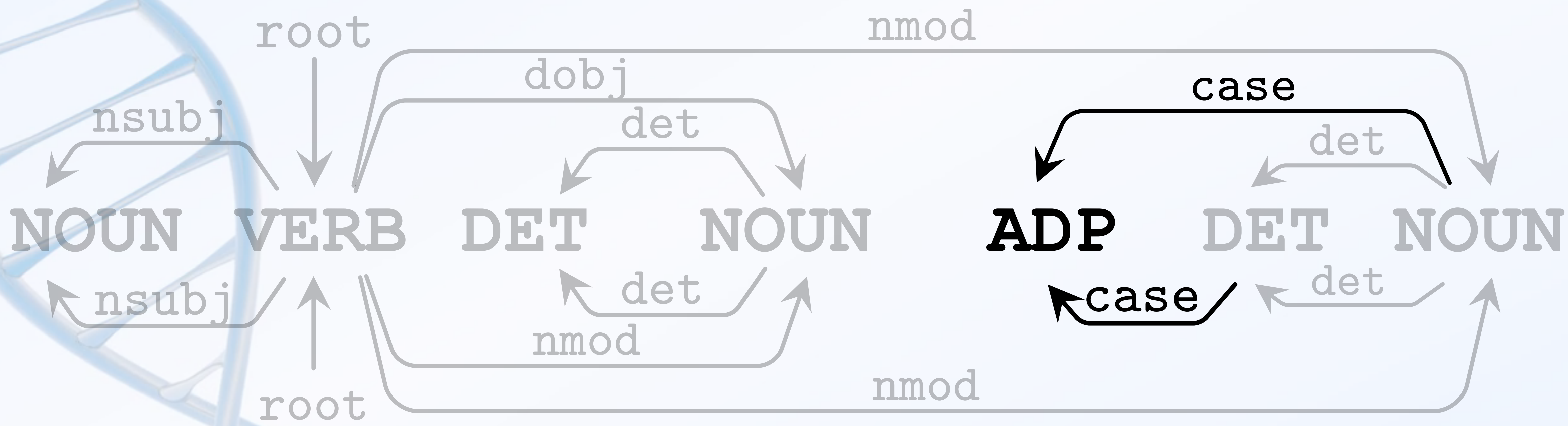
UAS	✓	✓	✓	✓
LAS	✓	✓	✓	✗

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

JOHNS HOPKINS UNIVERSITY

How to evaluate parses?

Gold



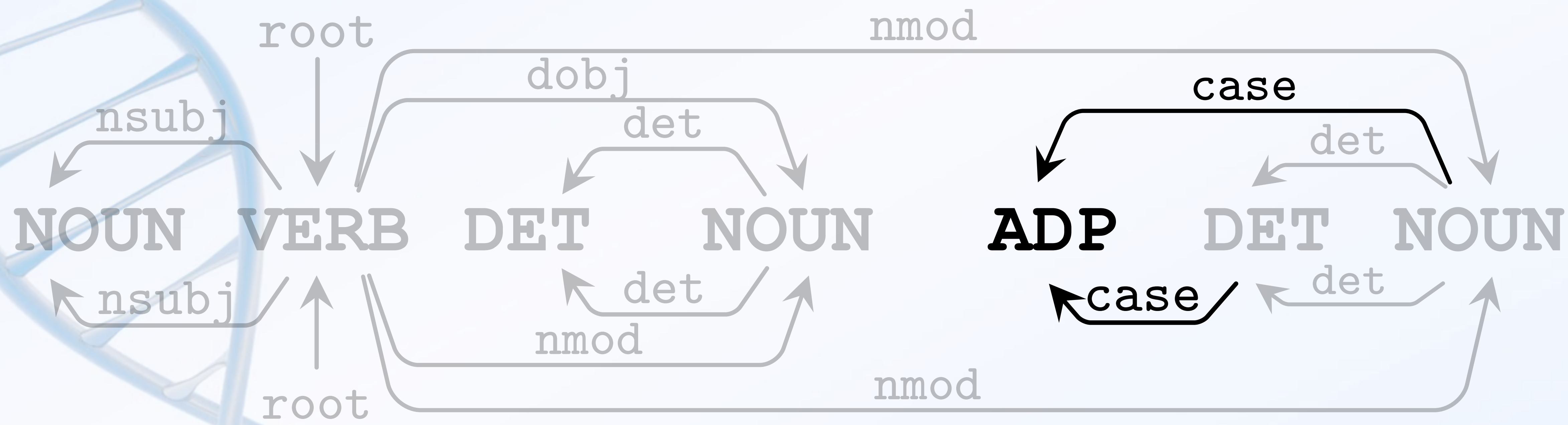
Pred.

UAS	✓	✓	✓	✓
LAS	✓	✓	✓	✗

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

How to evaluate parses?

