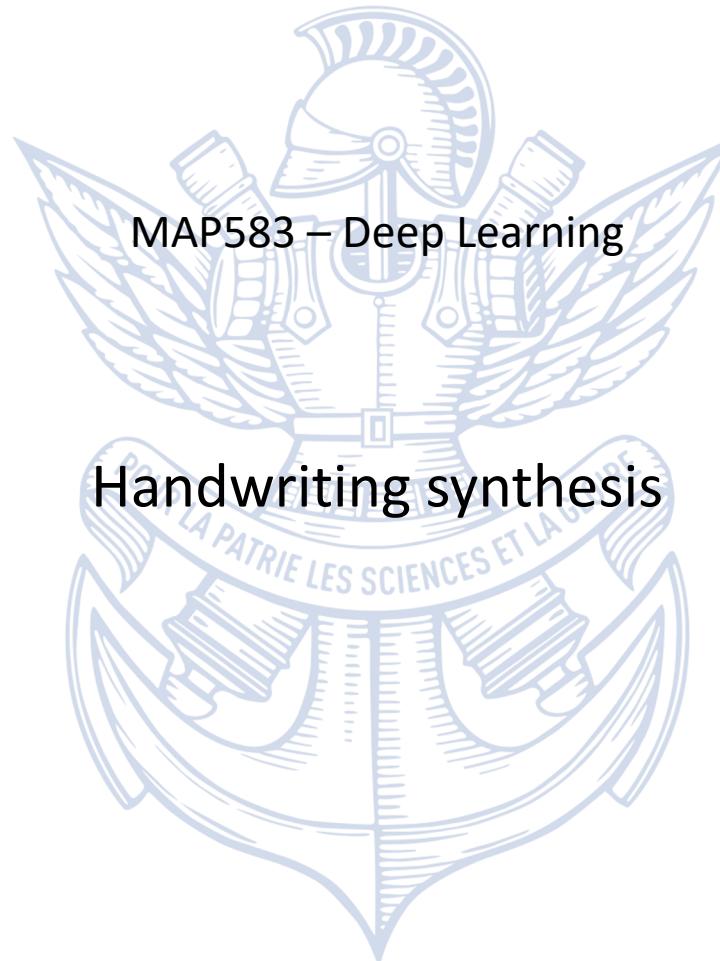


Presentation Deep Learning Project

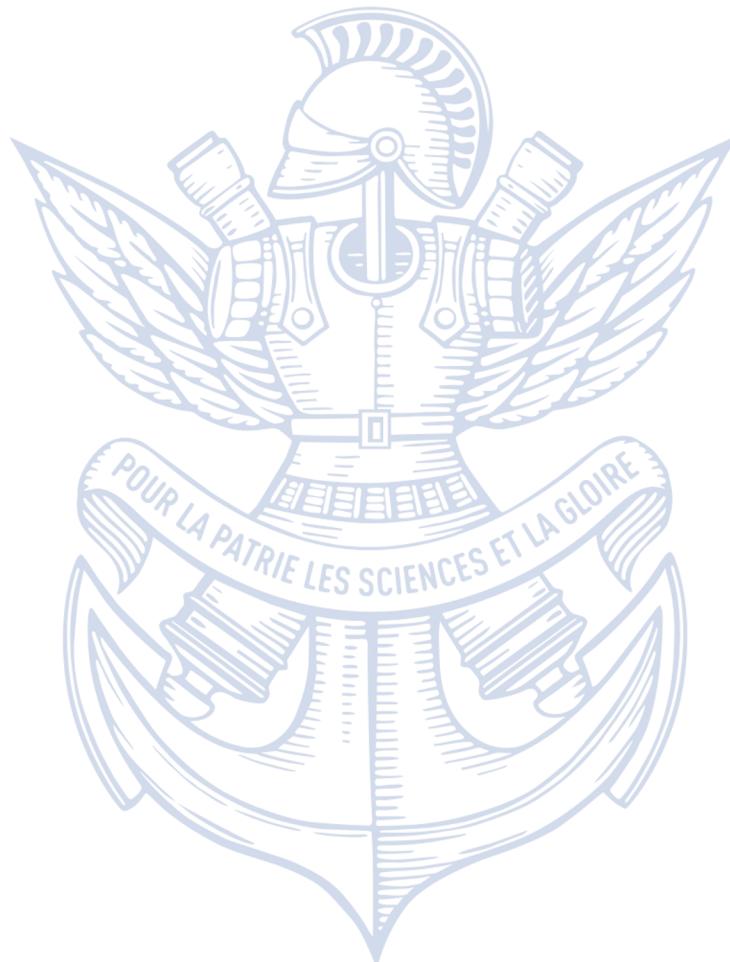


Nicolas HUOT & Alexis BOULEY



March 2022

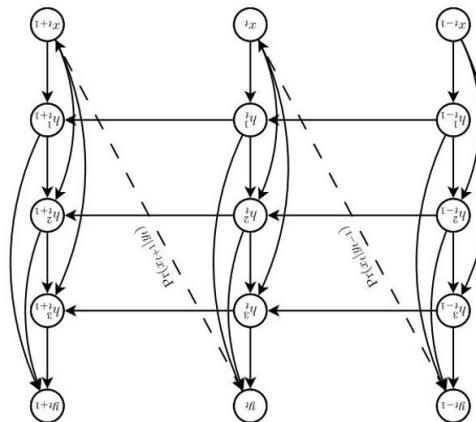
Introduction



Handwriting synthesis

Input:

« a thing of beauty is joy forever »



Output: a thing of beauty is joy forever

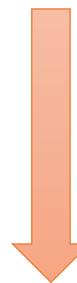
Goal



1) Goal

Design a writing smoother

For a moment, he didn't recognise him.



For a moment, he didn't recognise him.

Shakespeare texts generation

the c-on to the many.
coriolanus
that i c-w me say their deserves
coriolanus
why your worthies the comes, there's when
for people by honans best coriolanus?
cen-cinize
i will by him ow

Contents

- 1) Prediction models
- 2) Sampling
 - a) Unbiased
 - b) Biased
 - c) Primed
- 3) Problem solving
 - a) Writing smoothenner
 - b) Shakespeare texts generation

Conclusion

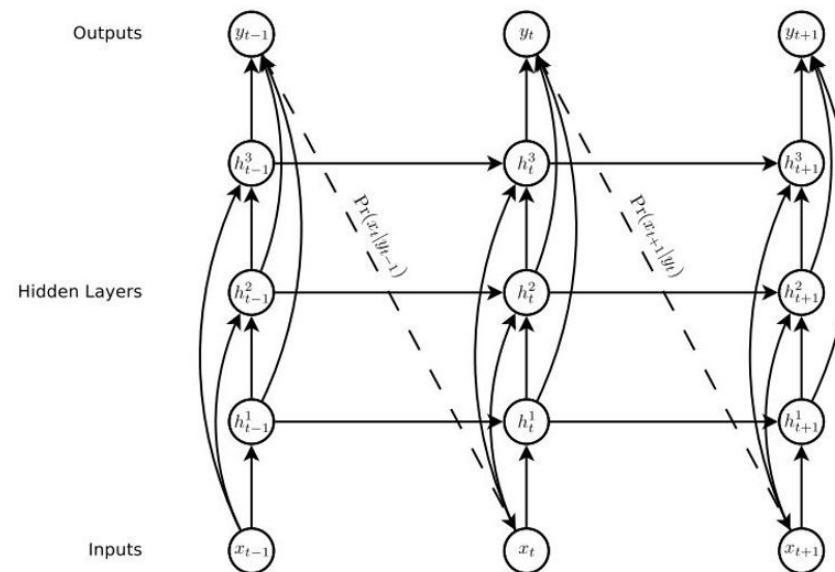


1) Prediction models



1) Prediction models

Unconditional model



$$\hat{y}_t = \left(\hat{e}_t, \{\hat{w}_t^j, \hat{\mu}_t^j, \hat{\sigma}_t^j, \hat{\rho}_t^j\}_{j=1}^M \right) = b_y + \sum_{n=1}^N W_{h^n y} h_t^n$$

