Course series: Deep Learning for Machine Translation

Introduction to Language and Translation

Lecture # 1

Hassan Sajjad and Fahim Dalvi
Qatar Computing Research Institute, HBKU

Language

A sequence of words that delivers a concept

Language properties

- Morphology
- Syntax
- Semantics

Language - Morphology

- Study of words
- How words are formed and their relationship with other words in a language

Language - Morphology

Play

Playing

Played

Morphology pertains to variations in words (“play”) to represent different states. “ing” and “ed” are suffixes that change what the root “play” conveys

Language - Morphology

- Morphologically poor languages
  - English, French
- Morphologically rich languages
  - Arabic, Hebrew, German

Arabic

وسنساعدهم wsnsAEdhm (and will we help them)

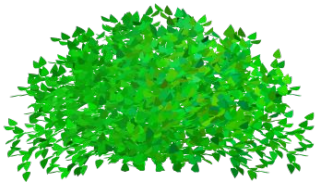
w(and) s(will) n(we) sAEd(help) hm(them)

Language - Semantics

- Study of the meaning of words and phrases in a language

Language - Semantics

Bush vs. bush



bank vs. bank

can vs. can

The same word can mean different things in different contexts, i.e. what other words surround it

Language - Syntax

ich habe einen Apfel gegessen

NOUN VERB

I ate an apple

VERB NOUN

Different languages have different ways of ordering components (objects, actions, etc.) in a sentence

Translation

Meaningful representation of one language in another language

English	He does not go home
Spanish	No va a su casa
Chinese	他不回家
German	Er geht ja nicht nach hause
Arabic	هو لا يذهب إلى البيت

Ingredients of an Automatic Translation System

- Parallel corpus
  - Pairs of sentences in two languages that represent the same meaning

Er geht ja nicht nach hause	He does not go home
Ich arbeite daran	I am working on it
⋮	⋮

Machine translation system learns from a large pool of parallel sentences

Ingredients of an Automatic Translation System

- Parallel corpus
  - Pairs of sentences in two languages that represent the same meaning

Source language	Target language
Er geht ja nicht nach hause	He does not go home
Ich arbeite daran	I am working on it

Machine translation system learns from a large pool of parallel sentences

Exercise

Find translation of words

Thanks to Fabienne Cap for sharing the exercise

What did we get?

- From a set of parallel sentences, we can
  - learn a dictionary
  - find ambiguous words
  - one to many and many to one translations
  - etc.

This is what an automatic translation system does!

Ingredients of an Automatic Translation System

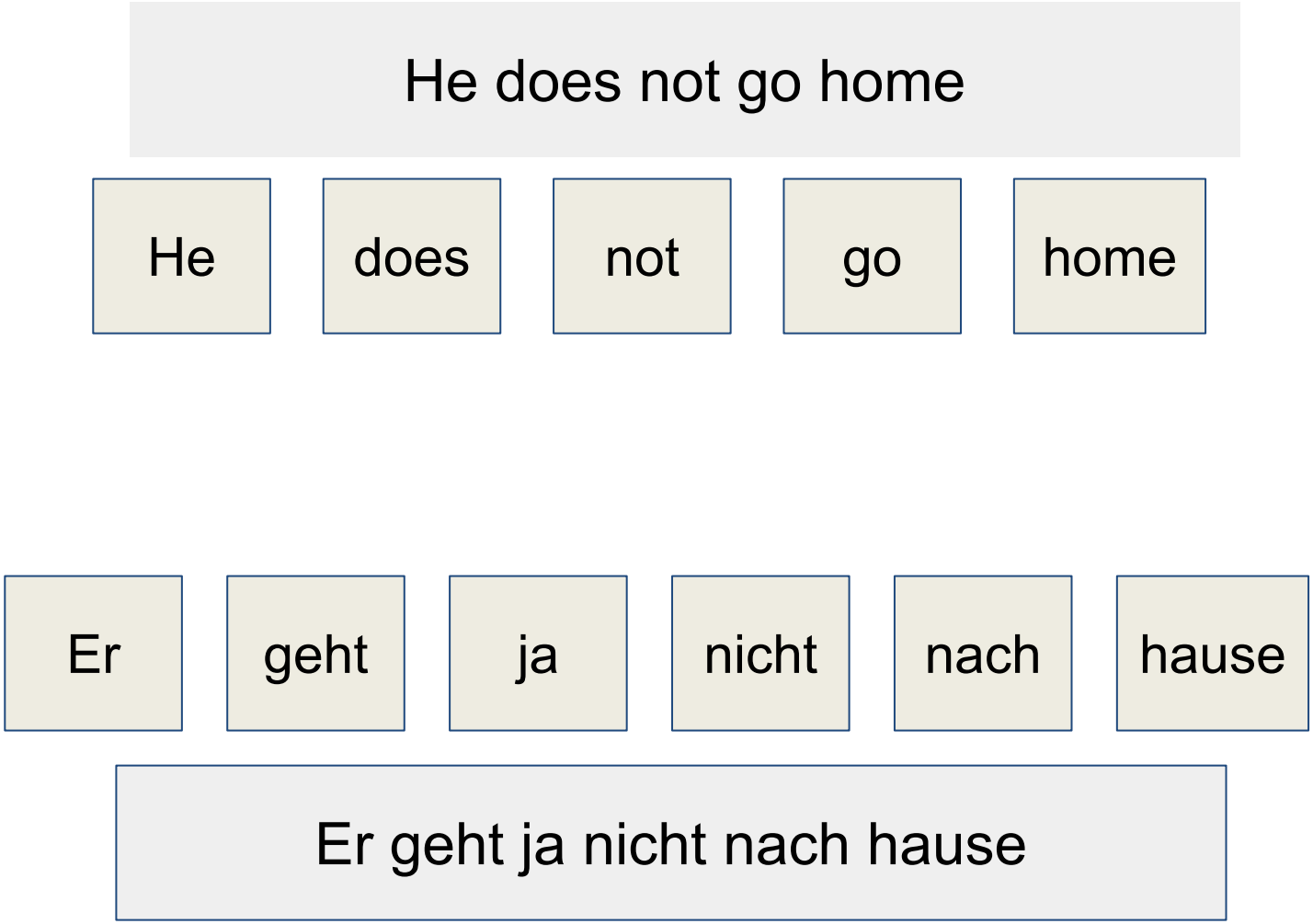
- Translation model
  - learn word level and phrase level translations
- Language model
  - fluency model
  - learn to generate fluent translations
- Decoder
  - translation generation component
  - produce a translation from a trained translation model and language model