Gold

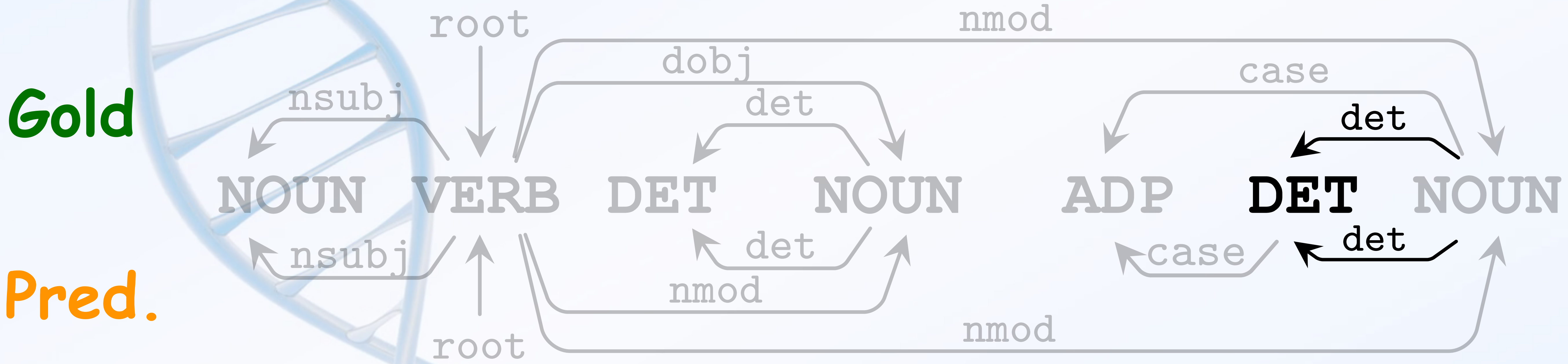


Pred.

UAS	✓	✓	✓	✓	✗
LAS	✓	✓	✓	✗	✗

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

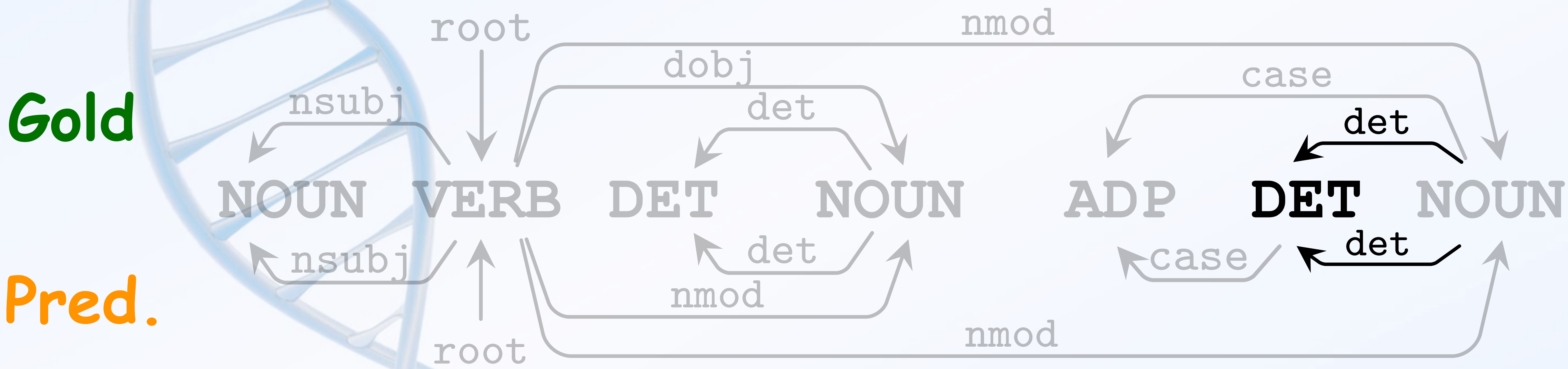
How to evaluate parses?



