



**University of  
Zurich** <sup>UZH</sup>

Institute of Computational Linguistics

---

# Machine Translation

## 14 Summary & QA

Mathias Müller



**University of  
Zurich** <sup>UZH</sup>

**Institute of Computational Linguistics**



# Course Summary

# 1 Introduction



## 2 Evaluation

Manual evaluation

+ more reliable

- costly

- slow

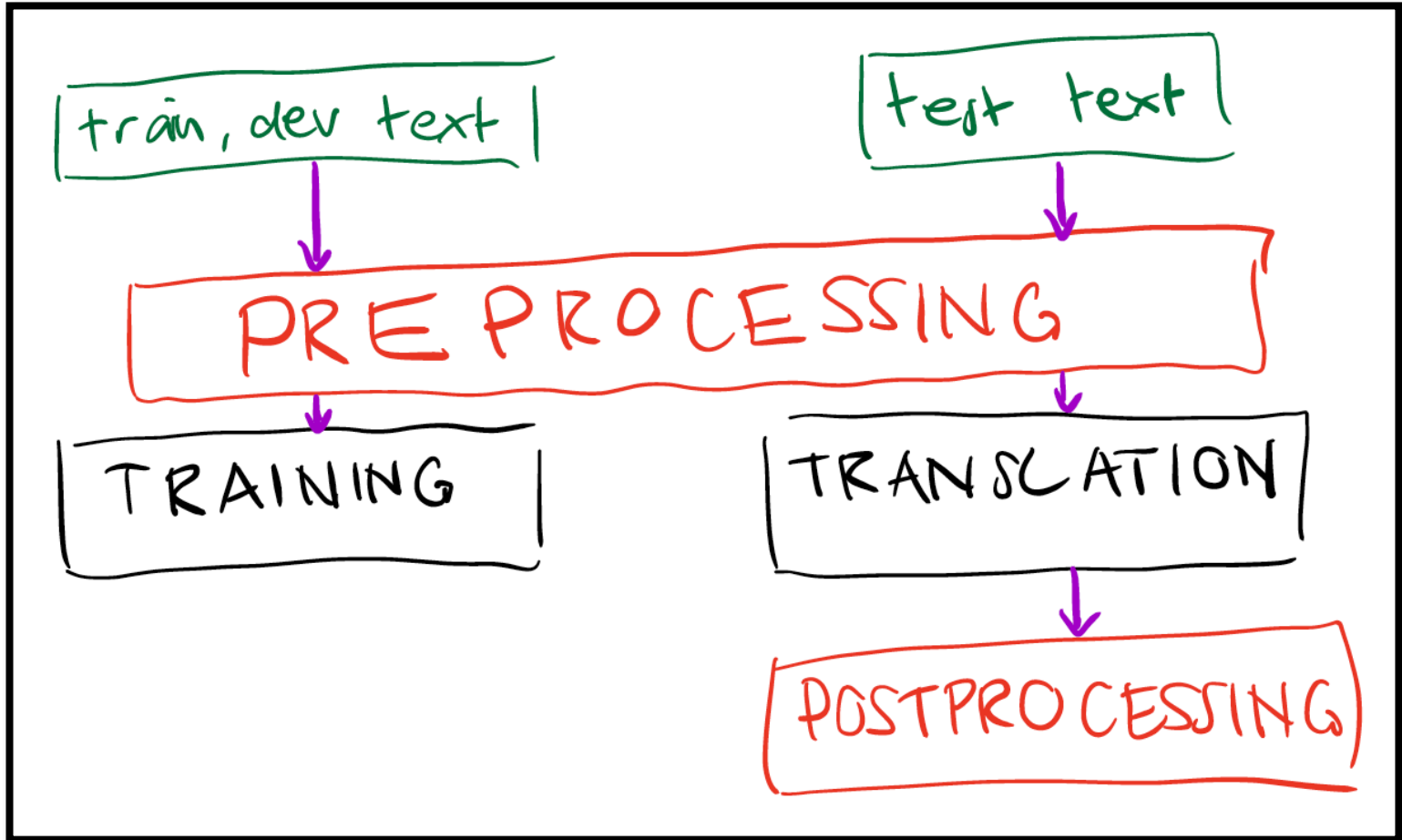
Automatic evaluation

- less reliable

+ cheap

+ fast

# 3 Preprocessing



# 4 Statistical translation

## What a language model looks like

igram

- contains ngrams of a certain order, together with their probability

target ngrams	probability
natürlich ,	0.000071
er ist	0.000065
ist sehr	0.000083

## What a translation model looks like

phrase tab

phrases = ngrams

- contains phrases in two languages, together with their translation probability

source phrases	target phrases	probability
natürlich	of course	0.7
natürlich	natural	0.3
natürlich ,	of course ,	0.000031
Fallout 76	crap game	0.679