$$e_t = \frac{1}{1 + \exp(\hat{e}_t)} \implies e_t \in (0, 1)$$

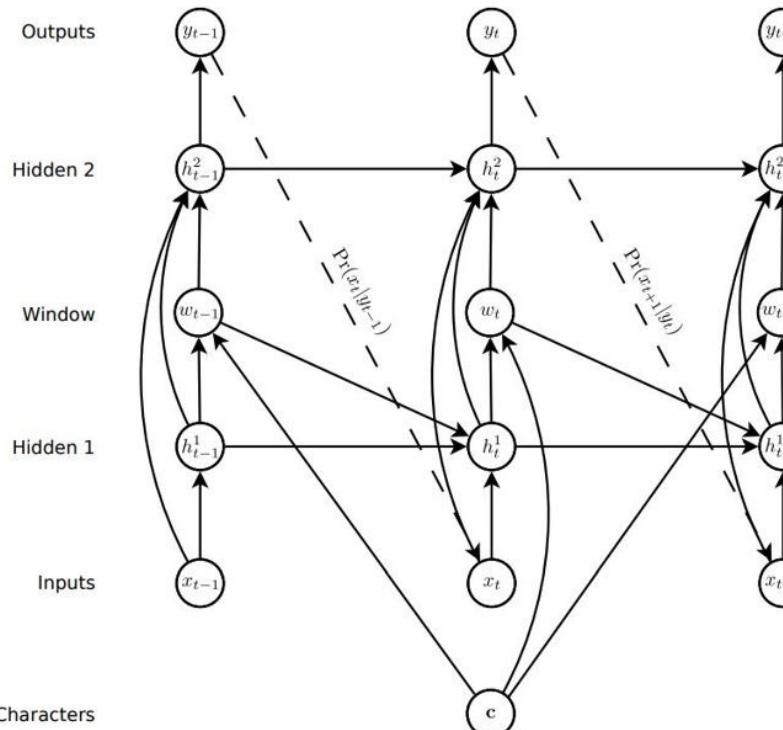
$$\pi_t^j = \frac{\exp(\hat{\pi}_t^j)}{\sum_{j'=1}^M \exp(\hat{\pi}_t^{j'})} \implies \pi_t^j \in (0, 1), \quad \sum_j \pi_t^j = 1$$

$$\mu_t^j = \hat{\mu}_t^j \implies \mu_t^j \in \mathbb{R}$$

$$\sigma_t^j = \exp(\hat{\sigma}_t^j) \implies \sigma_t^j > 0$$

$$\rho_t^j = \tanh(\hat{\rho}_t^j) \implies \rho_t^j \in (-1, 1)$$

Conditional model



$$\phi(t, u) = \sum_{k=1}^K \alpha_t^k \exp \left(-\beta_t^k (\kappa_t^k - u)^2 \right)$$

$$w_t = \sum_{u=1}^U \phi(t, u) c_u$$

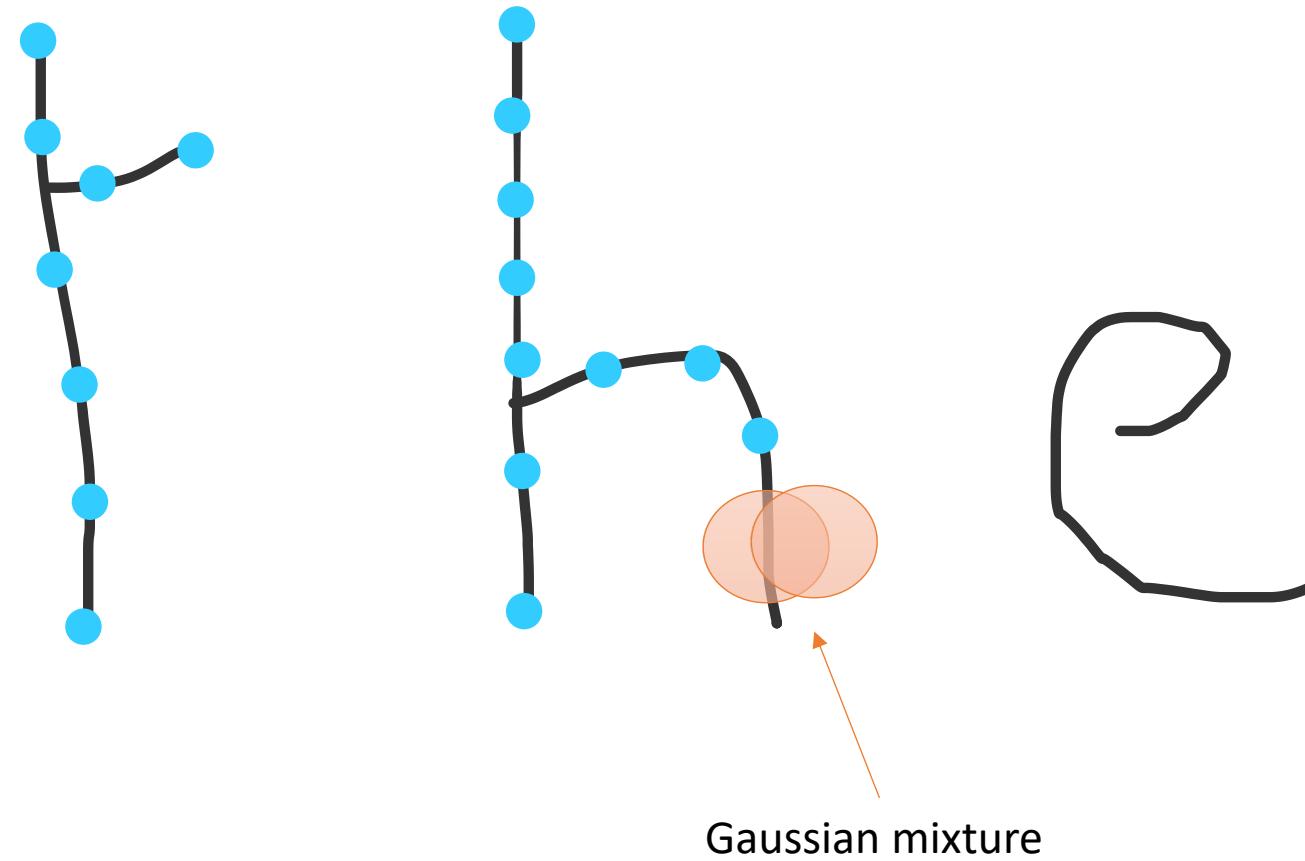
$$(\hat{\alpha}_t, \hat{\beta}_t, \hat{\kappa}_t) = W_{h^1 p} h_t^1 + b_p$$

