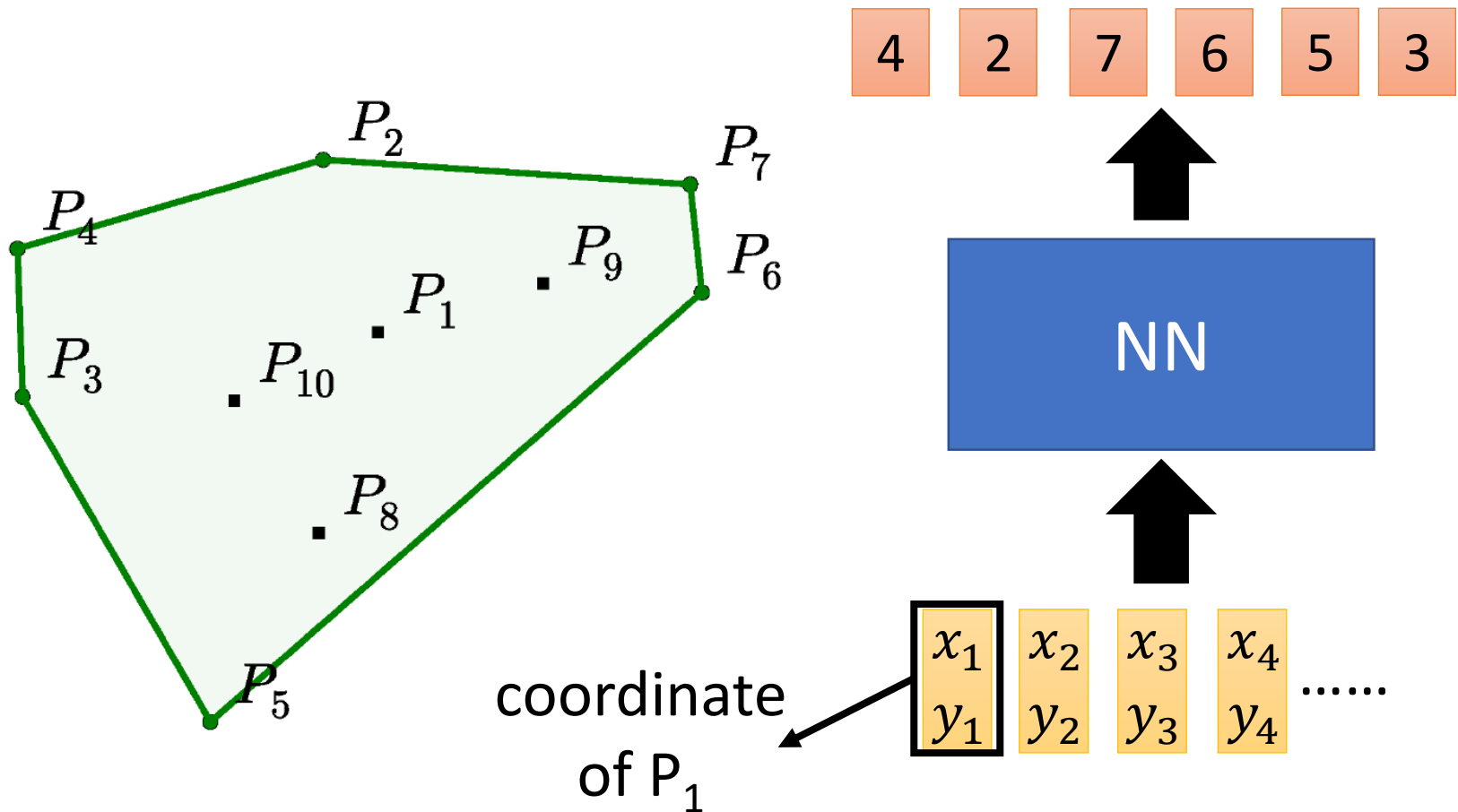# Pointer Network
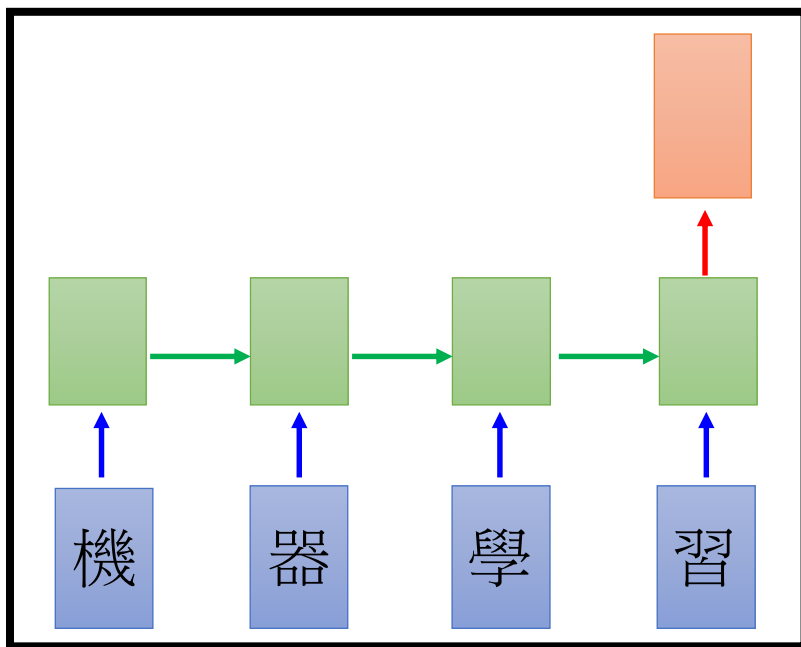
Sequence-to-sequence?