# 5 Math fundamentals

Matrix-vector multiplication: right

$$m_1 = [1 \ 2 \ 3]$$

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{bmatrix}$$

$2 \times 3$

$$\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

3

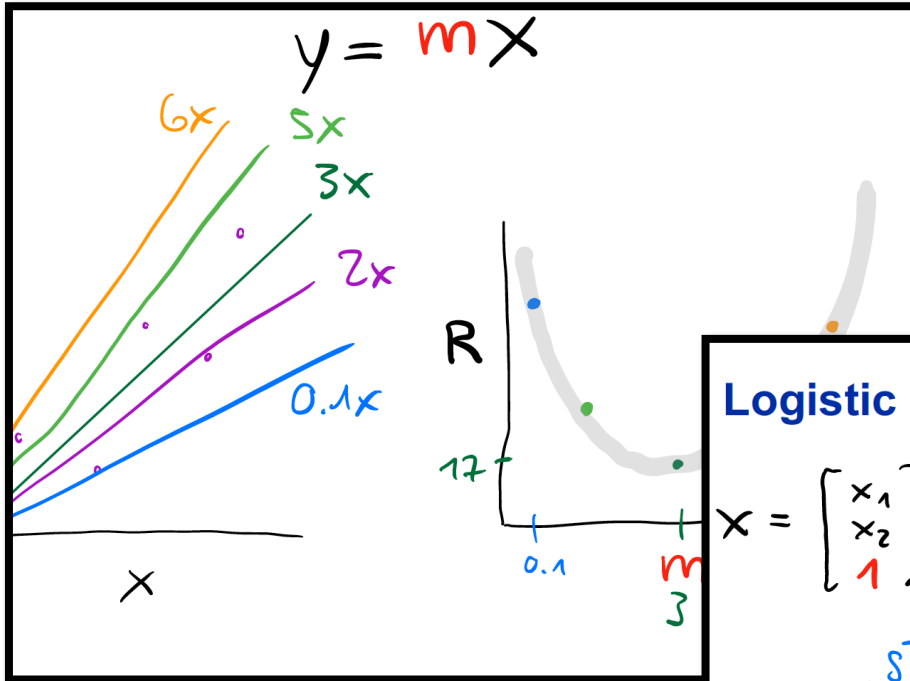
$$M\vec{a} = \begin{bmatrix} (m_1)^T \cdot \vec{a} \\ (m_2)^T \cdot \vec{a} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \frac{1 \times 1 + 2 \times 2 + 3 \times 3}{1} \\ \frac{0 \times 1 + (-1) \times 2 + 1 \times 3}{-2 \quad 3} \end{bmatrix}$$

$$\begin{bmatrix} 14 \\ 1 \end{bmatrix}$$

# 6 Linear models



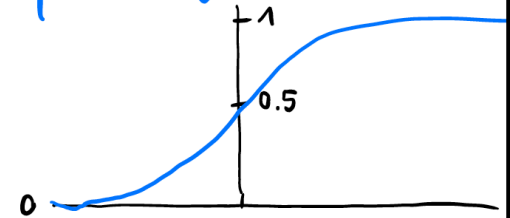
## Logistic Regression

$$x = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \quad c = \begin{bmatrix} c_1 \\ c_2 \\ b \end{bmatrix}$$

sigmoid

$$y = \sigma(\vec{c} \cdot \vec{x})$$

"squashing"