UAS	✓	✓	✓	✓	✗
LAS	✓	✓	✓	✗	✗

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

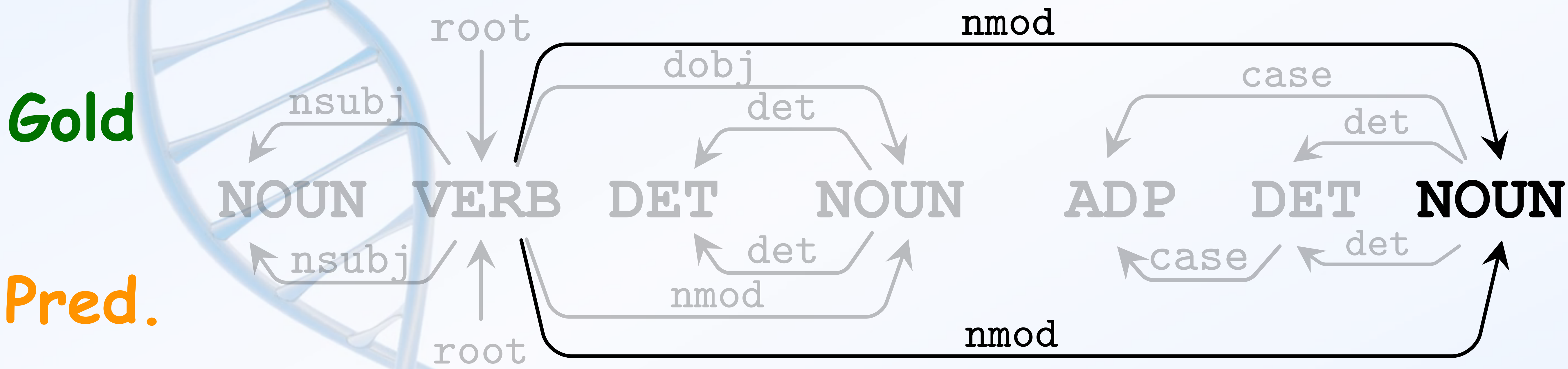
How to evaluate parses?



UAS	✓	✓	✓	✓	✗	✓	
LAS	✓	✓	✓	✗	✗	✓	

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

How to evaluate parses?

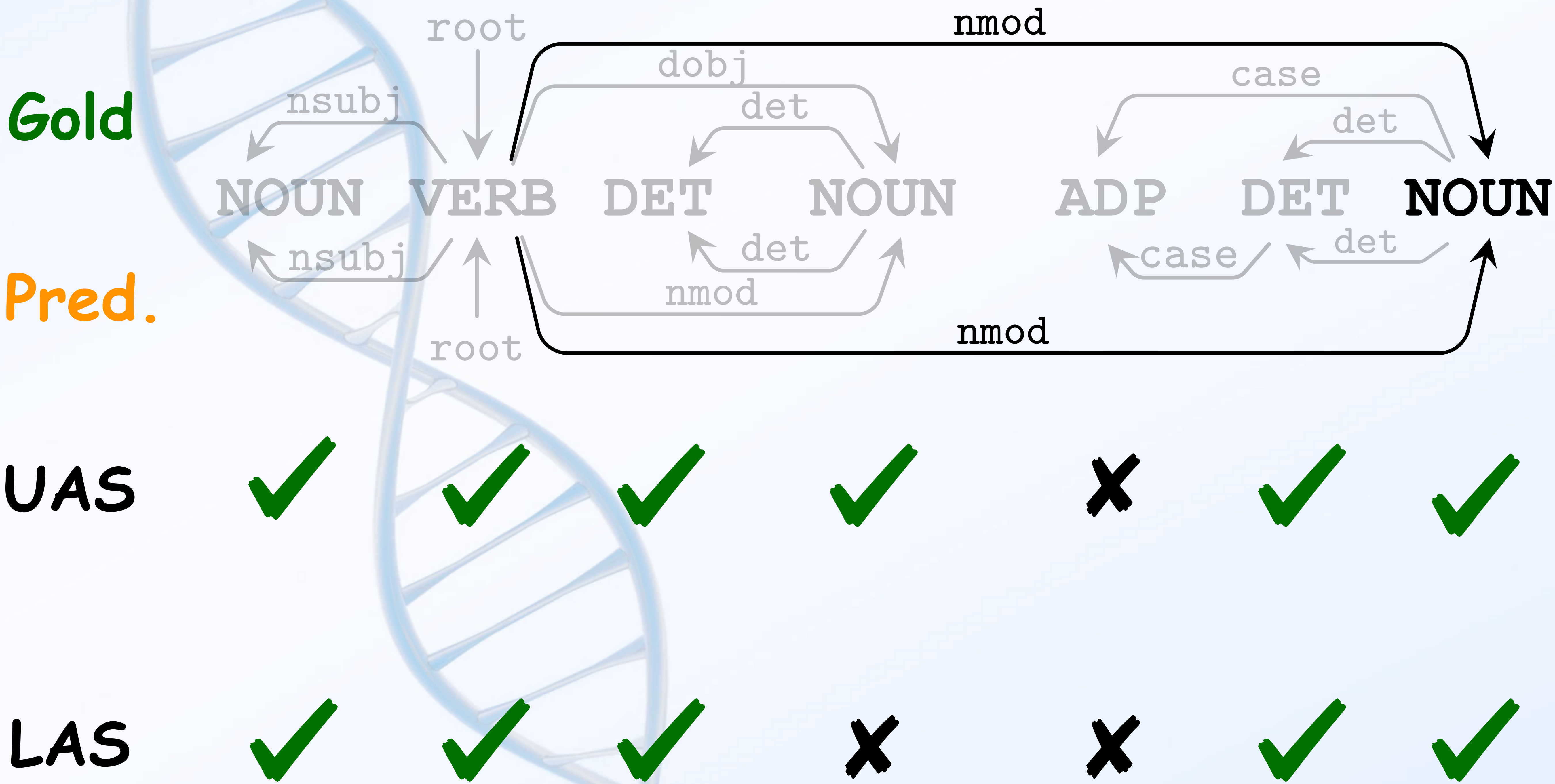


UAS	✓	✓	✓	✓	✗	✓
LAS	✓	✓	✓	✗	✗	✓

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

JOHNS HOPKINS UNIVERSITY

How to evaluate parses?

