

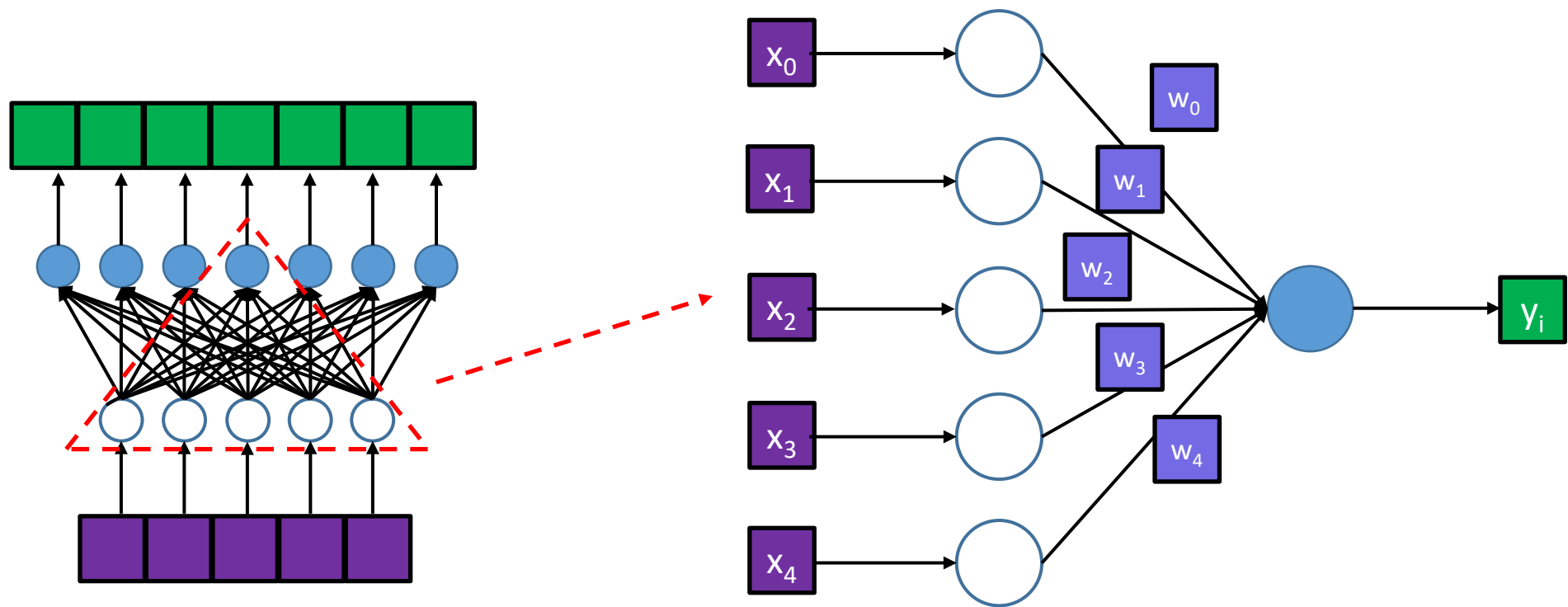


# AttentionとSelf-Attentionの仕組み

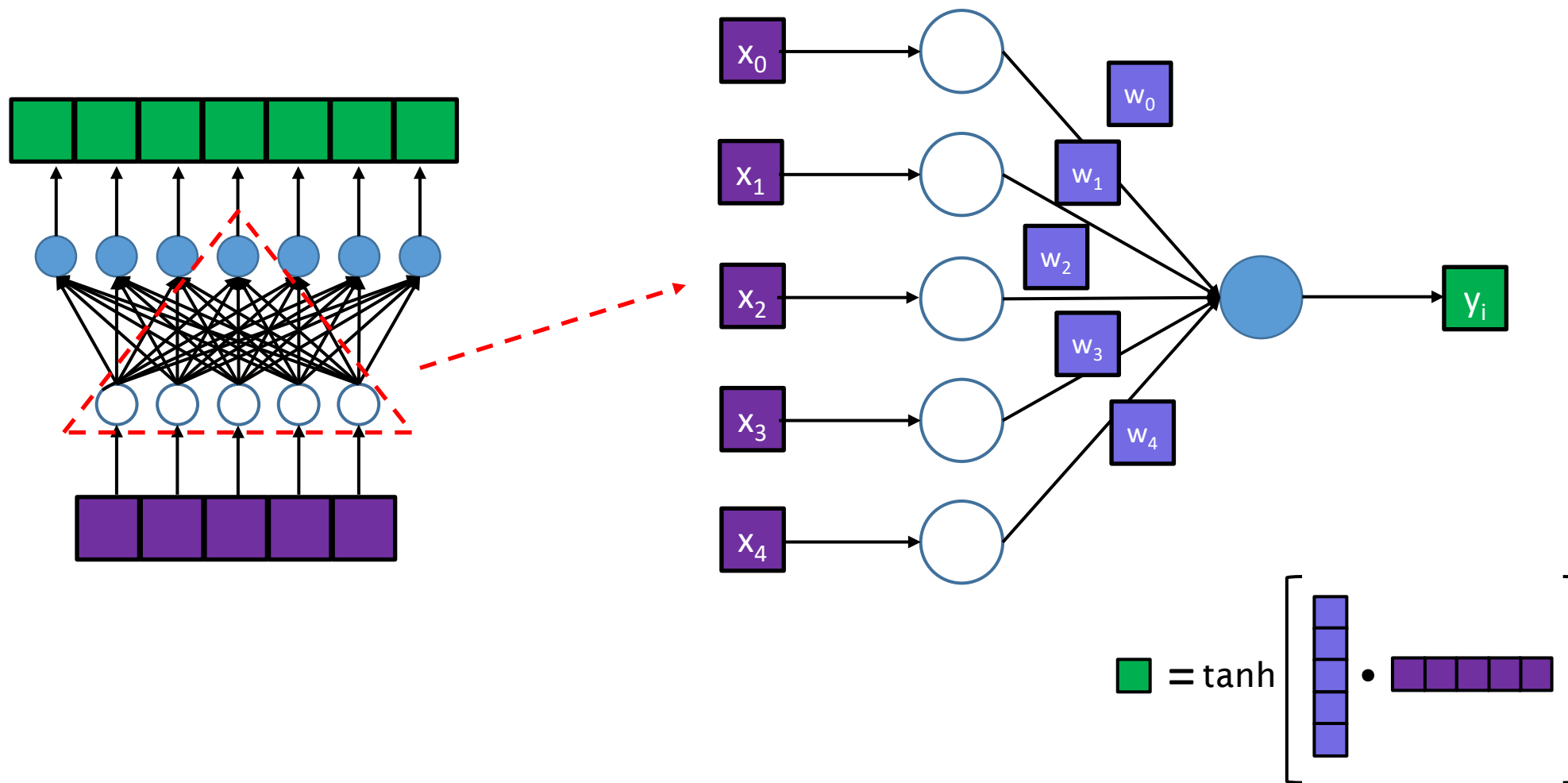
徐 宏坤

2021.04.23

# ニューラルネットワーク(DNN)

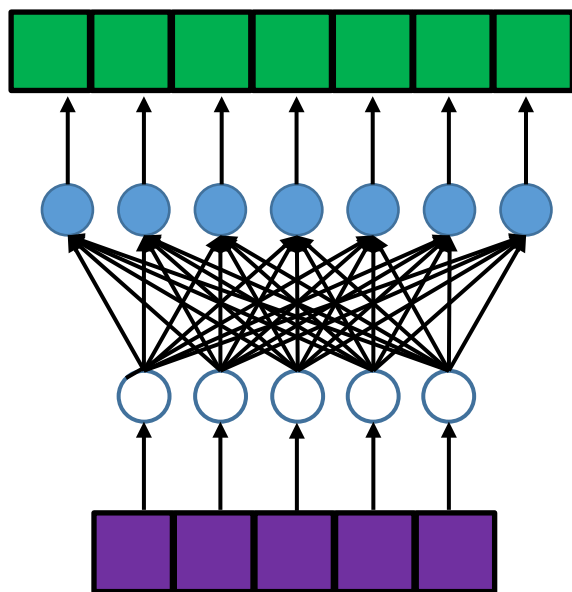


$$y_i = \tanh(x_0w_0 + x_1w_1 + x_2w_2 + x_3w_3 + x_4w_4)$$



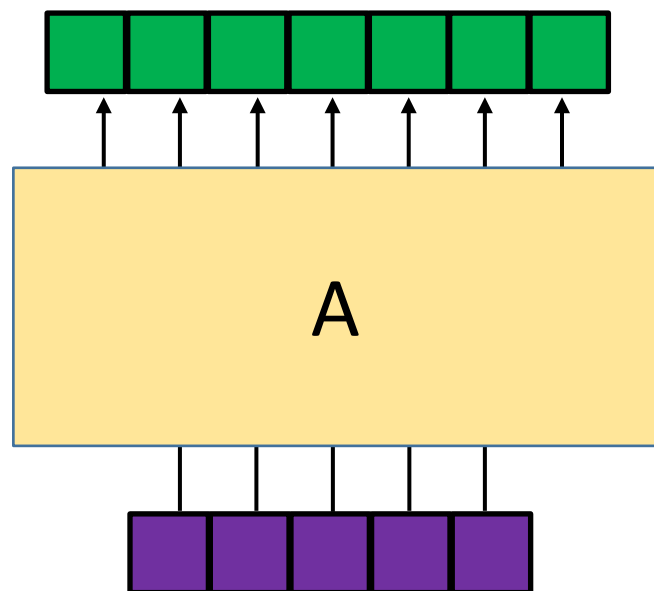


# ニューラルネットワーク(DNN)



$$\begin{bmatrix} \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \end{bmatrix} = \tanh \left[ \begin{bmatrix} \text{blue grid} \\ \text{blue grid} \\ \text{blue grid} \\ \text{blue grid} \\ \text{blue grid} \end{bmatrix} \cdot \begin{bmatrix} \text{purple} \\ \text{purple} \\ \text{purple} \\ \text{purple} \end{bmatrix} \right]$$