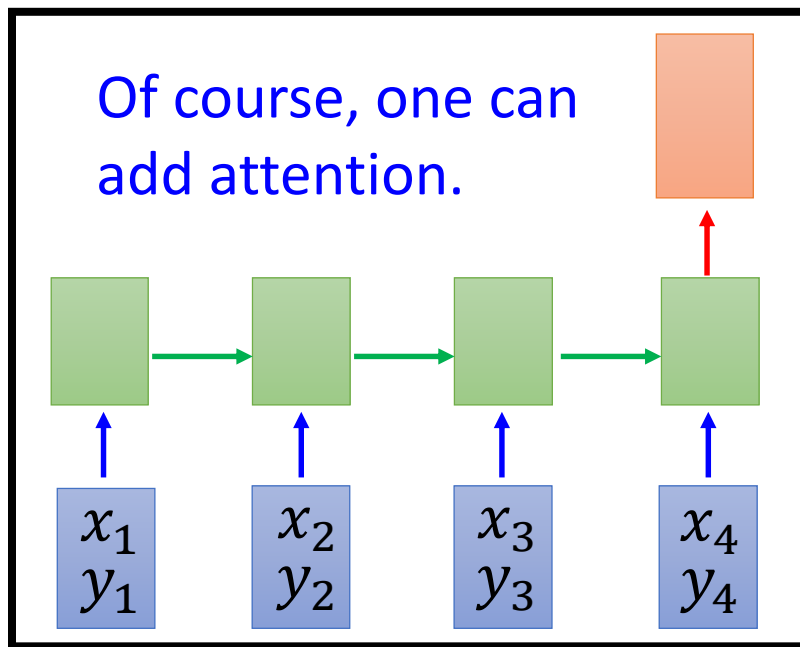**_Encoder_**

Decoder

Sequence-to-sequence?

Problem?