Word Alignment

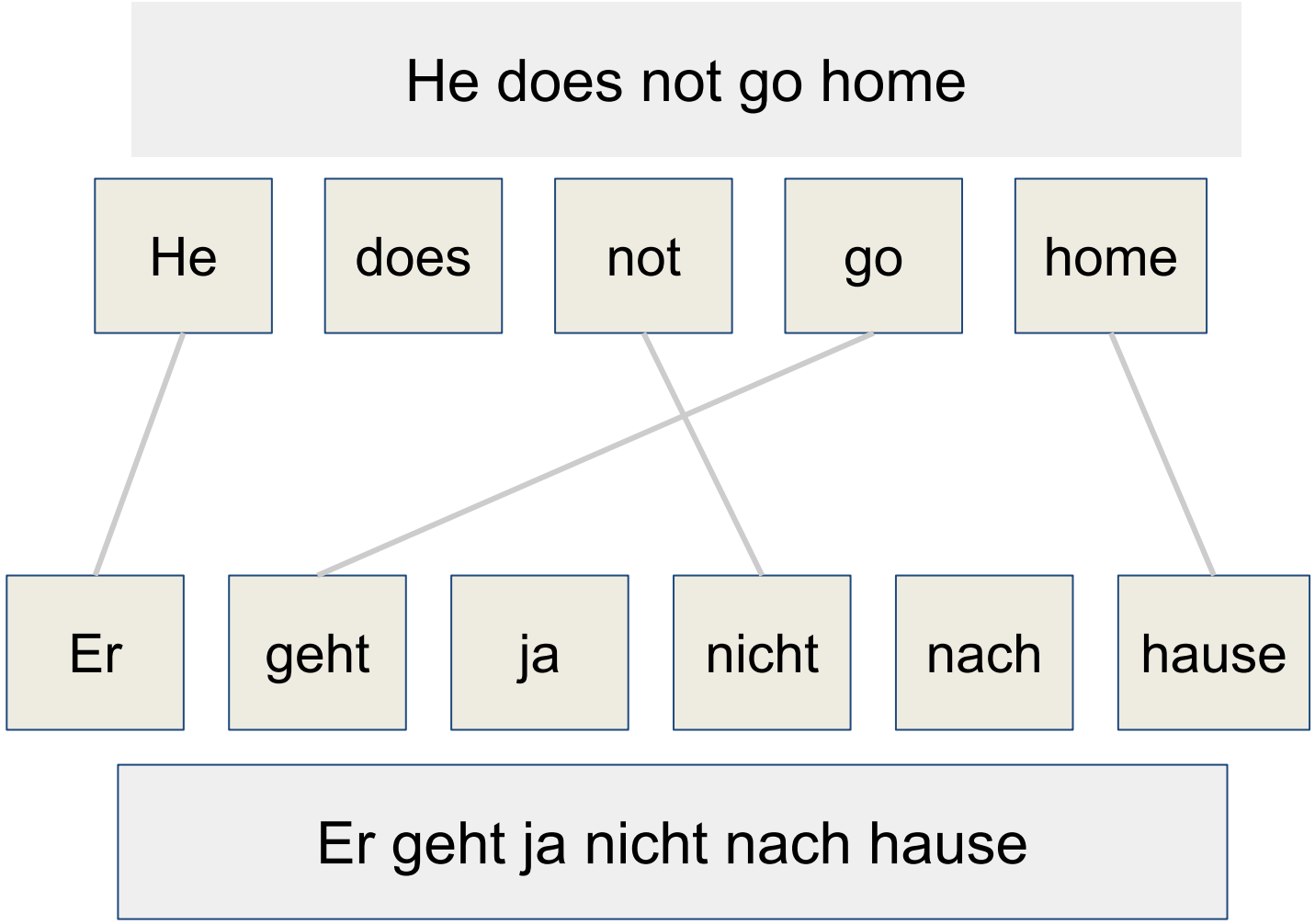
He does not go home

Er geht ja nicht nach hause

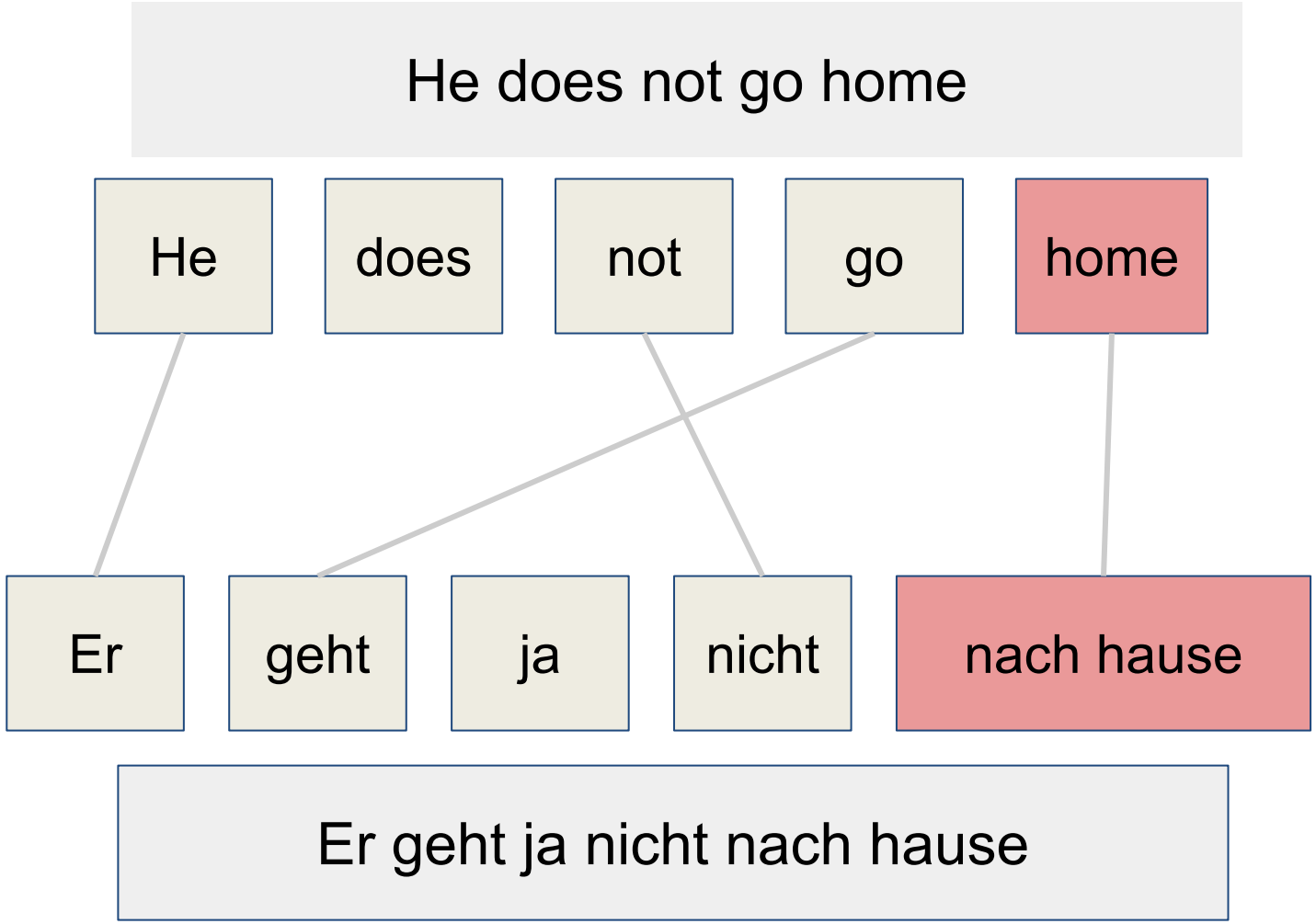
Word Alignment



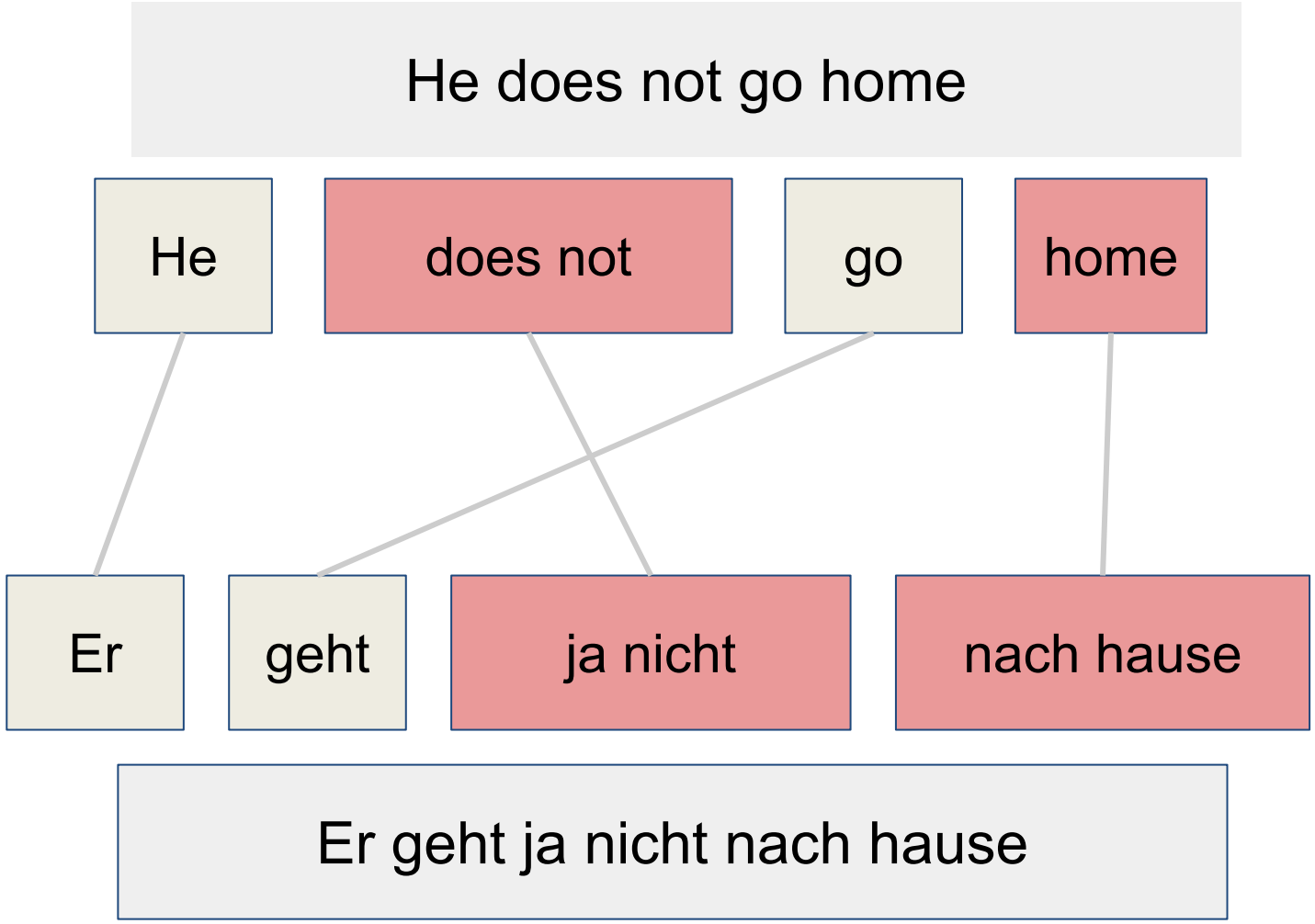
Word Alignment



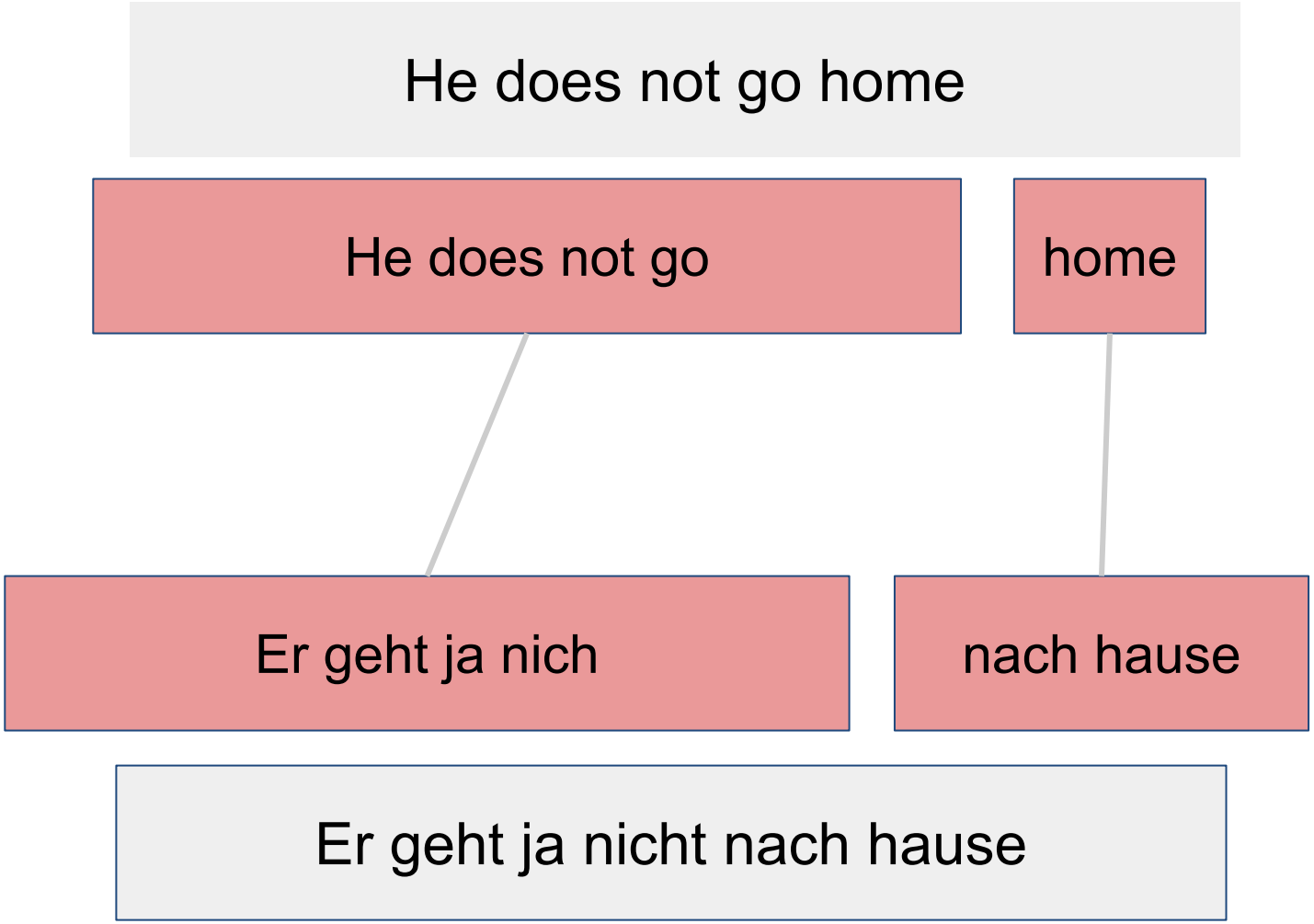
Phrase Alignment



Phrase Alignment



Phrase Alignment



Decoding

Process of generating translation of an input sentence based on a trained model

Let’s decode a few examples: (exercise)

Decoding