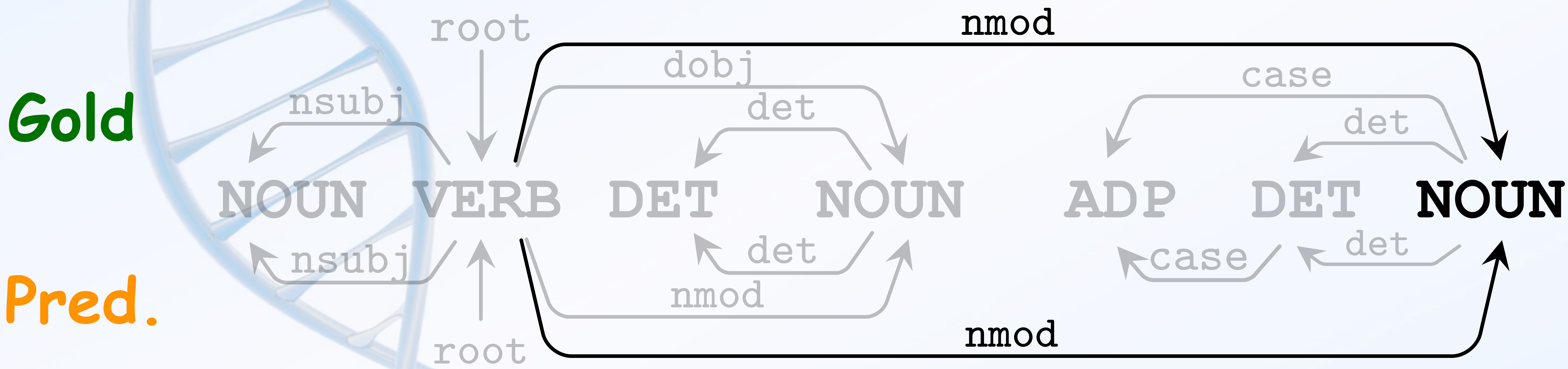


UAS: Unlabelled Attachment Score

LAS: Labelled Attachment Score

JOHNS HOPKINS
UNIVERSITY

How to evaluate parses?