$$\alpha_t = \exp(\hat{\alpha}_t)$$

$$\beta_t = \exp(\hat{\beta}_t)$$

$$\kappa_t = \kappa_{t-1} + \exp(\hat{\kappa}_t)$$

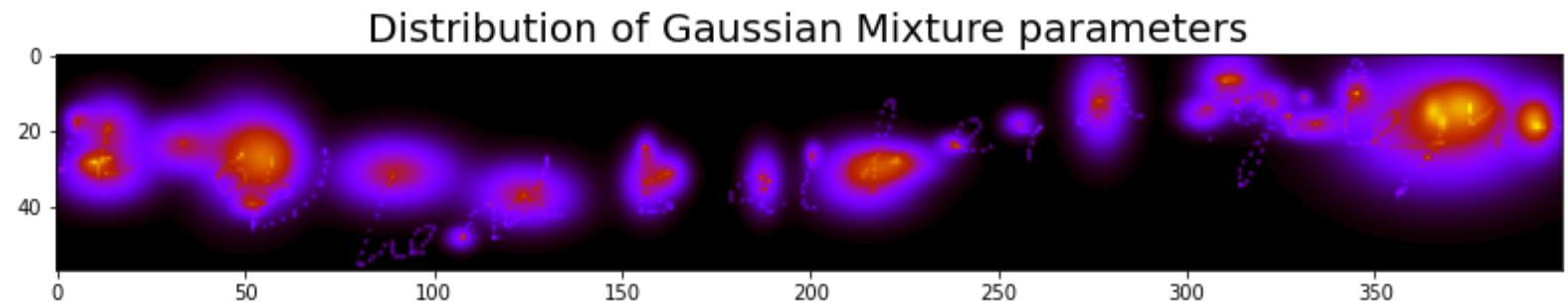
Next point prediction



1) Prediction models

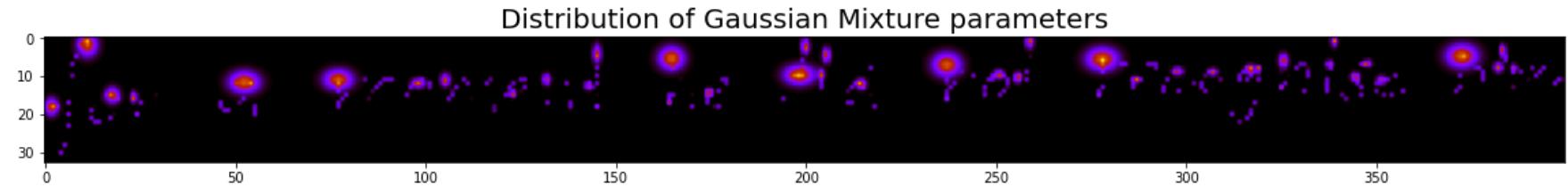
Prediction visualisation

Unconditional model



If o we hənd is al. fər d təske, i-

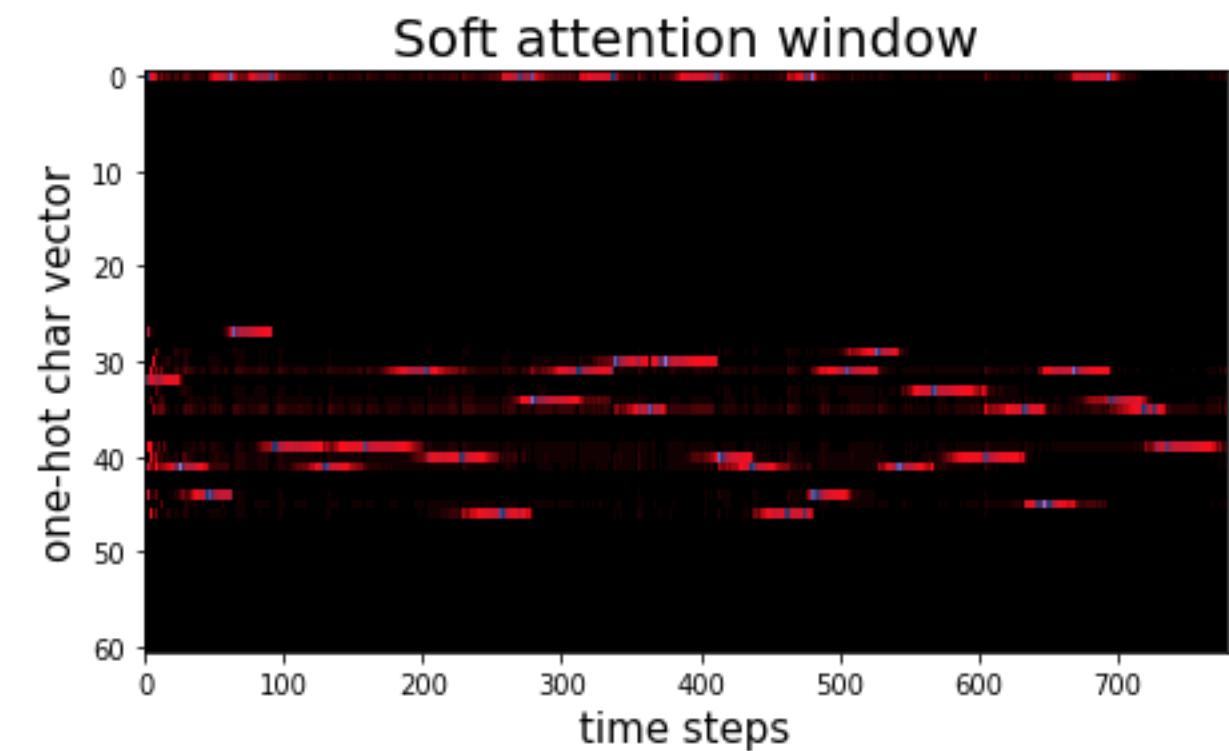
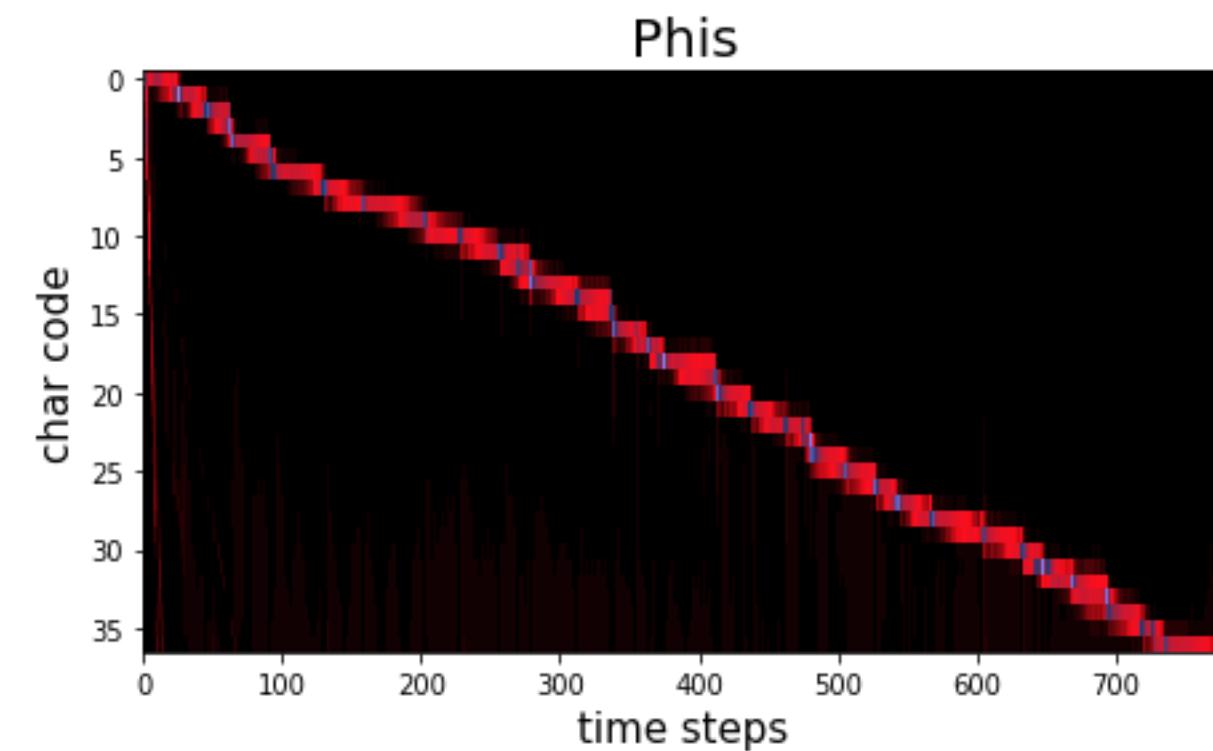
Conditional model