Process of generating translation of an input sentence based on a trained model

Let’s decode a few examples: (exercise)

Did you keep multiple options?

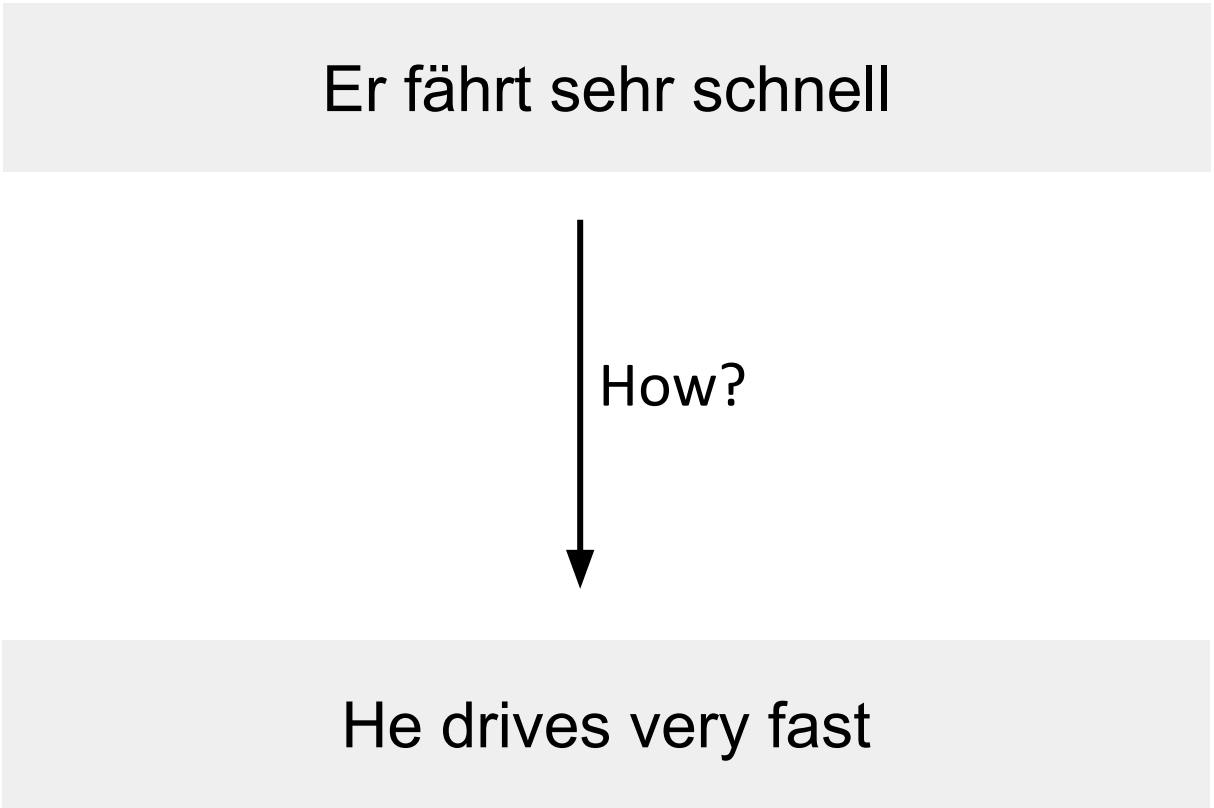
Decoding

Process of generating translation of an input sentence based on a trained model

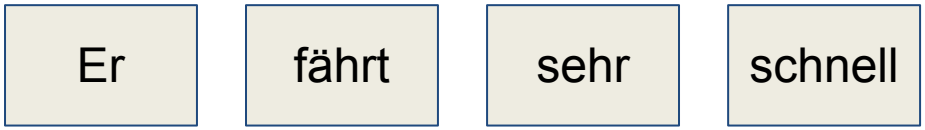
Let’s decode a few examples: (exercise)

Did you keep multiple options? called **Beam search**

Decoding (Translation generation)

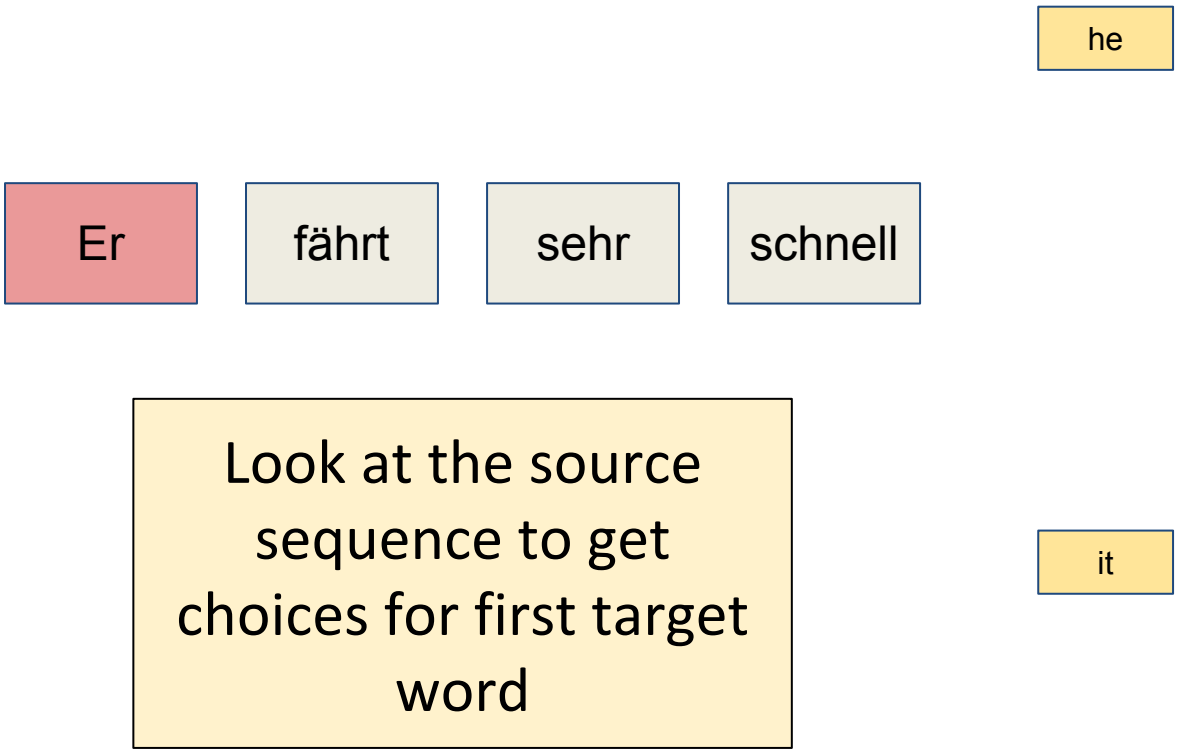


Decoding (Translation generation)