$\{1, 2, 3, 4, \text{END}\}$

Of course, one can add attention.
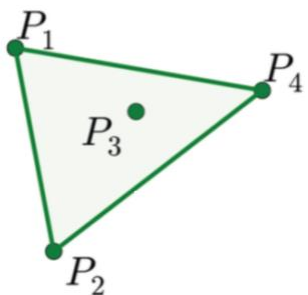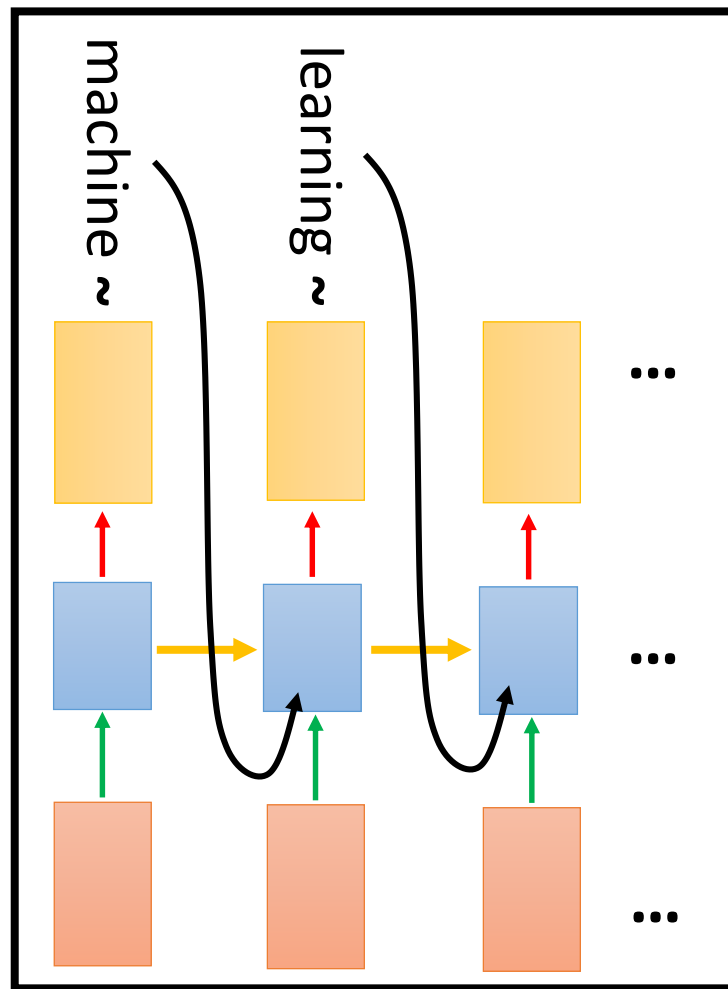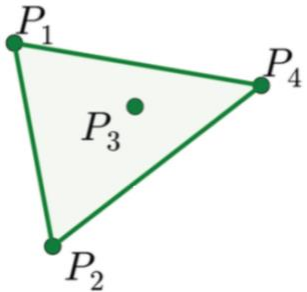
Encoder
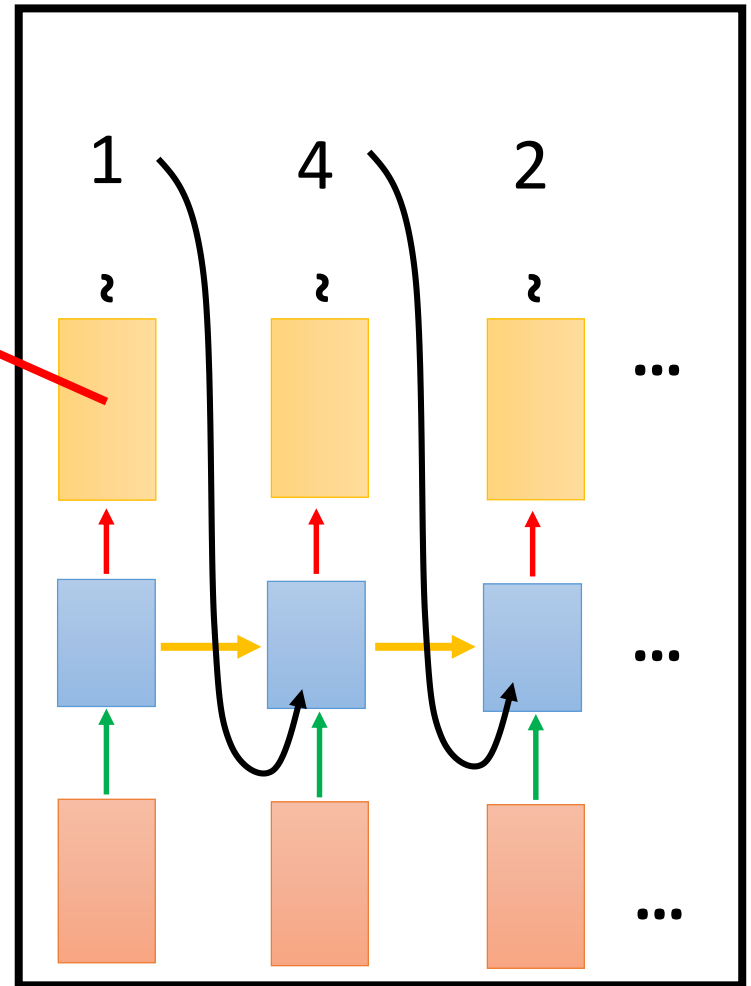
Decoder