$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

interpretation:  $y$  is the probability of the positive class



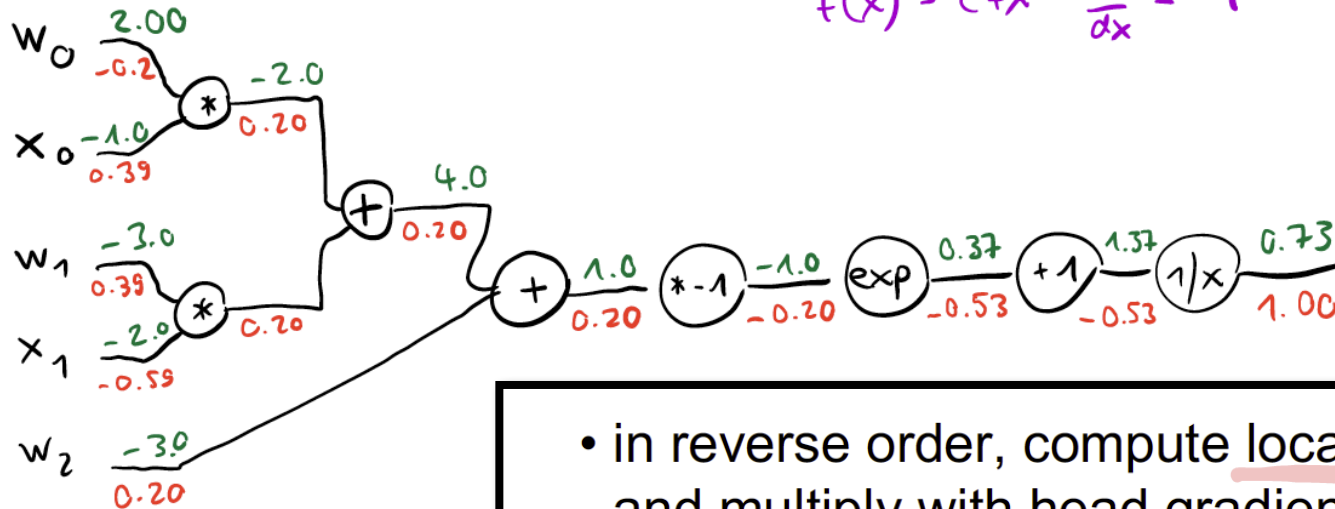
# 7 Feed-forward neural networks

## Computational graph view

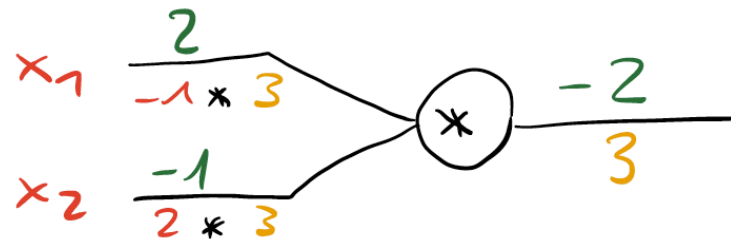
## RULES

$$f(x) = \frac{1}{x} \quad \frac{df}{dx} = -1/x^2$$

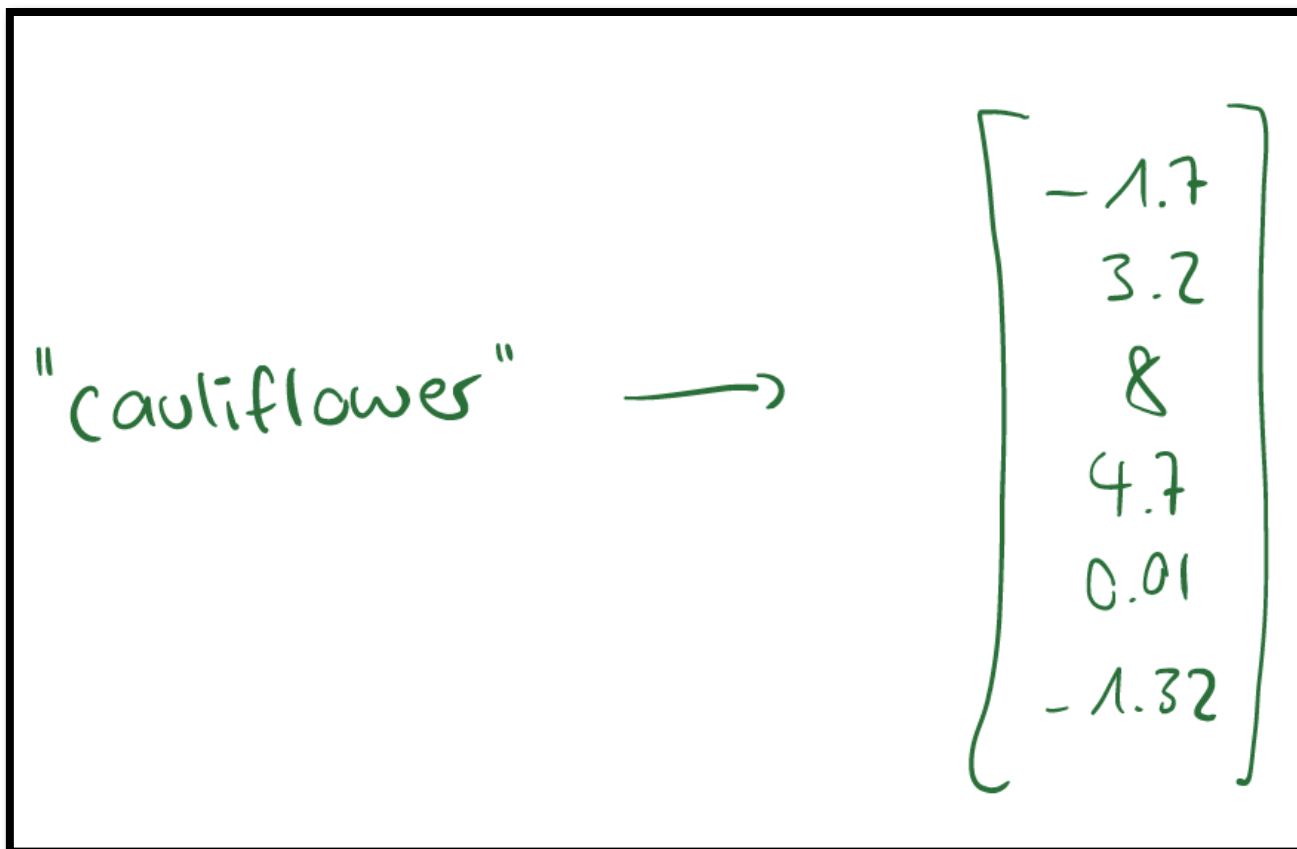