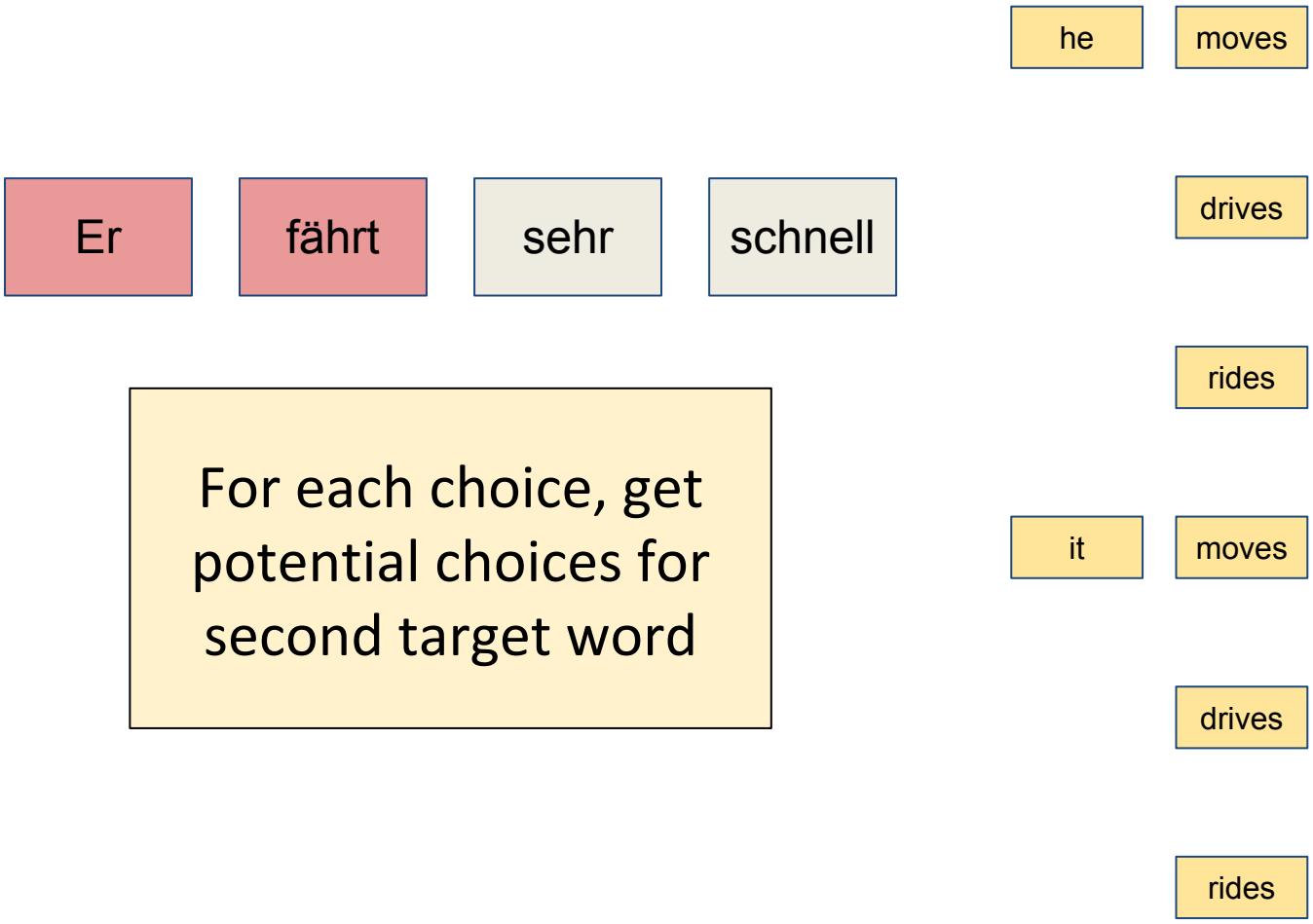


Translation is generated from left to right

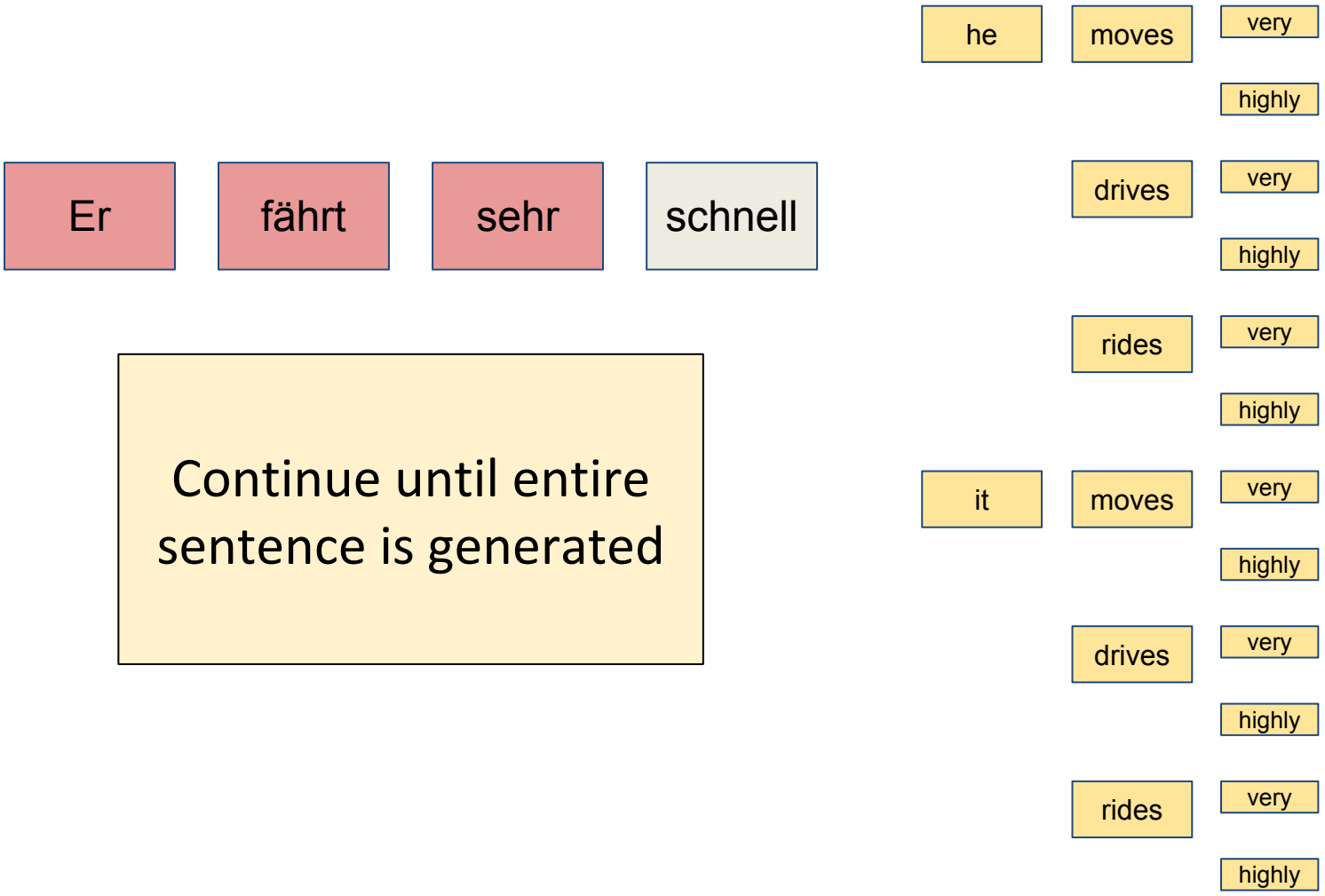
Decoding (Translation generation)



Decoding (Translation generation)