Pointer Network

$\begin{matrix} x_0 \\ y_0 \end{matrix}$: END

Attention
Weight

key

0.5 $\longleftarrow$ z^0

$h^0 \rightarrow h^1 \rightarrow h^2 \rightarrow h^3 \rightarrow h^4$

$\begin{matrix} x_0 \\ y_0 \end{matrix}$ $\begin{matrix} x_1 \\ y_1 \end{matrix}$ $\begin{matrix} x_2 \\ y_2 \end{matrix}$ $\begin{matrix} x_3 \\ y_3 \end{matrix}$ $\begin{matrix} x_4 \\ y_4 \end{matrix}$

Pointer Network

$\begin{matrix} x_0 \\ y_0 \end{matrix}$: END

Output: 1

argmax from this distribution

What decoder can output depends on the input.

| 0.0 | 0.5 | 0.3 | 0.2 | 0.0 |

$h^0 \rightarrow h^1 \rightarrow h^2 \rightarrow h^3 \rightarrow h^4$

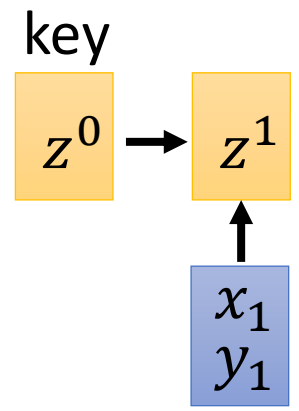$\begin{matrix} x_0 \\ y_0 \end{matrix}$ $\begin{matrix} x_1 \\ y_1 \end{matrix}$ $\begin{matrix} x_2 \\ y_2 \end{matrix}$ $\begin{matrix} x_3 \\ y_3 \end{matrix}$ $\begin{matrix} x_4 \\ y_4 \end{matrix}$

key

$z^0 \rightarrow z^1$

$\begin{matrix} x_1 \\ y_1 \end{matrix}$
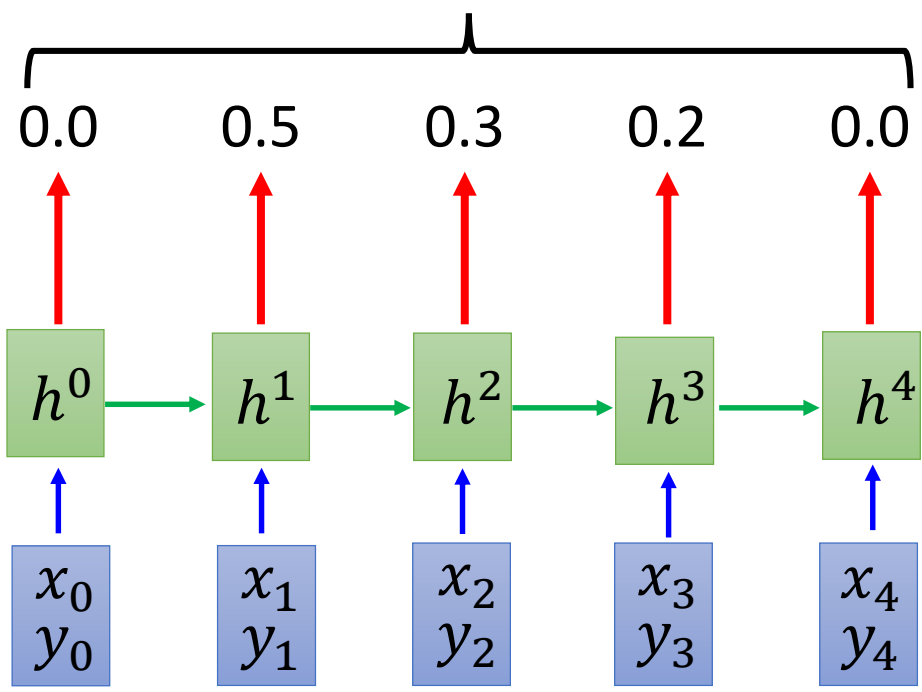
Pointer Network

x_0
y_0 : END
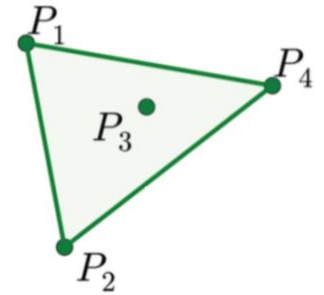
Output: 4

$\approx$

argmax from this distribution

| 0.0 | 0.0 | 0.1 | 0.2 | 0.7 |

$h^0 \rightarrow h^1 \rightarrow h^2 \rightarrow h^3 \rightarrow h^4$

x_0
y_0

x_1
y_1

x_2
y_2

x_3
y_3

x_4
y_4

What decoder can output depends on the input.