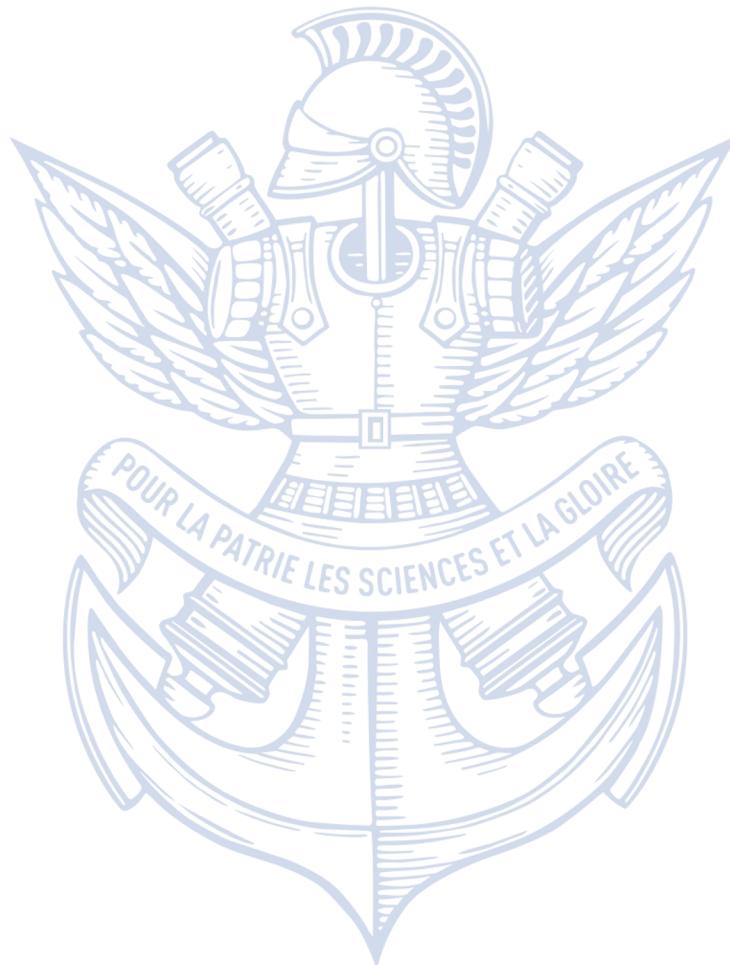
-for a moment he did not recognise him

Attention mechanism for conditional model

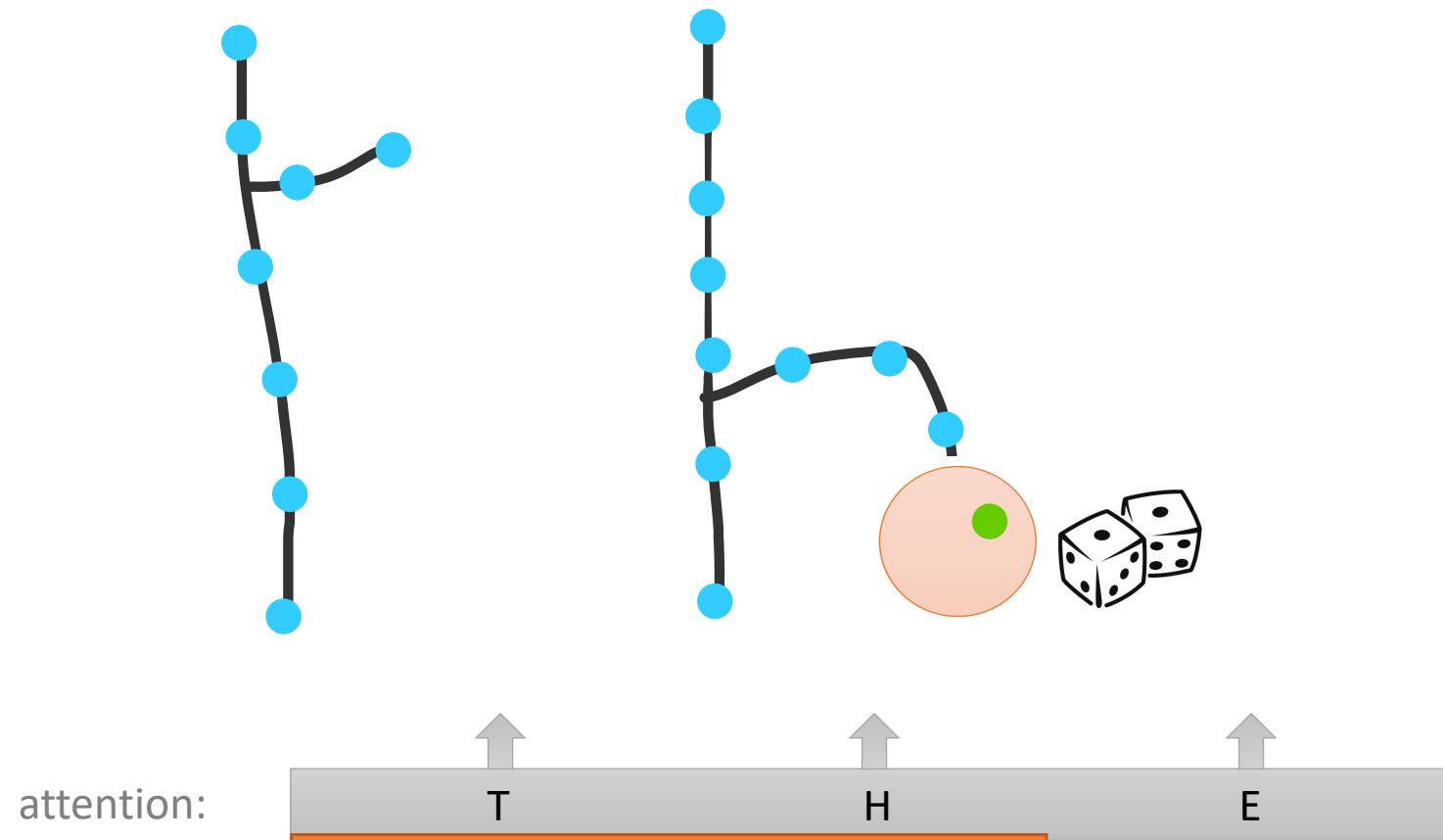
for a moment he did not recognise him



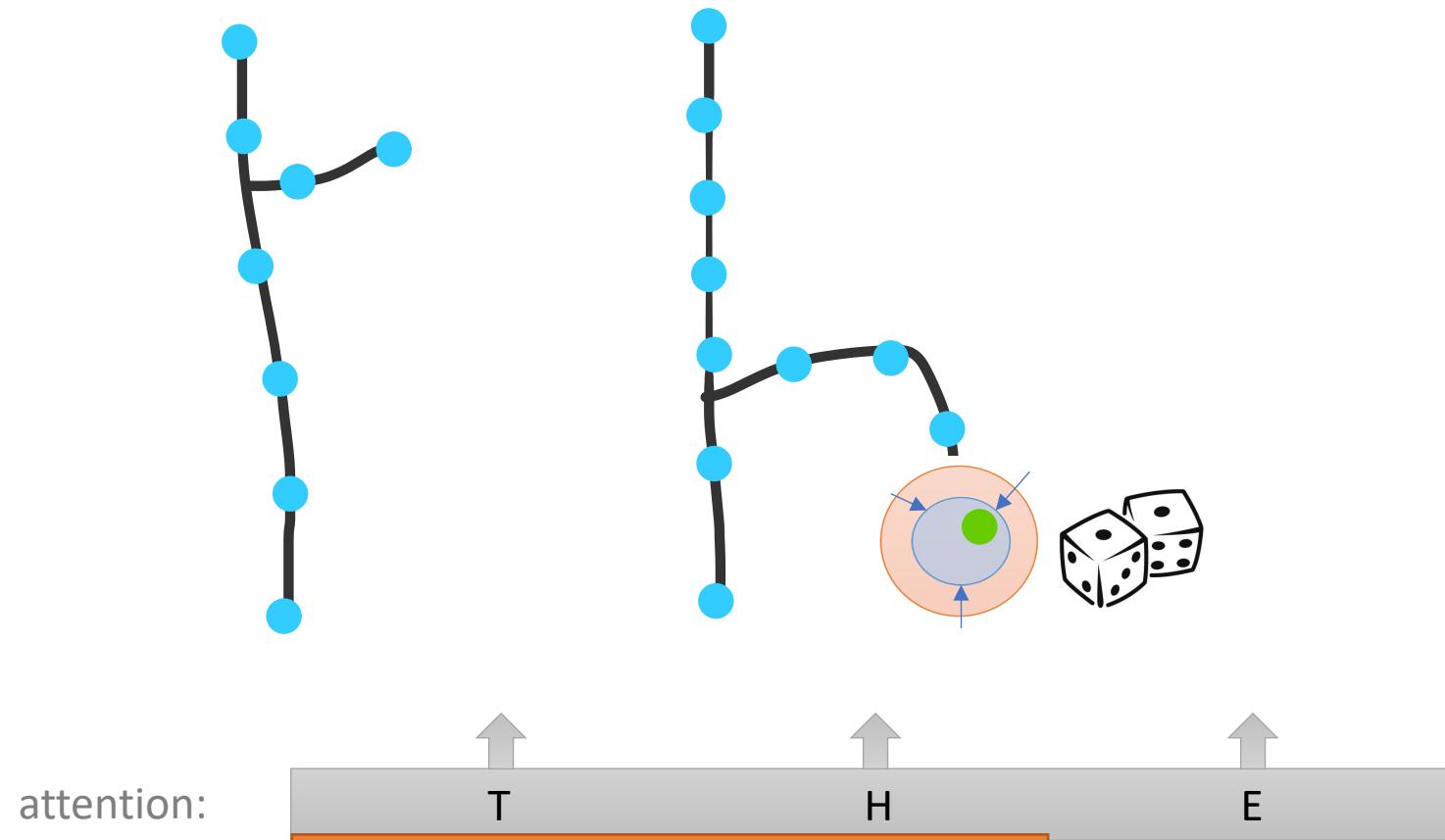
2) Sampling



Unbiased sampling



Biased sampling



Trade-off : Diversity / Cleanliness

b=0

on a dragon he did not recognise him

Unreadable

b=0,4

On a morning he did not recognise him

b=0,8

for a moment he did not recognise him

b=1,2

for a moment he did not recognise him

b=1,6