$$f(x) = c+x \quad \frac{df}{dx} = 1$$



- in reverse order, compute local gradient and multiply with head gradient



## 8 Word embeddings



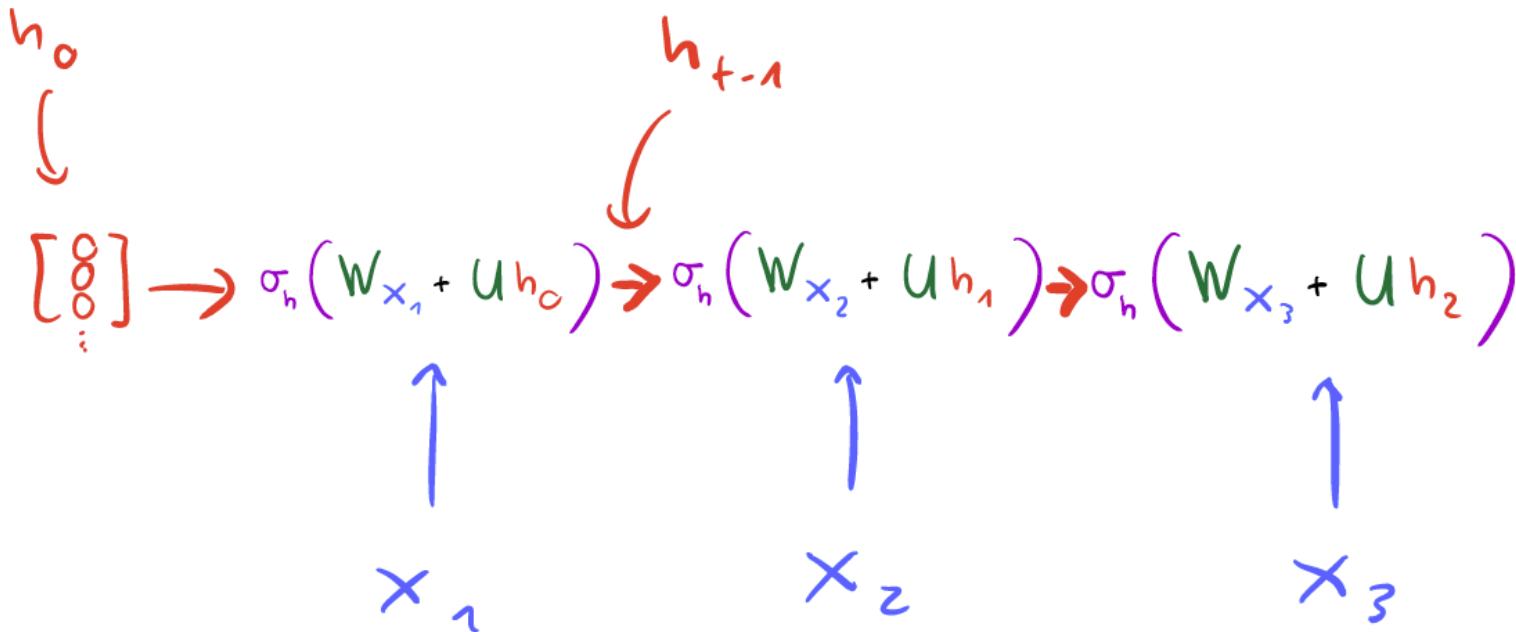
# 8 RNNs

Elman ("Vanilla") RNN

## Unfolding / unrolling RNN graph

$$h_t = \text{RNN}(x_t, h_{t-1})$$

Key observation: RNNs are just very deep FFNNs where layers **share weights**