key

$z^0 \rightarrow z^1 \rightarrow z^2$ ......

x_1
y_1

x_4
y_4
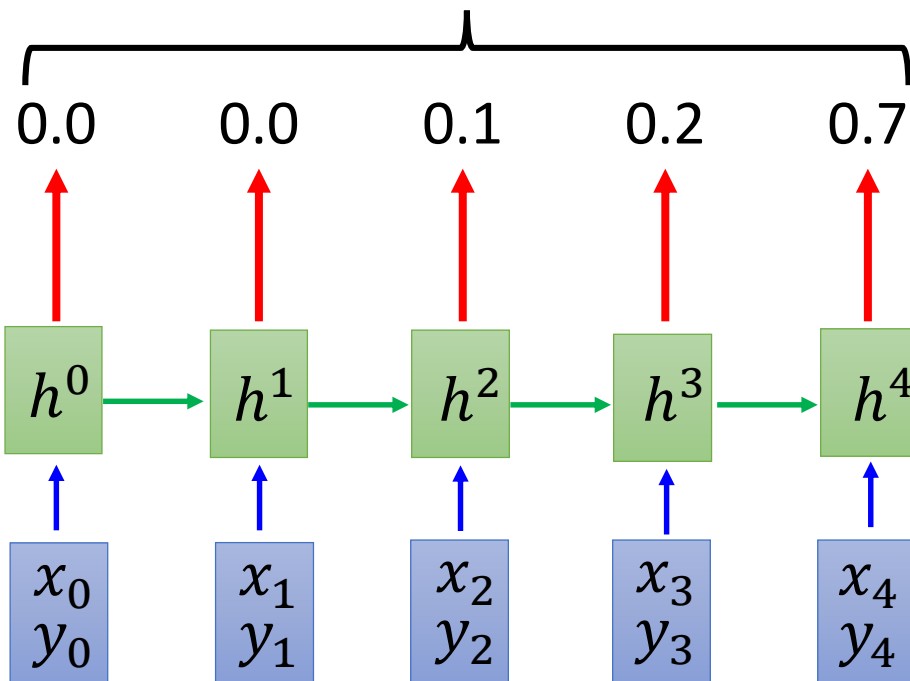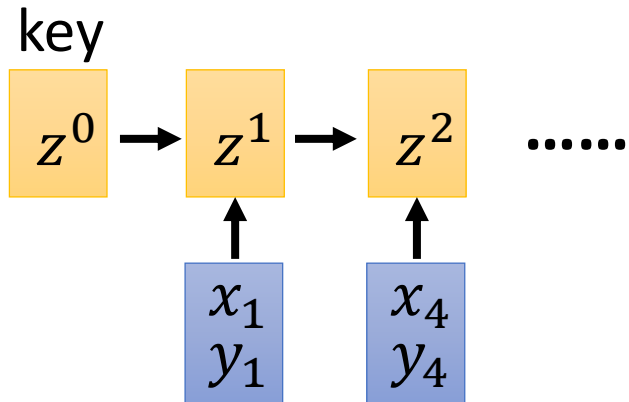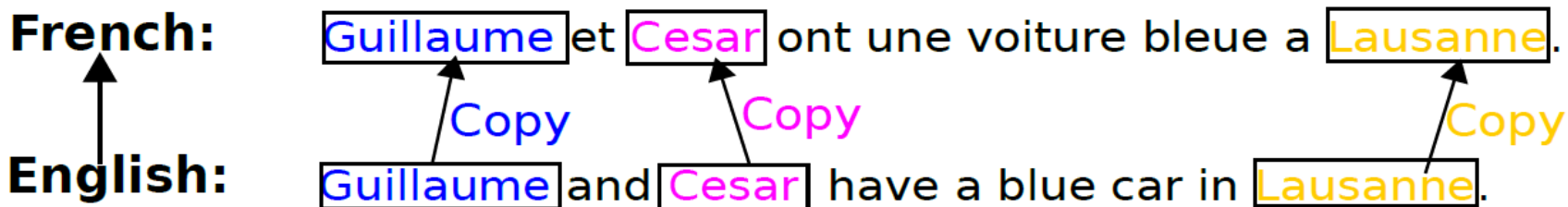
The process stops when "END" has the largest attention weights.