for a moment he did not recognise him

b=2

for a moment he did not recognise him

b=2,4

for a moment he did not recognise him

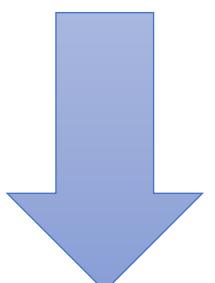
b=2,8

for a moment he did not recognise him

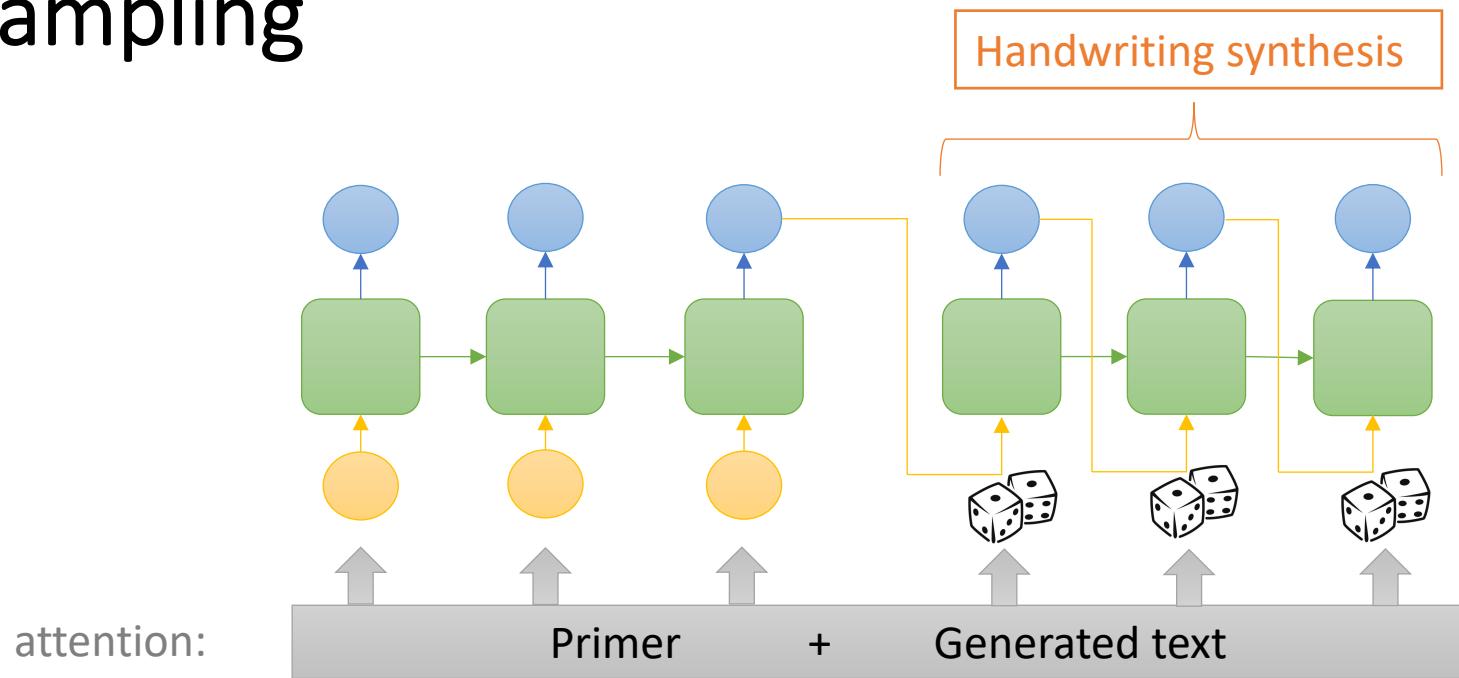
b>2,8

for a moment he did not recognise him

Lack of diversity



Primed sampling



For a moment, he didn't recognise him.

But he too



Ideas for improving primed sampling

- Using a high bias

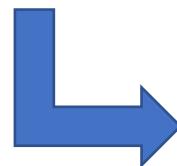


a

Problem at startup:

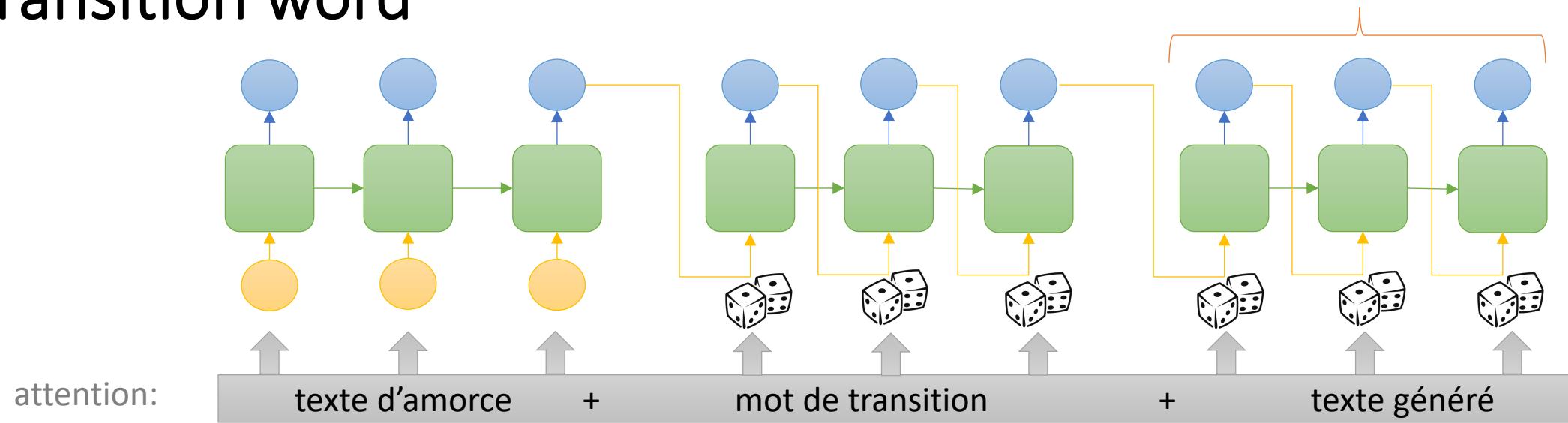
Tw