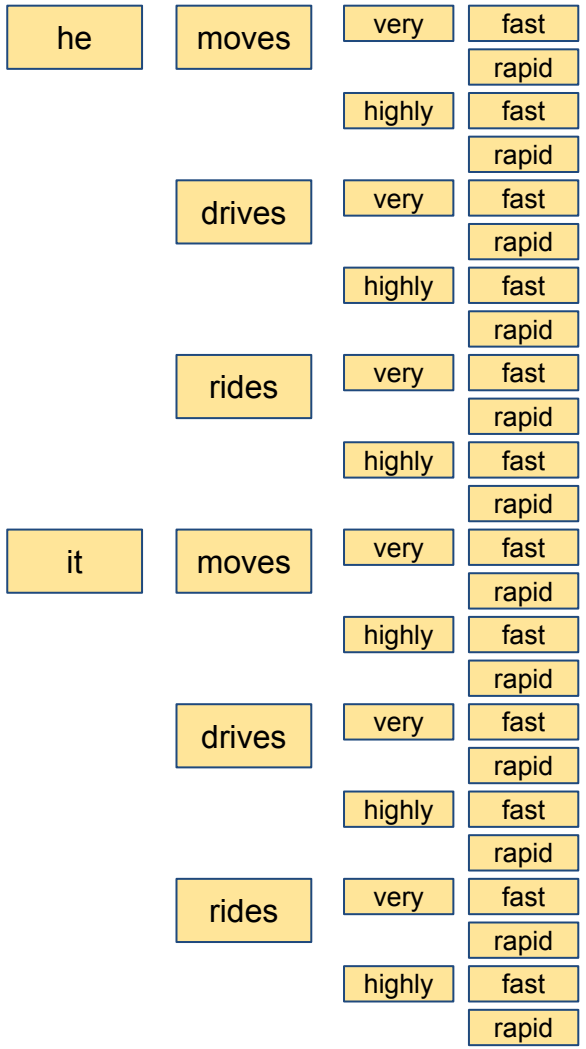
Decoding (Translation generation)



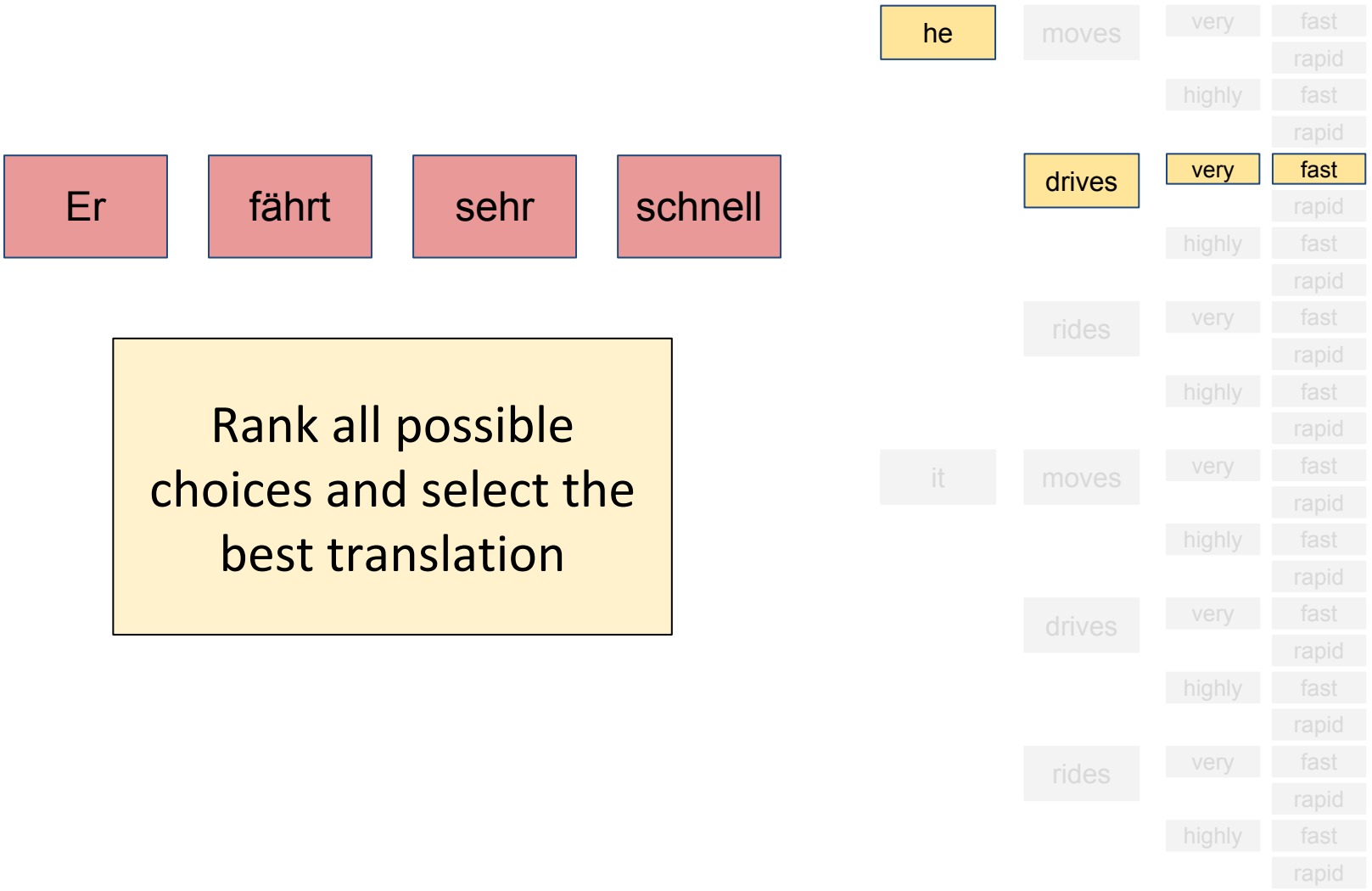
Decoding (Translation generation)

Er | fährt | sehr | schnell

Continue until entire sentence is generated



Decoding (Translation generation)



Evaluation

How good is a translation in terms of meaning and fluency?

- Human evaluation
- Automatic evaluation

Evaluation

Ten translations of a Chinese sentence:

这个机场的安全由以色列方面负责.

Israeli officials are responsible for airport security.

Israel is in charge of the security at this airport.

The security work for this airport is the responsibility of the Israeli government.

Israeli side was in charge of the security of this airport.

Israel is responsible for the airport's security.

Israel is responsible for safety work at this airport.

Israel presides over the security of the airport.

Israel took charge of the airport security.

The safety of this airport is taken charge of by Israel.

This airport's security is the responsibility of the Israeli security officials.

Evaluation

Ten translations of a Chinese sentence:

这个机场的安全由以色列方面负责.

Israeli officials are responsible for airport security.

Israel is in charge of the security at this airport.

The security work for this airport is the responsibility of the Israeli government.

Israeli side was in charge of the security of this airport.

Israel is responsible for the airport's security.

Israel is responsible for safety work at this airport.

Israel presides over the security of the airport.

Israel took charge of the airport security.

The safety of this airport is taken charge of by Israel.

This airport's security is the responsibility of the Israeli security officials.

Human Evaluation

- Given a source sentence and/or a reference sentence, rate the translation output

- Adequacy - does the translation hold the meaning as in the source sentence?
- Fluency - Is it a grammatically and syntactically fluent sentence?

Human Evaluation

- Slow and expensive
- Inconsistency among evaluators
- Need of bilingual speakers
- Not practical (for every system run)

Automatic Evaluation

Given a **reference** translation

- automatically score a translation
  - adequacy and fluency

- BLEU
- Meteor
- TER
- WER
- …

Automatic Evaluation

Given a **reference** translation

- automatically score a translation
  - adequacy and fluency

- BLEU (most commonly used metric)
- Meteor
- TER
- WER
- …

Automatic Evaluation - BLEU

- Considers two things between translation output and reference
  - ngram overlap
  - brevity penalty (length difference)

$$\text{BLEU} = \min\left(1, \frac{\text{output length}}{\text{reference length}}\right)\left(\prod_{i=1}^{4} \text{precision}_i\right)^{\frac{1}{4}}$$

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

- Considers two things between translation output and reference
  - ngram overlap
  - brevity penalty (length difference)

$$\text{BLEU} = \min\left(1, \frac{\text{output length}}{\text{reference length}}\right)\left(\prod_{i=1}^{4} \text{precision}_i\right)^{\frac{1}{4}}$$

Penalizes shorter sentences

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

- Considers two things between translation output and reference
  - ngram overlap
  - brevity penalty (length difference)

$$\text{BLEU} = \min\left(1, \frac{\text{output length}}{\text{reference length}}\right)\left(\prod_{i=1}^{4} \text{precision}_i\right)^{\frac{1}{4}}$$

precision of ngrams of size 1 to 4

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output	Israeli officials are responsible for security

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output	Israeli officials are responsible for security

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output	Israeli officials are responsible for security

4gram matches 1gram match

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output	Israeli officials are responsible for security

Metric	Output
precision (1gram)	6/6
precision (2gram)	4/5
precision (3gram)	3/4
precision (4gram)	2/3
brevity penalty	6/7
BLEU	68%

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output 1	Israeli officials responsibility of airport safety
Output 2	airport security Israeli officials are responsible

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output 1	Israeli officials responsibility of airport safety
Output 2	airport security Israeli officials are responsible