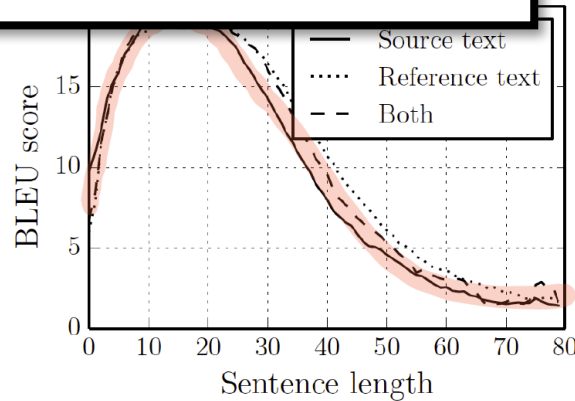
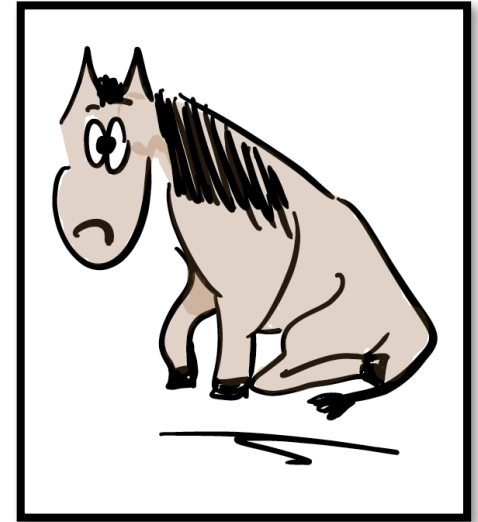
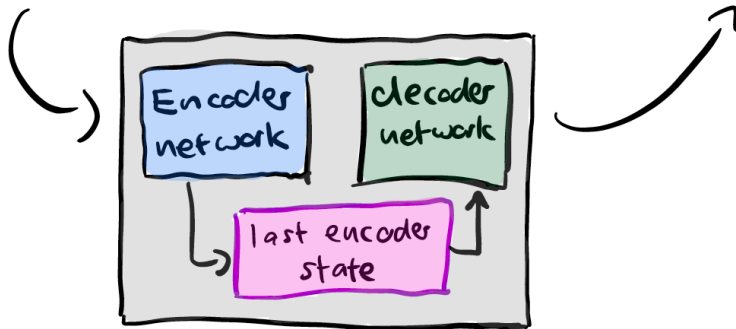
# 9 Tensorflow



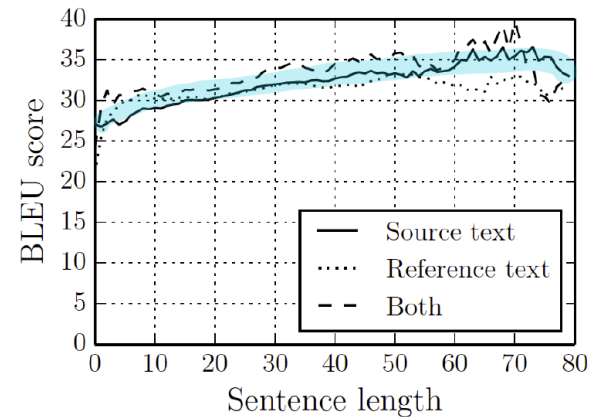
# 10 Encoder-decoder models

He is out  
of his mind.

Er ist  
ausser sich.



NMT\*



SMT

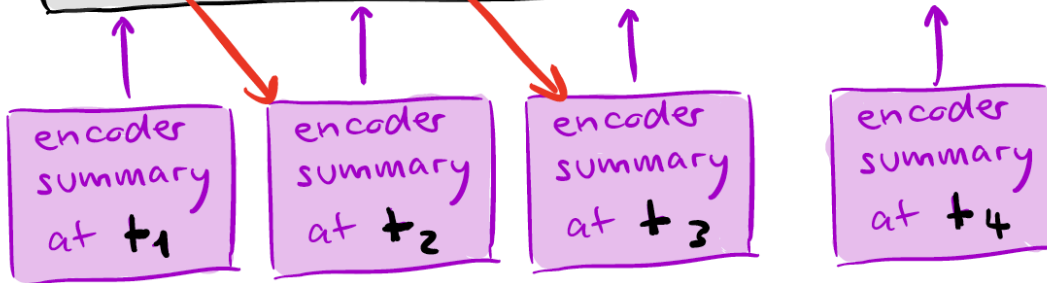
# 11 Attention



A woman is throwing a frisbee in a park.