UAS	$\frac{6}{7}$	✓	✓	✓	✓	✗	✓	✓
LAS	$\frac{5}{7}$	✓	✓	✓	✗	✗	✓	✓

UAS: Unlabelled Attachment Score
LAS: Labelled Attachment Score

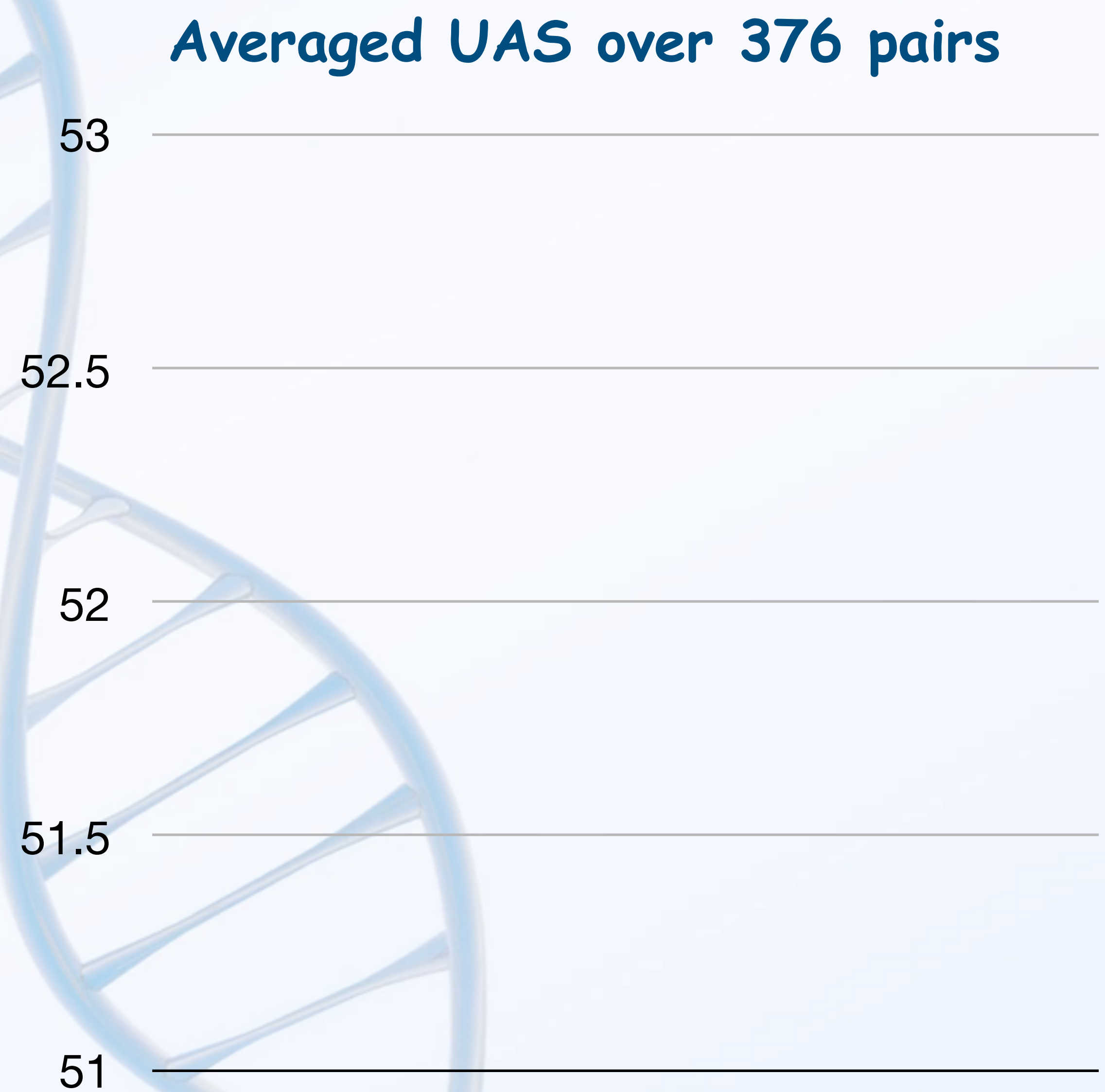
Johns Hopkins University

Does our method work?

Does our method work?

Overall YES!

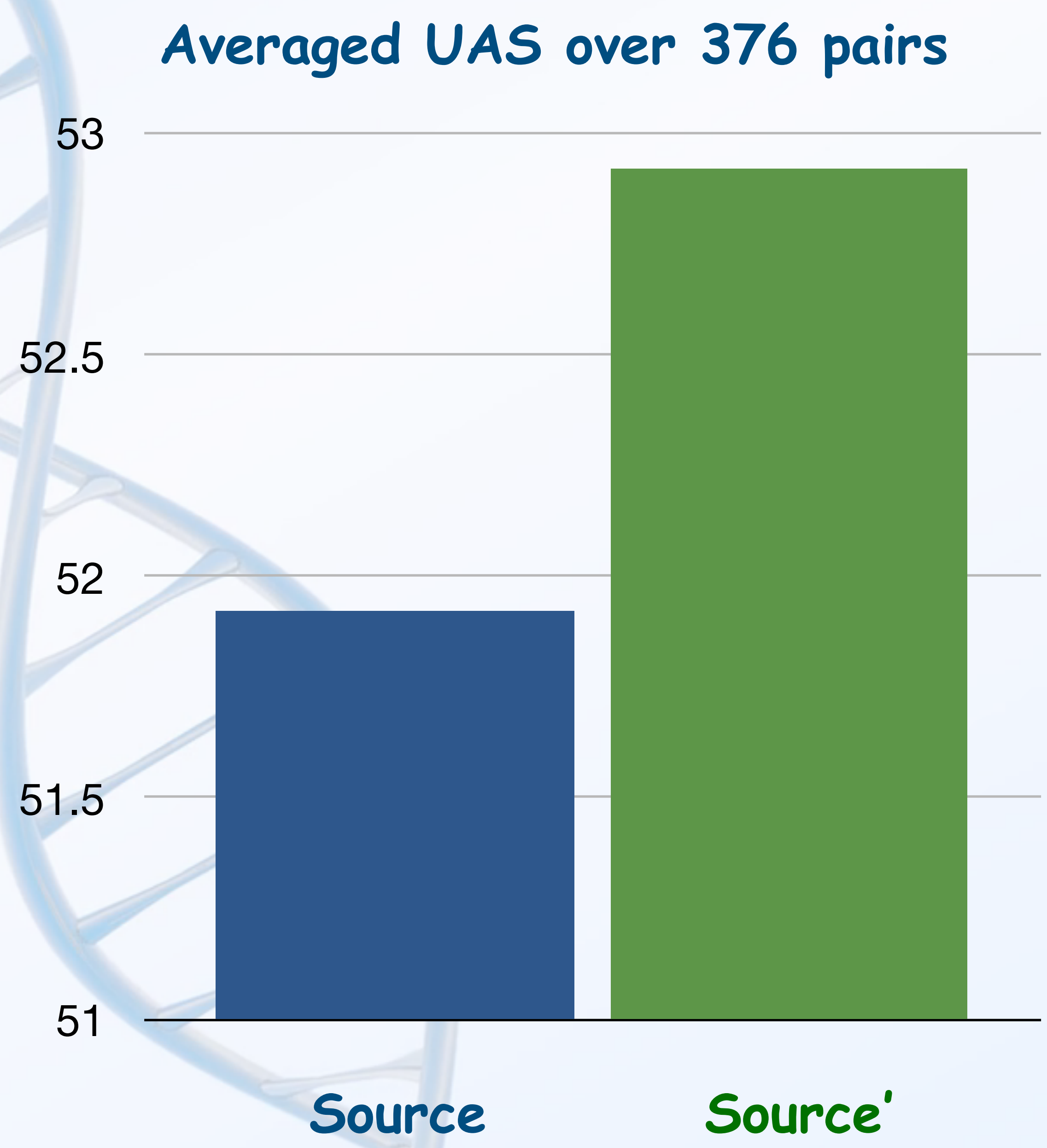
Does our method work?