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output 1	Israeli officials responsibility of airport safety 2gram match 1gram match
Output 2	airport security Israeli officials are responsible 2gram match 4gram match

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

Reference	Israeli officials are responsible for airport security
Output 1	Israeli officials responsibility of airport safety 2gram match 1gram match
Output 2	airport security Israeli officials are responsible 2gram match 4gram match

Metric	Output 1	Output 2
precision (1gram)	3/6	6/6
precision (2gram)	1/5	4/5
precision (3gram)	0/4	2/4
precision (4gram)	0/3	1/3
brevity penalty	6/7	6/7
BLEU	0%	52%

http://mt-class.org/jhu/slides/lecture-evaluation.pdf

Automatic Evaluation - BLEU

- Efficient
- Cheap
- Easy to integrate

Automatic Evaluation - BLEU

- Efficient
- Cheap
- Easy to integrate

Issues

- ignores
  - synonyms
  - partial matches (morphologically rich languages)
- penalizes short translations only
- high score may not result in better system

Automatic Evaluation - BLEU

- Efficient
- Cheap
- Easy to integrate

Issues

- ignore
  - sync
  - part anguages)
- penalizes short translations only
- high score may not result in better system

Still widely used :)

Lecture Summary

- Language
  - syntax, semantics, morphology
- Translation
  - parallel corpus
  - word and phrase alignment
  - generation
- Evaluation
  - human vs. automatic

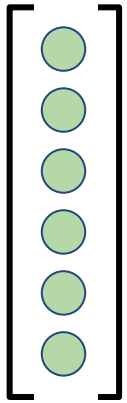
Extras

Linear Algebra in 5 minutes

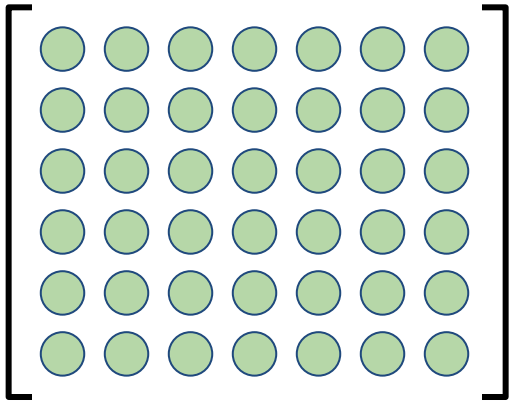
Need to know a lot of linear algebra to study and create new algorithms,

but *not to use* the algorithms!

Linear Algebra in 5 minutes

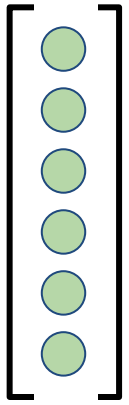


Vector



Matrix

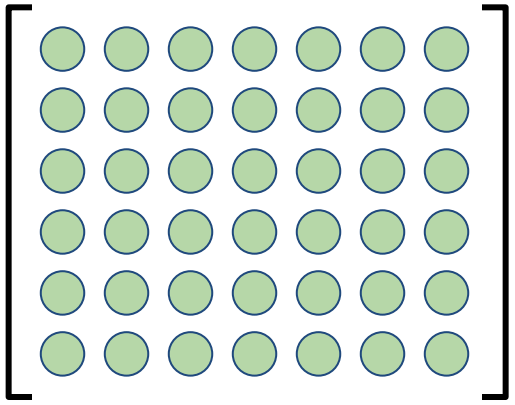
Linear Algebra in 5 minutes



Vector

length = 6

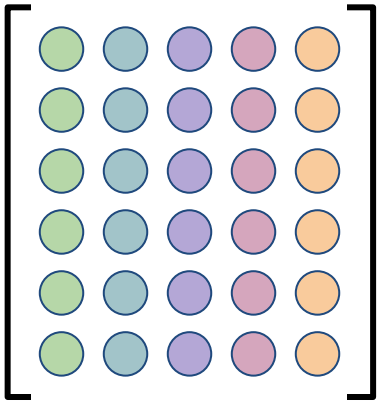
[6 x 1]



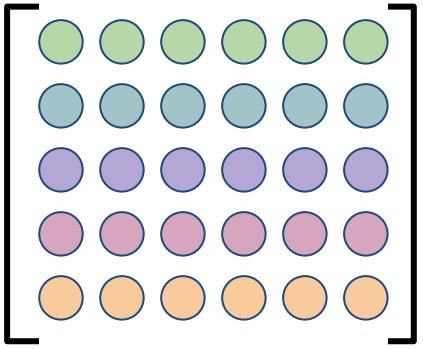
Matrix

[6 x 7]

Linear Algebra in 5 minutes



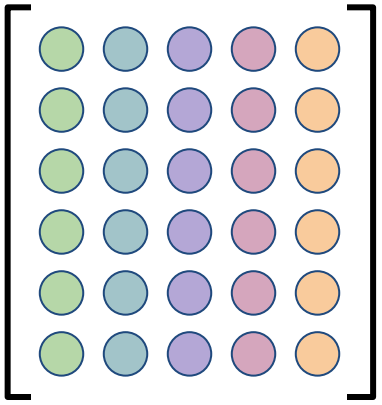
Matrix



Matrix
Transpose

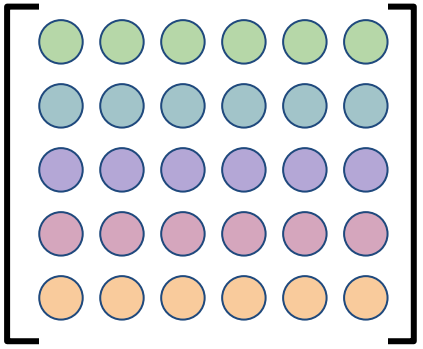
Linear Algebra in 5 minutes

[6 x 5]



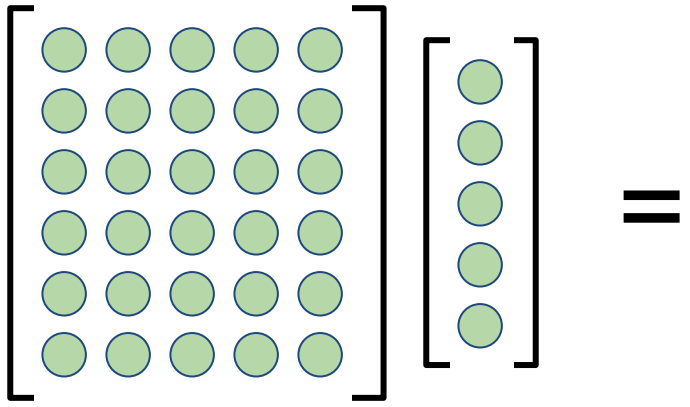
Matrix

[5 x 6]



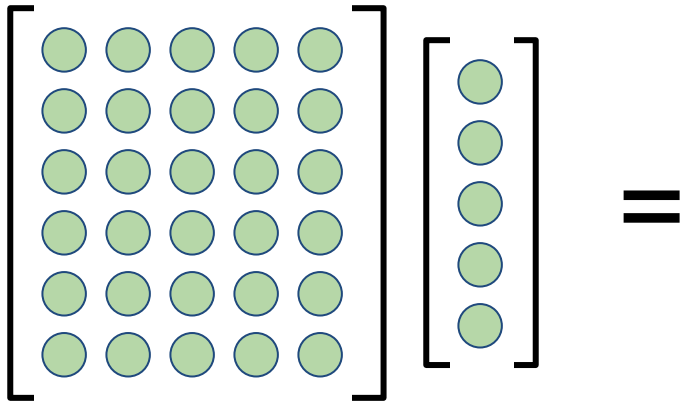
Matrix
Transpose

Linear Algebra in 5 minutes



Matrix-Vector product

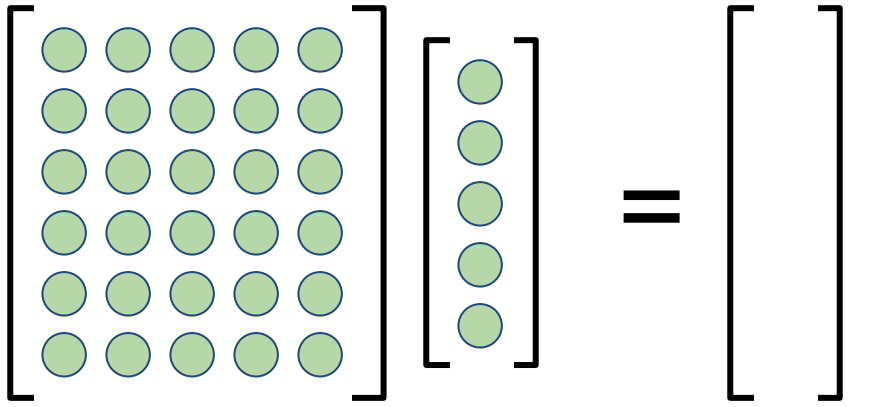
Linear Algebra in 5 minutes



[6 x 5] [5 x 1]