**Weighted sum of all encoder states**

$t_1$        $t_2$        $t_3$        $t_4$   
Er      ist      auser      sich



[  
⋮  
]

# 12 Decoding

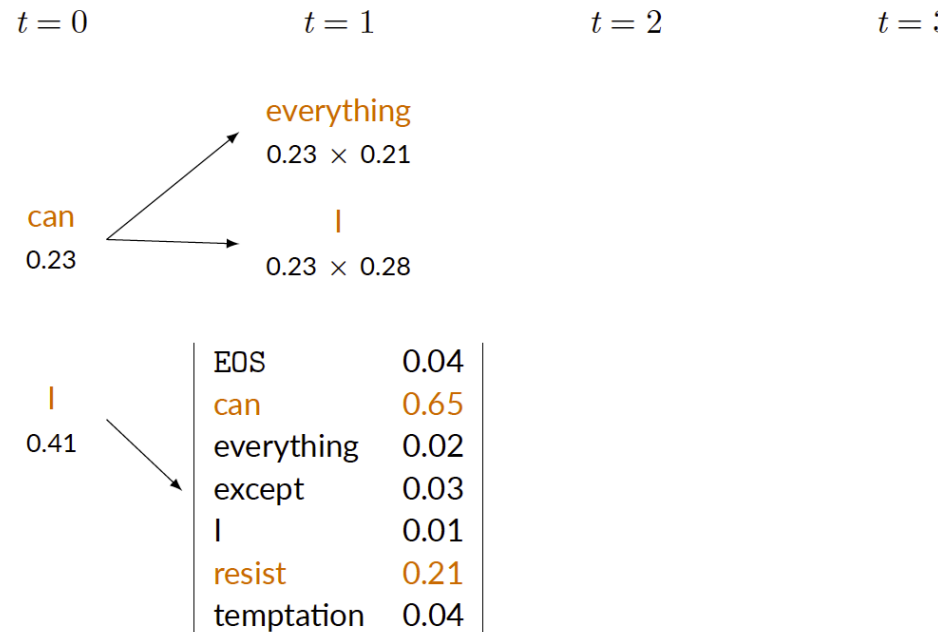
## Grösse des Suchraums: Beispiel

- $V = \{ \text{EOS, can, everything, except, I, resist, temptation} \}$
- $|V| = 7$
- $\text{max\_len} = 15$
- Grösse des Suchraums:  $7^{15} = 4,747,561,509,943$  Sätze

→ Welches ist der beste Satz?

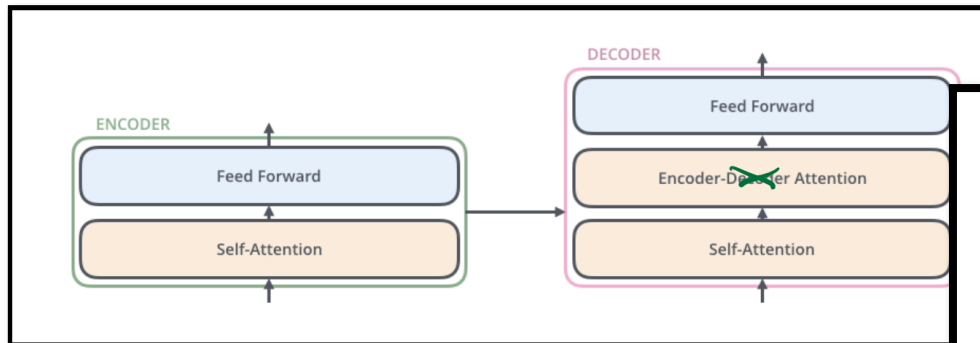
## Random Sampling: Beispiel

- $V = \{ \text{EOS, can, everything, except, I, resist, temptation} \}$
- Eingabe: ich kann allem widerstehen, ausser der Versuchung
- Mögliche Ausgaben (bei mehreren Durchläufen):
  - everything except temptation I can resist
  - I can resist everything except temptation
  - except temptation I can resist everything
  - can I resist everything except temptation
  - I resist can everything except temptation



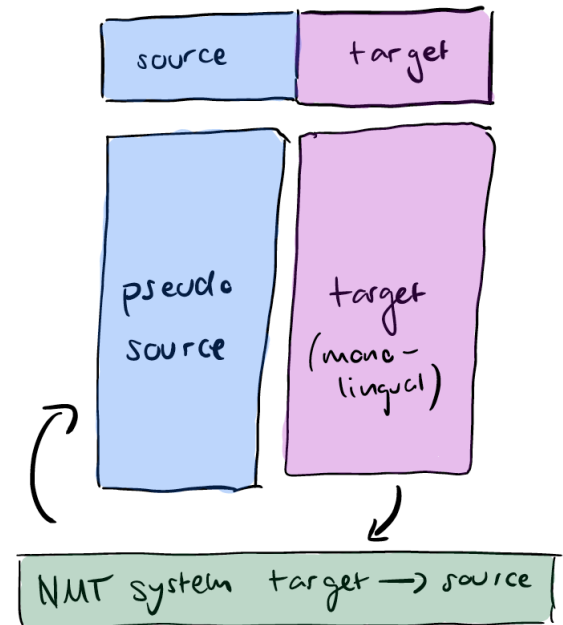
# 13 "Methods that make current NMT work well"