- Attention manipulation by concentrating it on the first character
- Only keep the hidden state and give only the generated sentence for the condition



in amoment, he didn't recognis-nnn

Transition word



Primer:

For a moment, he didn't recognise him.

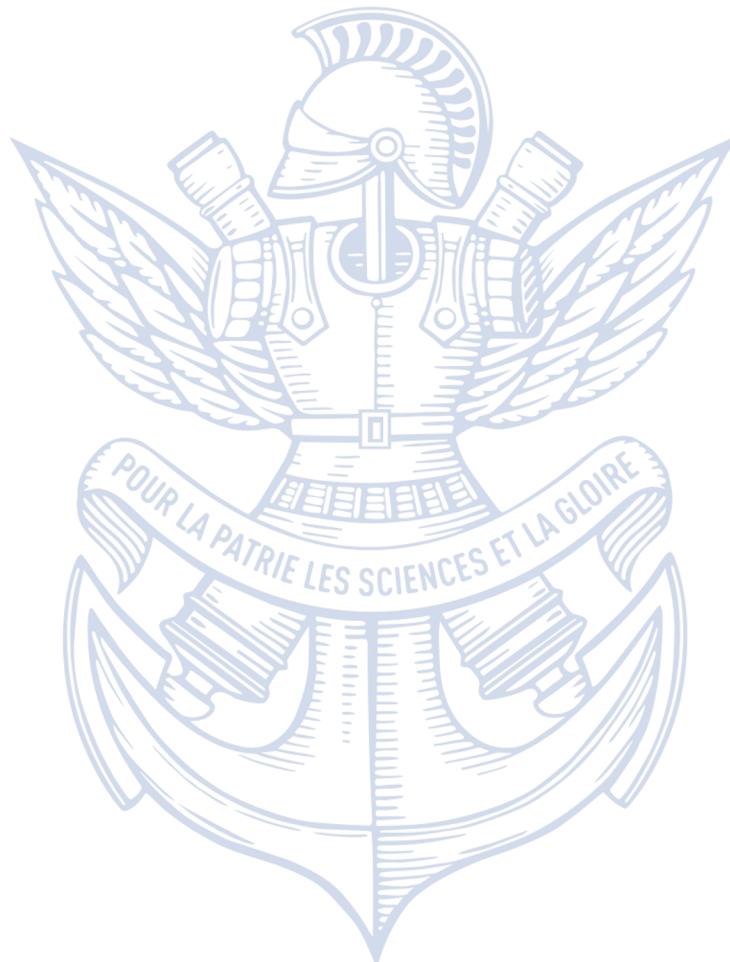
Handwriting synthesis without transition word:

For a moment, he didn't recognise him.

Handwriting synthesis with transition word:

~~trans~~ For a moment, he didn't recognise him.

3) Problems solving



Writing smoother

For a moment, he didn't recognise him.