Applications - Summarization
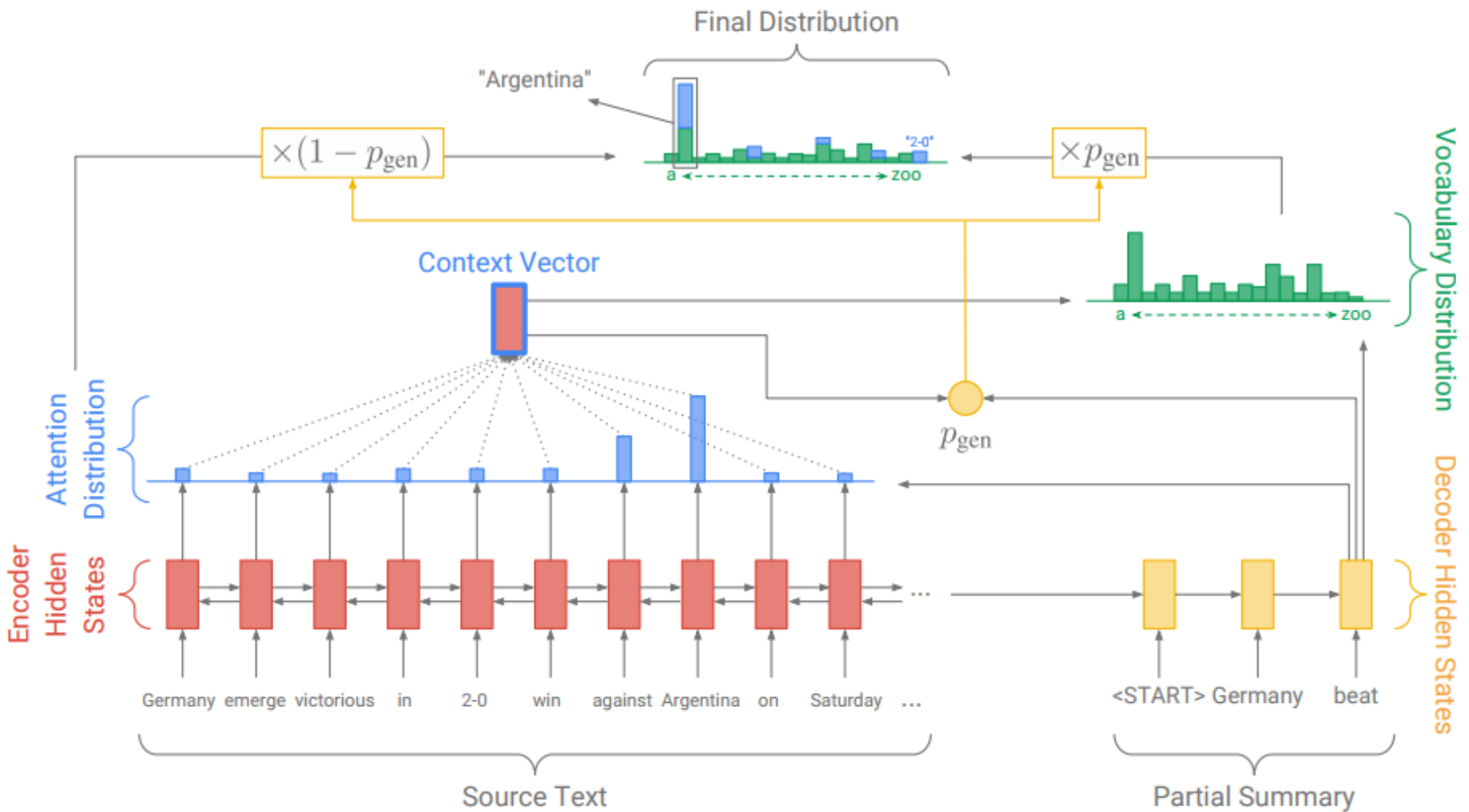
More Applications

Machine Translation

French: Guillaume et Cesar ont une voiture bleue a Lausanne.

Copy — Copy — Copy

English: Guillaume and Cesar have a blue car in Lausanne.

Chat-bot

User: X寶你好，我是庫洛洛

Machine:庫洛洛你好，很高興認識你