Overall YES!

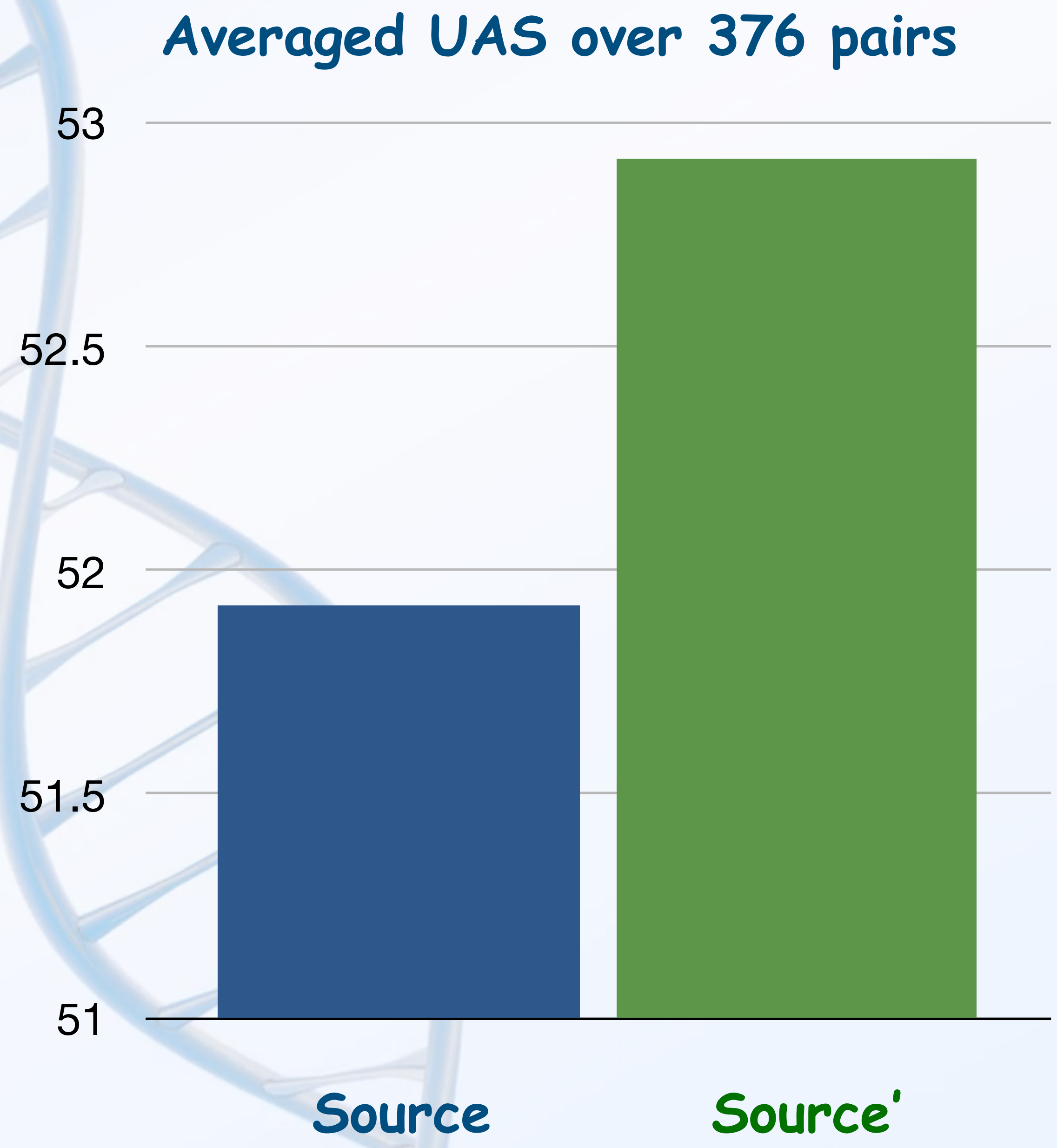
JOHNS HOPKINS UNIVERSITY

Does our method work?