+ « For a moment, he didn't recognise him »

Generation

Handwriting synthesis model

Sélection

True handwritten input

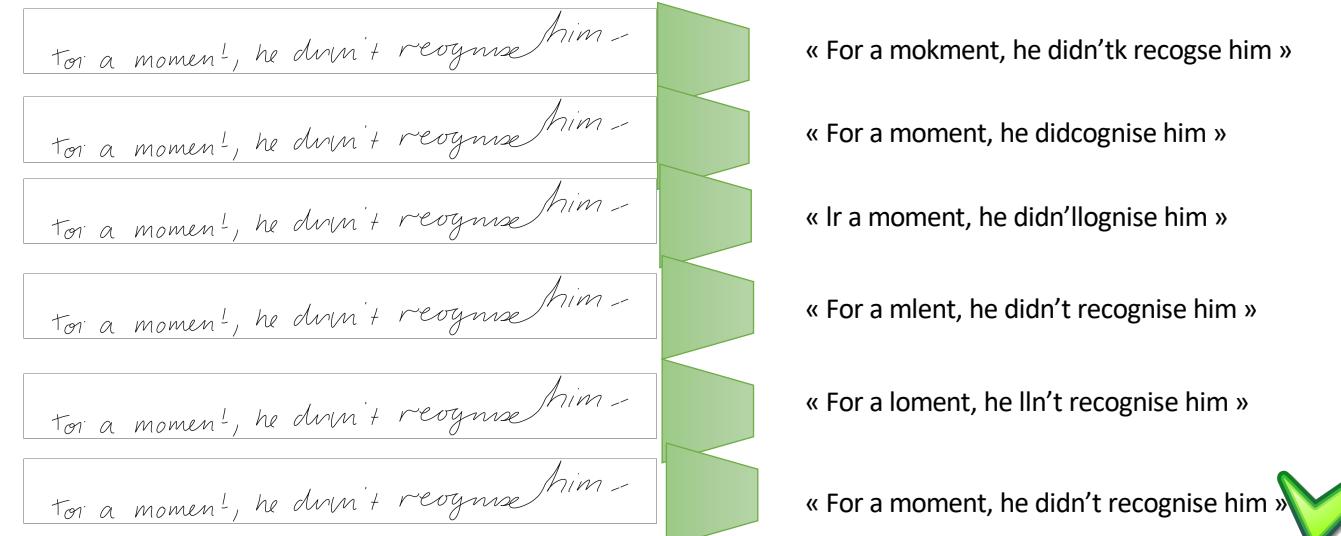
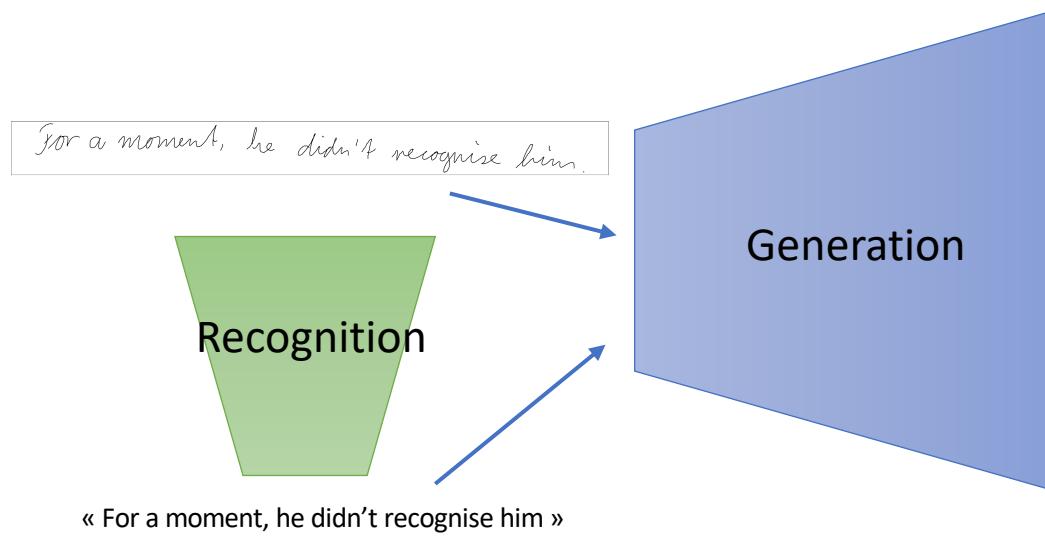
For a moment, he didn't recognise him.

For a moment, he didn't recognise him.

For a moment, he didn't recognise him -

Selection:

Idea 1: use a online handwriting recognition model



Model too complex to be implemented quickly and public code unusable

Fast Multi-language LSTM-based Online Handwriting Recognition
Victor Carbinho · Pedro Gonnet · Thomas Deselaers ·
Henry A. Rowley · Alexander Daryin · Marcos Calvo · Li-Lun Wang ·
Daniel Keysers · Sandro Foug · Philippe Gervais

Received: Nov. 2018 / Accepted: 24-Jan-2020

Abstract: We describe an online handwriting system that is able to support 102 languages using a deep neural network. This is a significant improvement over our previous Segment-and-Decode-based system and reduced the error rate by 20%-40% relative to the state-of-the-art results on IAM-Online. Our new system is based on a sequence-to-sequence architecture and achieves state-of-the-art results on IAM-Online for both the open and closed domains. The main idea is to use a sequence-to-sequence model trained on sequence recognition with a new input encoding using Bezier curves. This leads to up to 10x faster recognition than our previous system. In this paper we also present a series of experiments to determine the optimal configuration of our models and report the results of our work on a number of additional public datasets.

1 Introduction

In this paper we discuss online handwriting recognition. The goal is to take in the form of an ink, i.e. a list of touch or pen strokes, output the textual interpretation of this input. A stroke is a sequence of points (x, y, t) where x and y are the coordinates and t is the time. Figure 1 illustrates example inputs to our online handwriting recognition system in different languages and writing. The first column shows examples in English with different writing styles, with different types of characters (text, numbers, symbols). The second column shows examples of multiple lines. The center column shows examples from five different alphabetic languages similar in structure to English. The left column shows examples from Georgian. The right column shows scripts that are significantly more complex than English. These scripts have a larger set of more complex characters, and users often overlap characters with one another. Korean, while an abjaditic language, groups letters in syllables leading to a large “alphabet” of syllables. Hand writing often consists of a writing “Shivershka” line and characters are often written in a single stroke. Some characters influence the written shape of the components. Arabic is written right-to-left (with embedded left-to-right accents) and some characters change shape depending on their position within a word. Finally are non-text Unicode symbols that we will also consider.

Online handwriting recognition has recently been gaining importance for multiple reasons: (a) An increasing number of mobile devices have touchscreens, giving access to computing devices, many exclusively using mobile devices with touchscreens. Many of these users are non-native English speakers who may not easily type as English, e.g. due to the size of the alphabet and the lack of a standard keyboard layout. (b) More and more large mobile devices with styluses are becoming



Fig. 1 Example inputs for online handwriting recognition in different languages. See text for details.

abjaditic language, group letters in syllables leading to a large “alphabet” of syllables. Hand writing often consists of a writing “Shivershka” line and characters are often written in a single stroke. Some characters influence the written shape of the components. Arabic is written right-to-left (with embedded left-to-right accents) and some characters change shape depending on their position within a word. Finally are non-text Unicode symbols that we will also consider.