## Transformer models



<http://jalammar.github.io/illustrated-transformer/>

## Back-translation





Type	Discussion topics	Author	Last modified	Marked	New	Posts
🔒	<a href="#">Sammel-Thread Fragen für Exam Q&amp;A</a>	Müller, Mathias	5/27/2019 10:55 AM	0	0	4
🔒	<a href="#">Informationen zur Prüfung</a>	Müller, Mathias	5/16/2019 4:01 PM	0	0	1
🔒	<a href="#">ÜB: Fehlermeldung bei Übersetzen auf R2D2</a>	Conrad, Sophia	5/15/2019 12:59 PM	0	0	3
🔒	<a href="#">Error while finding spec for 'socketye.translate' (ImportError: No module named 'socketye')</a>	Engelmann, Jasmin	5/27/2019 10:49 AM	0	0	2
🔒	<a href="#">Fehlermeldung bei Attention-Visualisierung</a>	Borrello, Roberto	5/22/2019 11:22 AM	0	0	7
🔒	<a href="#">Ü6A3: vocab.trg.0.json nicht vorhanden</a>	Conrad, Sophia	5/17/2019 11:57 PM	0	0	3
🔒	<a href="#">sacredBLEU: The input and reference stream(s) were of different lengths.</a>	Angliker, Vivien	5/17/2019 2:32 PM	0	0	3
🔒	<a href="#">Fehlermeldung beim Trainingsstart</a>	Suter, Benjamin	5/12/2019 10:37 PM	0	0	7
🔒	<a href="#">Source-Sequenz umkehren</a>	Suter, Benjamin	5/12/2019 6:44 PM	0	0	5
🔒	<a href="#">Loss-Berechnung: Fragen über Fragen</a>	Säuberli, Andreas	5/12/2019 4:06 PM	0	0	7
🔒	<a href="#">Syntaxfehler in train.py</a>	Steiger, Melvin Samson	5/12/2019 11:00 AM	0	0	3
🔒	<a href="#">No module error</a>	Stein, Varvara	5/11/2019 9:52 PM	0	0	7
🔒	<a href="#">ResourceExhaustedError</a>	Stein, Varvara	5/11/2019 9:49 PM	0	0	3
🔒	<a href="#">Billing issue</a>	Stein, Varvara	5/11/2019 9:44 PM	0	0	5
🔒	<a href="#">ValueError</a>	Engelmann, Jasmin	5/11/2019 9:33 AM	0	0	3
🔒	<a href="#">Übersetzung auf GitHub oder im zip?</a>	Forster, Martina	5/10/2019 5:25 PM	0	0	2
🔒	<a href="#">Fehlermeldung bei MultiRNNCell</a>	Kiener, Sarah Elisabeth	5/9/2019 9:37 PM	0	0	3
🔒	<a href="#">Preprocessing Übung 5</a>	Forster, Martina	5/9/2019 12:18 PM	0	0	5
🔒	<a href="#">FileNotFound</a>	Forster, Martina	5/8/2019 5:49 PM	0	0	5
🔒	<a href="#">CUDA_ERROR_NO_DEVICE</a>	Borrello, Roberto	5/6/2019 10:05 PM	0	0	7
🔒	<a href="#">Erinnerung an Übung 2</a>	Borrello, Roberto	5/5/2019 3:12 PM	0	0	3
🔒	<a href="#">Fehlermeldung bei Versuch Training zu starten</a>	Conrad, Sophia	5/5/2019 2:11 PM	0	0	6
🔒	<a href="#">Dropout oder Early Stopping</a>	Borrello, Roberto	5/4/2019 3:53 PM	0	0	3
🔒	<a href="#">Serverkosten, Guthaben und Neustarts</a>	Martinez, Dominik	5/3/2019 11:10 AM	0	0	5
🔒	<a href="#">BPE-Modell vergrößern &amp; BLEU Score</a>	Forster, Martina	5/2/2019 5:28 PM	0	0	2
🔒	<a href="#">pip install funktioniert nicht</a>	Beuret, Debora	5/2/2019 4:02 PM	0	0	7
🔒	<a href="#">Übung 4 API GCP Quotas</a>	Chi, Margaret	4/29/2019 5:53 PM	0	0	7
🔒	<a href="#">Eigene Änderungen</a>	Engelmann, Jasmin	4/27/2019 4:47 PM	0	0	2
🔒	<a href="#">Training starten</a>	Engelmann, Jasmin	4/26/2019 2:53 PM	0	0	5
🔒	<a href="#">.bashrc-Datei wird nicht gefunden</a>	Suter, Benjamin	4/25/2019 6:22 PM	0	0	3
🔒	<a href="#">Problem with GitHub's maximal file size</a>	Borrello, Roberto	4/24/2019 11:17 AM	0	0	3
🔒	<a href="#">Google Cloud Guthaben</a>	Forster, Martina	4/24/2019 11:00 AM	0	0	2
🔒	<a href="#">Übung4, Datenset</a>	Kiener, Sarah Elisabeth	4/23/2019 1:26 PM	0	0	3
🔒	<a href="#">Romanesco freiwilliges Update zu Version 0.2</a>	Müller, Mathias	4/23/2019 9:29 AM	0	0	1
🔒	<a href="#">Tensorboard starten</a>	Engelmann, Jasmin	4/23/2019 9:00 AM	0	0	2
🔒	<a href="#">Anzahl Types</a>	Forster, Martina	4/5/2019 10:56 AM	0	0	3
🔒	<a href="#">Training Zeit</a>	Stein, Varvara	3/24/2019 10:11 PM	0	0	5
🔒	<a href="#">Installation von sockey</a>	Steiger, Melvin Samson	3/24/2019 9:19 PM	0	0	7
🔒	<a href="#">Pip auf dem Server</a>	Klimashevskva, Olga	3/24/2019 1:50 PM	0	0	5
🔒	<a href="#">Übung 2, CPU-Kerne</a>	Kiener, Sarah Elisabeth	3/24/2019 12:18 PM	0	0	4
🔒	<a href="#">Übung 2, Fehlermeldung beim Erstellen einer virtuellen Umgebung</a>	Engelmann, Jasmin	3/23/2019 3:25 PM	0	0	2
🔒	<a href="#">Übung 2: Korpus nicht parallel</a>	Nagele, Lorenz	3/19/2019 12:51 PM	0	0	4
🔒	<a href="#">Outputs Übung 2</a>	Borrello, Roberto	3/18/2019 1:48 PM	0	0	6
🔒	<a href="#">BPE Fehler</a>	Forster, Martina	3/17/2019 5:01 PM	0	0	2
🔒	<a href="#">Übung 1.1 - Aggregation</a>	Klimashevskva, Olga	3/11/2019 10:18 PM	0	0	3
🔒	<a href="#">BLEU-Score: "computed over the entire test set"</a>	Säuberli, Andreas	3/5/2019 9:54 PM	0	0	4
🔒	<a href="#">Übung 1 Aufgabe 1.2</a>	Conrad, Sophia	3/1/2019 5:16 PM	0	0	2
🔒	<a href="#">BLEU-Score</a>	Beuret, Debora	2/28/2019 12:26 PM	0	0	2
🔒	<a href="#">Übung 1: Aufgabe 2: Merkwürdige BLEU-Werte</a>	Steiger, Melvin Samson	2/28/2019 11:31 AM	0	0	2
🔒	<a href="#">Übung 1: Einige Fragen</a>	Martinez, Dominik	2/27/2019 6:22 PM	0	0	1