Overall YES!

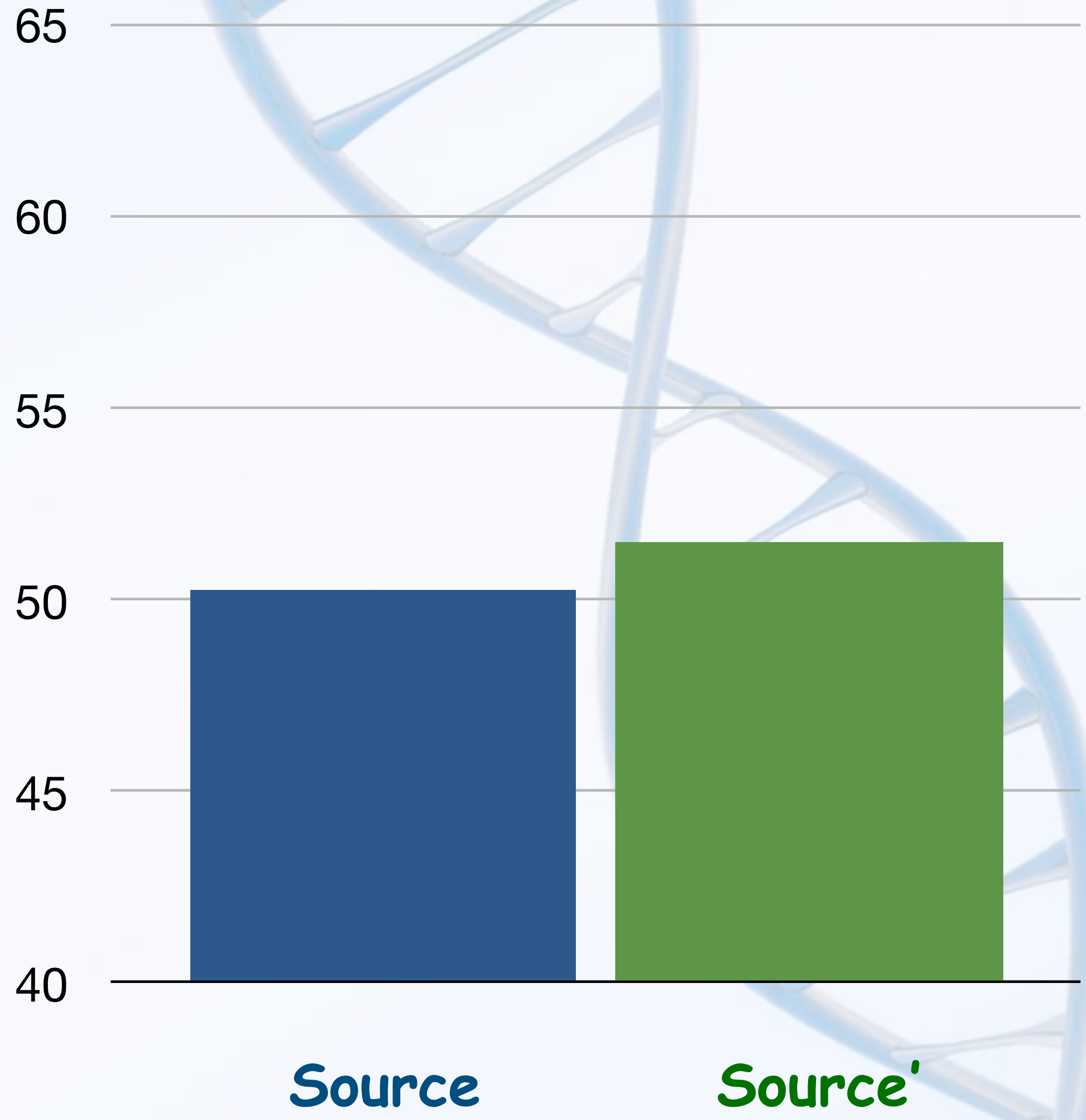
Depends on the Language Family



Depends on the Language Family

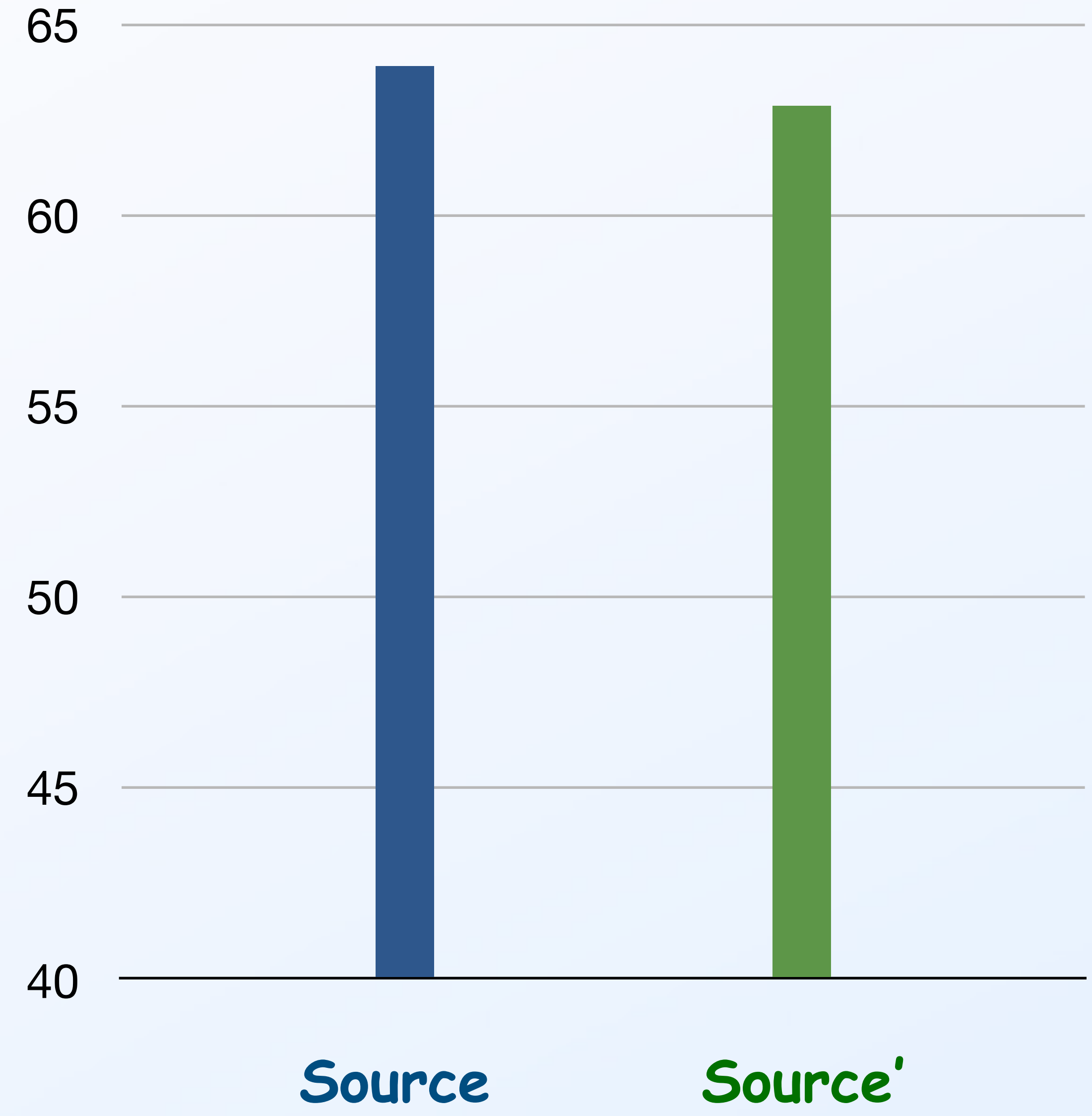
Different family

Averaged UAS over 330 pairs



Same family

Averaged UAS over 46 pairs



JOHNS HOPKINS
UNIVERSITY

Better Parsing

Better Parsing

- Fancier surface similarity function
  - Recurrent neural network language models capture longer context
  - Dynamic programming methods are no longer available.
  - We need approximate inference (sampling)!

Better Parsing

- Fancier surface similarity function
  - Recurrent neural network language models capture longer context
  - Dynamic programming methods are no longer available.
  - We need approximate inference (sampling)!

- Relax the requirement of POS-tags
  - Unavailable for a low-resource target language
  - Richer lexical information would be better than POS-tags
  - Cross-lingual unsupervised word embeddings (Ruder et al., 2017) would be useful

Better Parsing

- Fancier surface similarity function
  - Recurrent neural network language models capture longer context
  - Dynamic programming methods are no longer available.
  - We need approximate inference (sampling)!

- Relax the requirement of POS-tags
  - Unavailable for a low-resource target language
  - Richer lexical information would be better than POS-tags
  - Cross-lingual unsupervised word embeddings (Ruder et al., 2017) would be useful

- More efficient inference!
  - Enumerating over n! permutations
  - We could approximately sample from permutations (Eisner and Tromble, 2006)

THANKS!