Matrix-Vector product

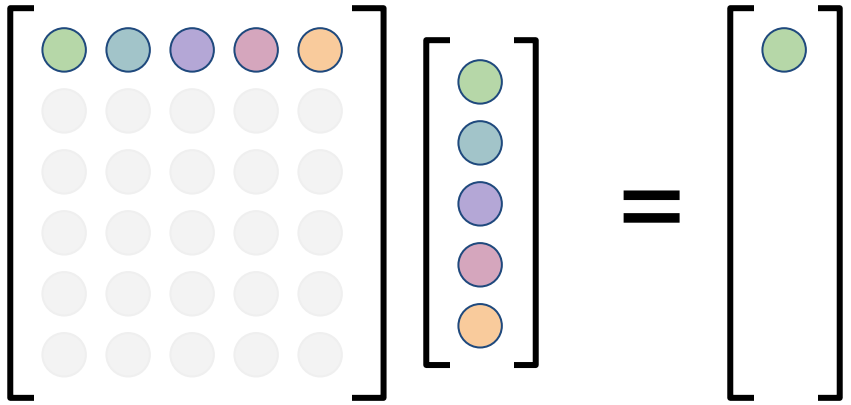
Linear Algebra in 5 minutes



[6 x 5] [5 x 1] [6 x 1]

Matrix-Vector product

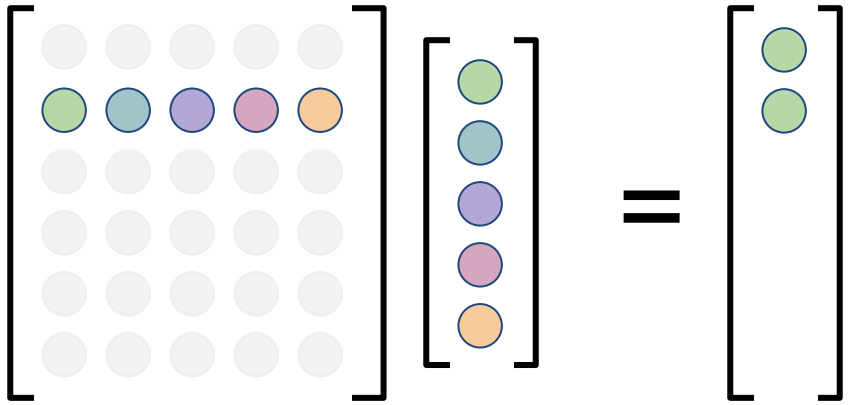
Linear Algebra in 5 minutes



[6 x 5] [5 x 1] [6 x 1]

Matrix-Vector product

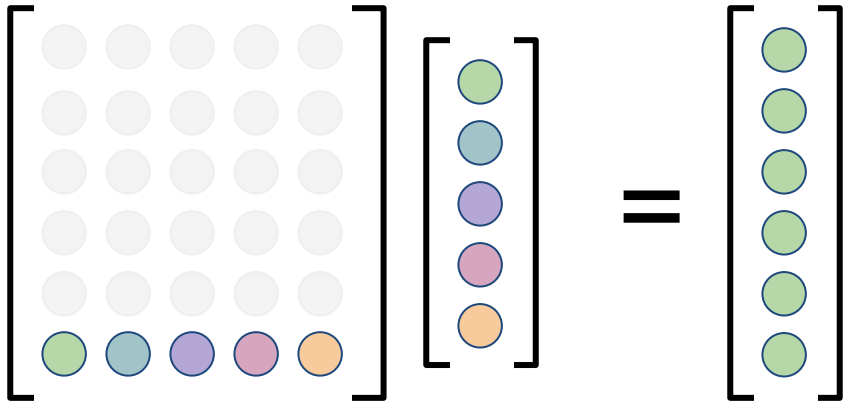
Linear Algebra in 5 minutes



[6 x 5] [5 x 1] [6 x 1]

Matrix-Vector product

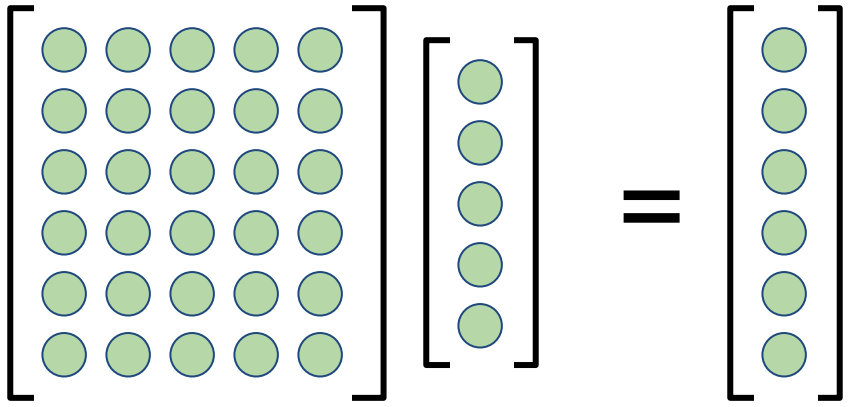
Linear Algebra in 5 minutes



[6 x 5] [5 x 1] [6 x 1]

Matrix-Vector product

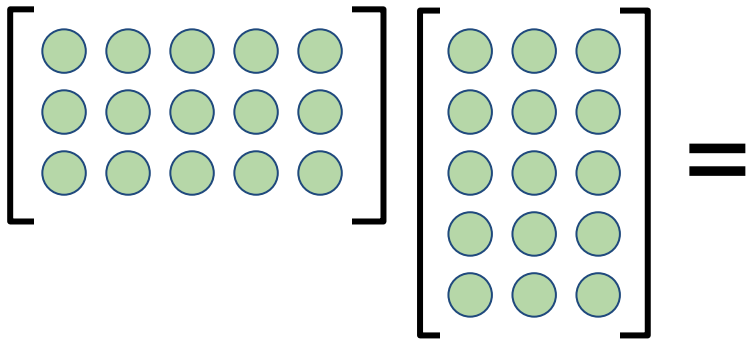
Linear Algebra in 5 minutes



[6 x 5] [5 x 1] [6 x 1]

Matrix-Vector product

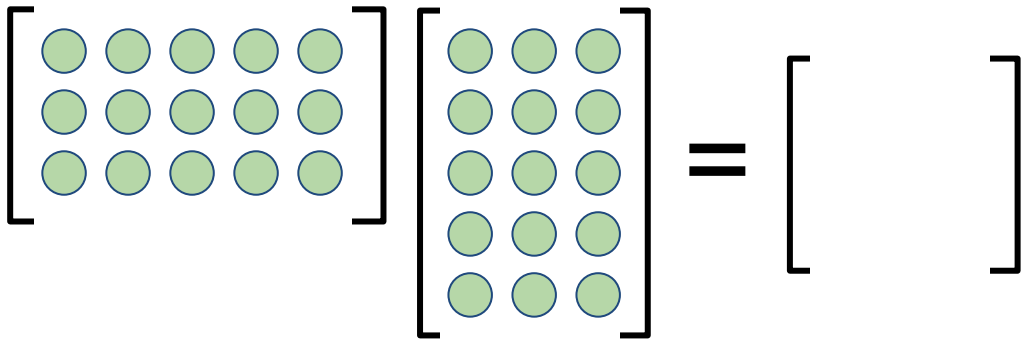
Linear Algebra in 5 minutes



[3 x 5] [5 x 3]

Matrix-Matrix product

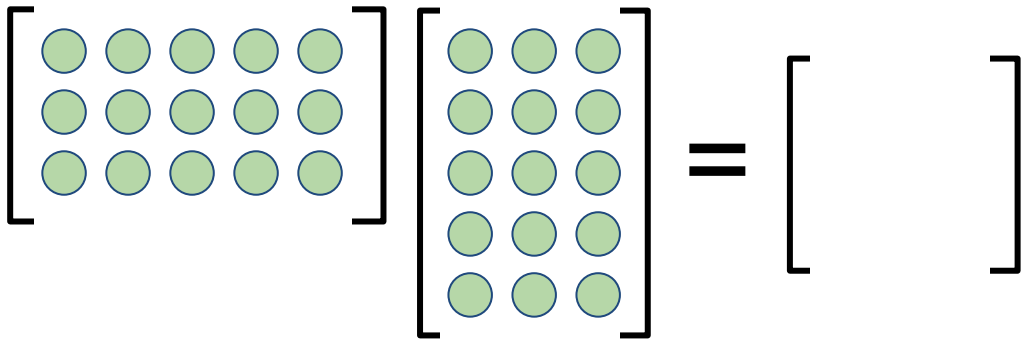
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

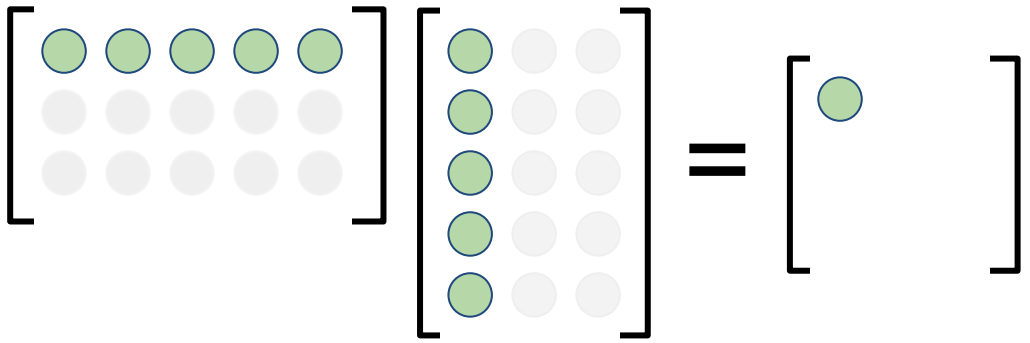
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