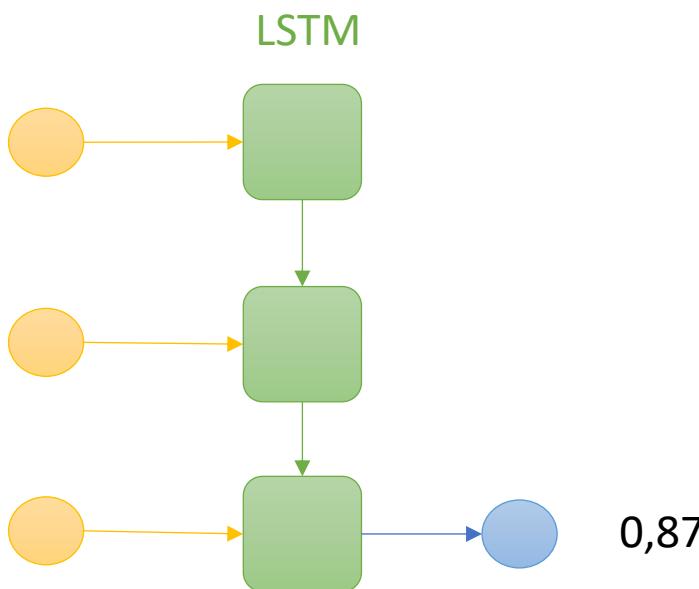
Online handwriting recognition has recently been gaining importance for multiple reasons: (a) An increasing number of mobile devices have touchscreens, giving access to computing devices, many exclusively using mobile devices with touchscreens. Many of these users are non-native English speakers who may not easily type as English, e.g. due to the size of the alphabet and the lack of a standard keyboard layout [10]. (b) More and more large mobile devices with styluses are becoming

All authors are with Google AI Research.
Email: vcarbinho@google.com, pgonnet@google.com,
tdeselaers@google.com, hrowley@google.com,
adaryin@google.com, mcalvo@google.com,
dkeysers@google.com, sfoug@google.com,
phgervais@google.com

Selection:

Idea 2: Create a recognition model to assess the handwriting quality

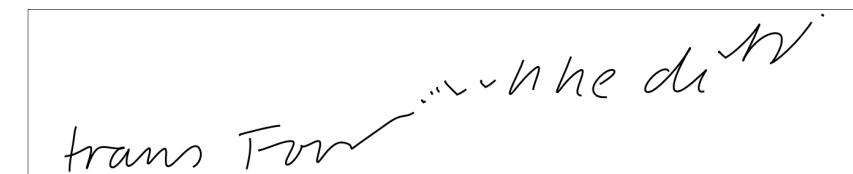
he didn't recognise him



Handmade database:



Boriammonium, he didn't recognise him



Last year saw the defeat of two

0

1

0

1

Selection:

Idea 3: Evaluating the quality of handwriting with the loss of the unconditional model

Barium ammonium, he didn't recognise him

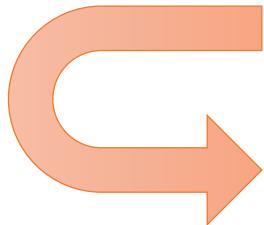
High loss

trans For ammonium
Barium ammonium

Low loss

Result:

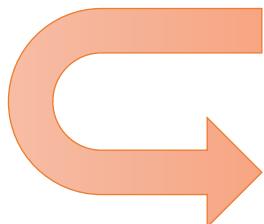
Smoothener



For a moment, he didn't recognise him.

For a moment, he didn't recognise him -

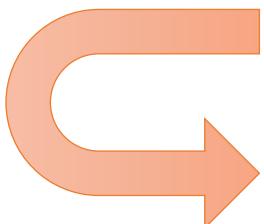
Smoothener



Last year saw the defeat of two

Last year saw the defeat of two

Smoothener

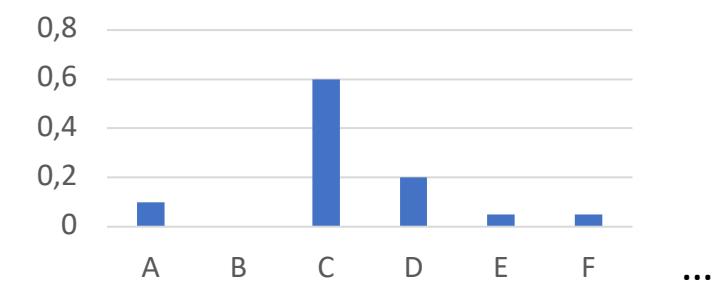
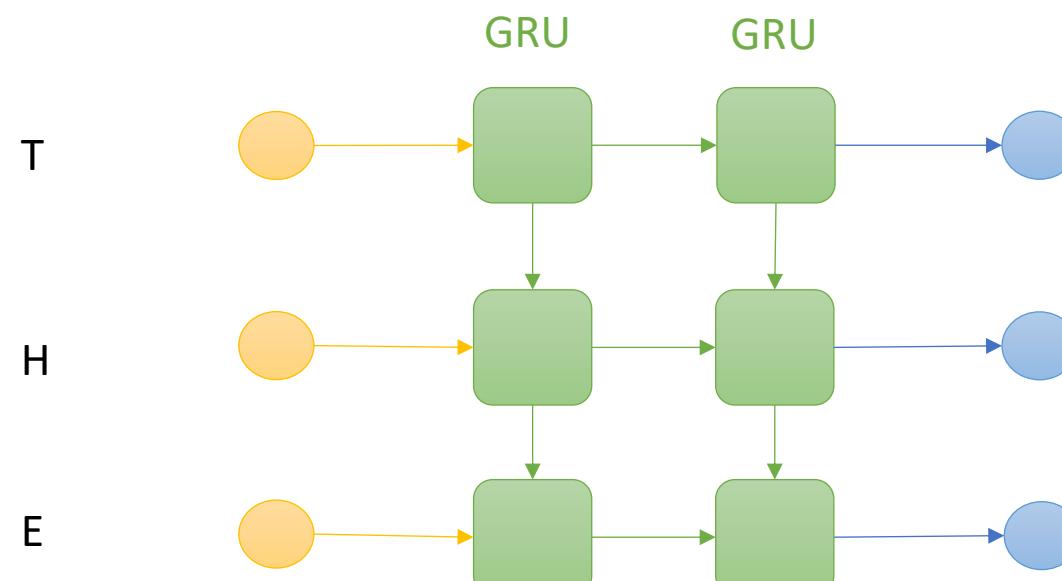


Commonwealth in which five

Commonwealth in which five

Generation of Shakespeare texts:

Architecture:



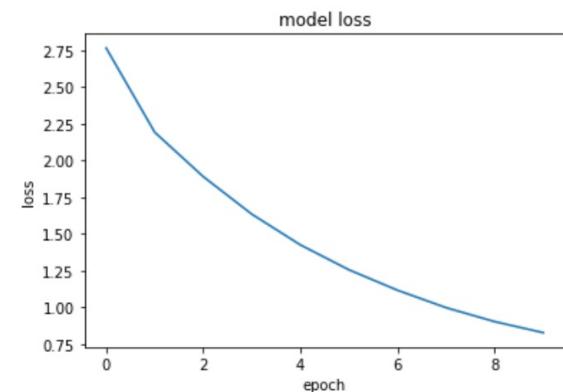
Data : <https://homl.info/shakespeare>

Result

Text generation

```
['the common to the many.',  
'',  
'coriolanus:',  
'that i know me say their deserves.',  
'',  
'coriolanus:',  
"why, your worthies the comes, there's when",  
'your people by honours best coriolanus?',  
'',  
'second citizen:',  
'i will by him ow']
```

Handwriting
generation



the c-on to the many.
coriolanus)
that i our me say their deserves
coriolanus)
why, your worthies the comes, there's when
for people by honans best coriolanus?
cer-cinize
i will by him ow