**University of  
Zurich** <sup>UZH</sup>

**Institute of Computational Linguistics**

**QA**

# Instructions

**Out of all questions,**

- Which one would you have the most trouble answering?
- Which one would you be most comfortable answering?

## Round 1

- Team 1 people go to their **trouble question**, Team 2 people go to their **comfortable question**
- Maximum number of people per question: **4**

## Round 2

- Vice versa!

# Exam

- Prüfung findet statt am **18. Juni 2019, 16:15 bis 18:00**. Achtung anderer Raum: **AND-2-48**.
- Prüfung lösen sollte 90 Minuten dauern, ihr habt maximal Zeit bis 18:00.
- Es ist eine sog. "Closed-Book"-Prüfung. Das heisst: es dürfen keine Unterlagen mitgenommen werden. Kein Laptop, kein Papier, kein Handy.
- Mitgenommen werden darf: Getränke, Stifte, Taschenrechner. Ein Taschenrechner ist nicht nötig. **Das Handy darf nicht als Taschenrechner benutzt werden.**
- Die Fragen sind in Deutsch geschrieben. Antworten darf man auf Deutsch oder Englisch. Leute, die nicht Deutsch-Muttersprachler\* innen sind, dürfen ein Wörterbuch mitnehmen. **Das Handy darf nicht als Wörterbuch benutzt werden.**

## Future plans

- Want more MT courses? Check our schedule for HS 2019.
- Planning a BA/MA thesis, Facharbeit or programming project in MT? Come talk to me, I would love to hear about it.
- Points of contact for MT at our institute: Martin Volk, Annette Rios, Samuel Läubli, Mathias Müller (more people arriving soon)