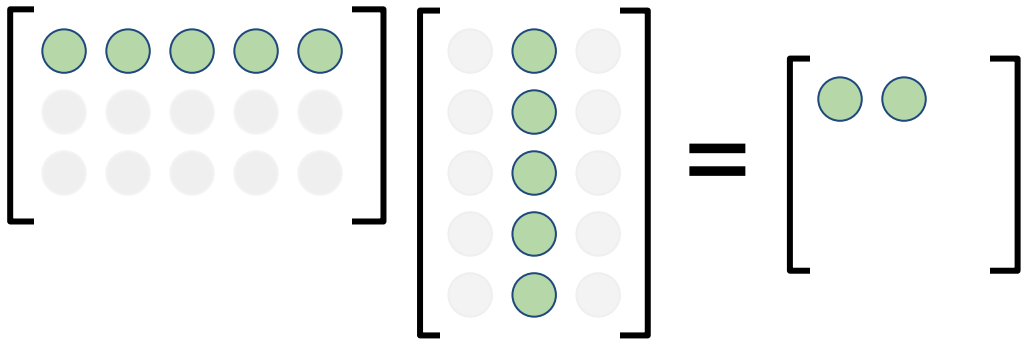
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

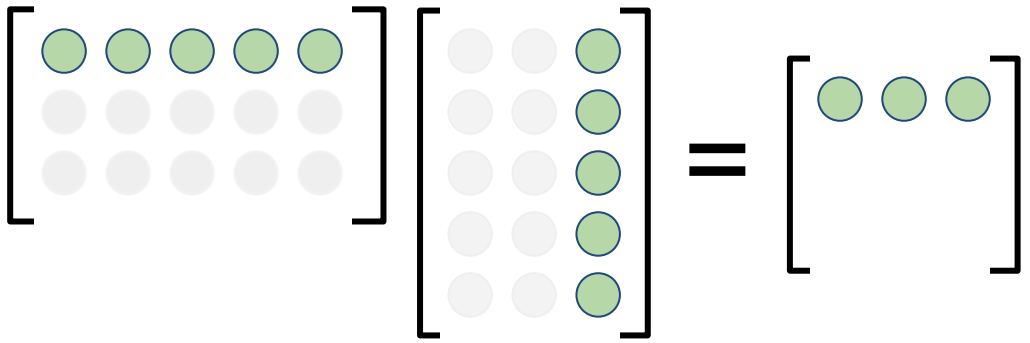
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

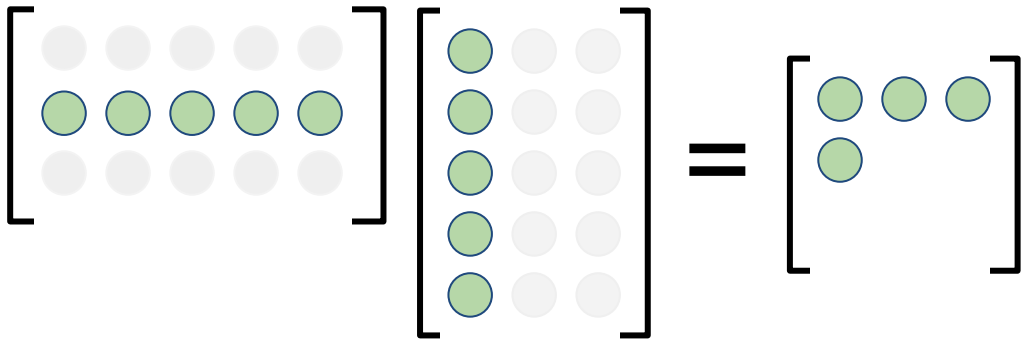
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

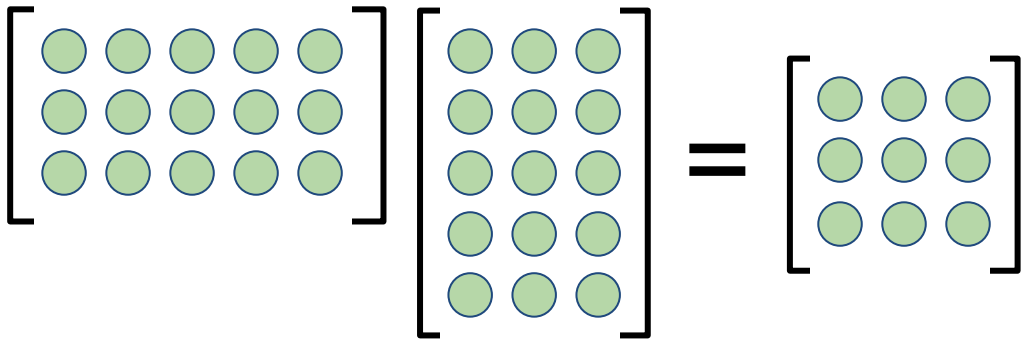
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

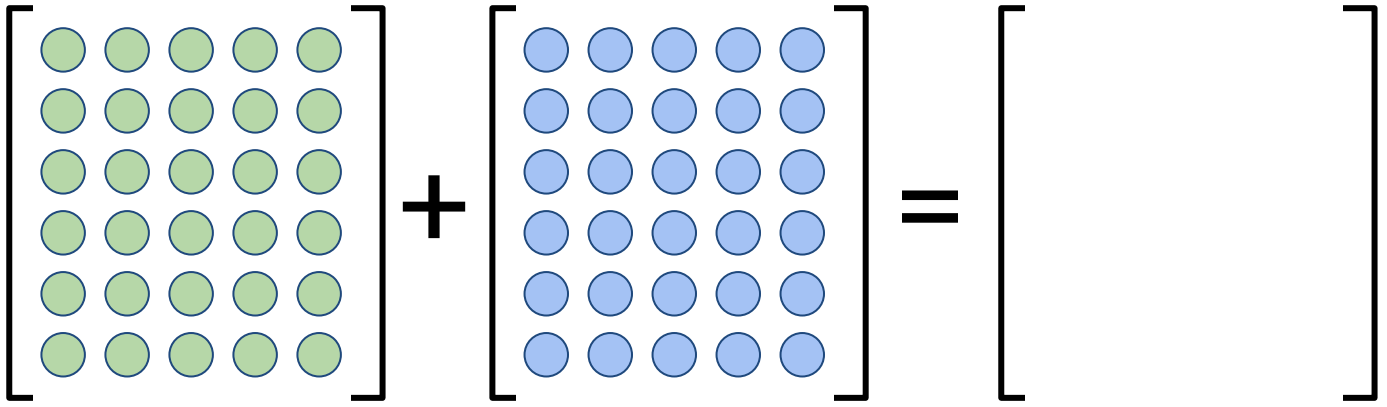
Linear Algebra in 5 minutes



[3 x 5] [5 x 3] [3 x 3]

Matrix-Matrix product

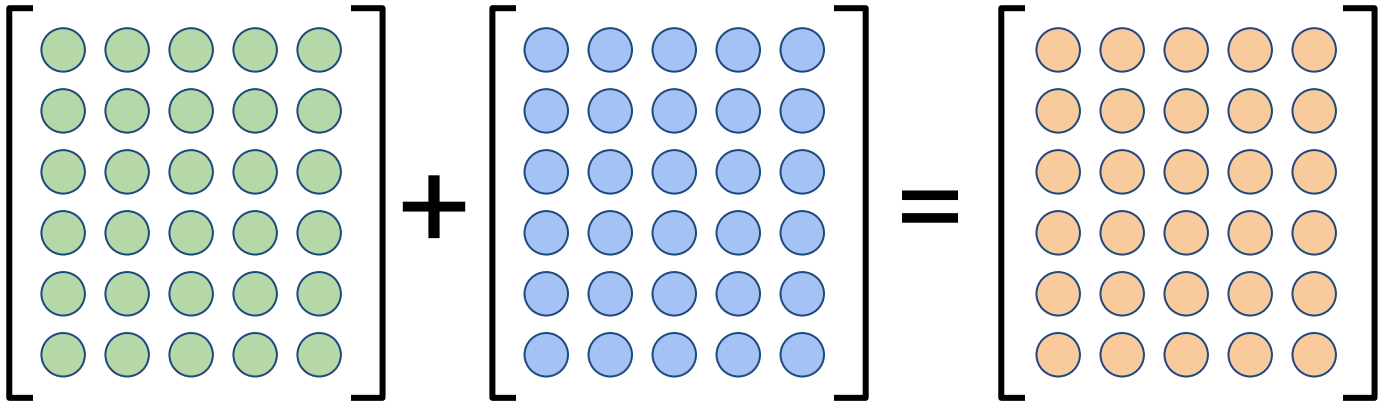
Linear Algebra in 5 minutes



[6 x 5] [6 x 5] [6 x 5]

Matrix
Addition/Subtraction

Linear Algebra in 5 minutes



[6 x 5] [6 x 5] [6 x 5]

Matrix
Addition/Subtraction