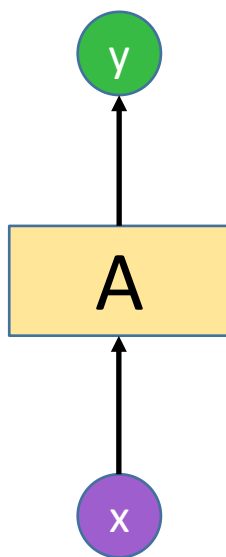
# ニューラルネットワーク(DNN)



$$\begin{bmatrix} \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \end{bmatrix} = \tanh \left[ \begin{bmatrix} \text{blue} & \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} & \text{blue} \end{bmatrix} \cdot \begin{bmatrix} \text{purple} \\ \text{purple} \\ \text{purple} \\ \text{purple} \end{bmatrix} \right]$$

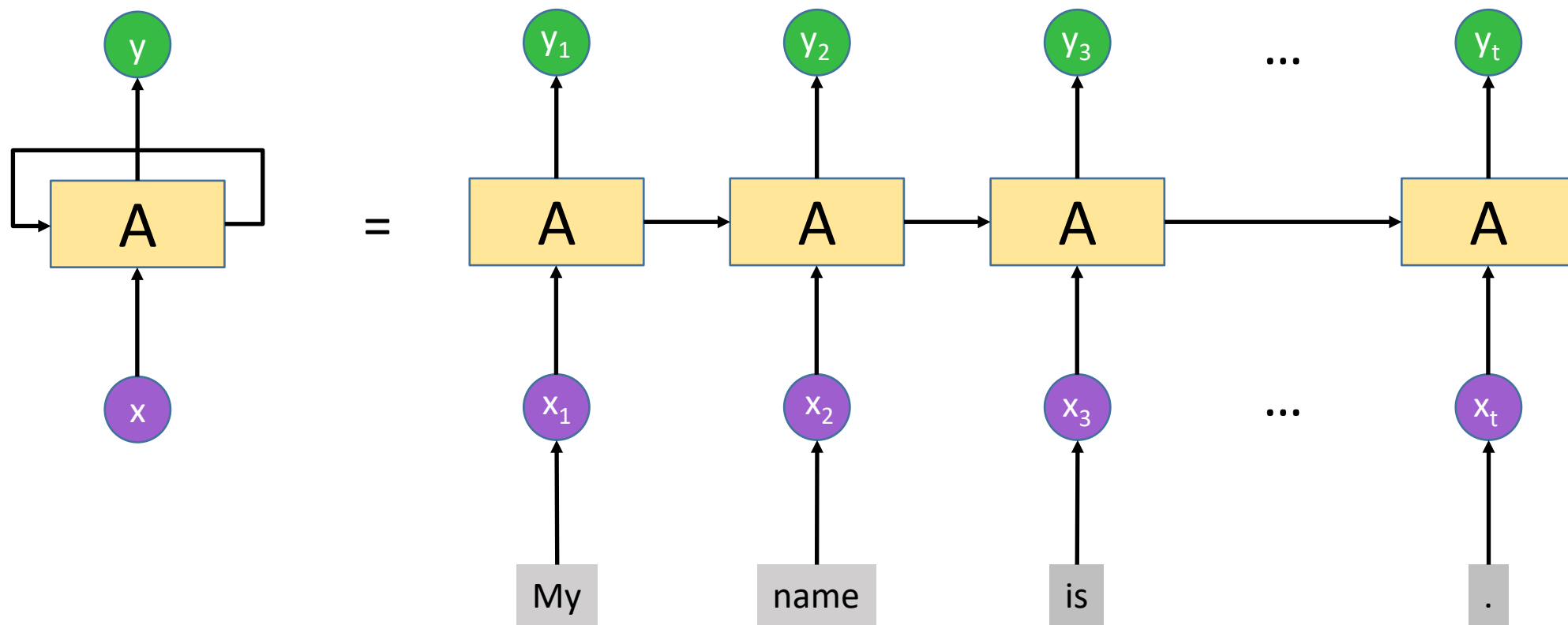


# ニューラルネットワーク (DNN)

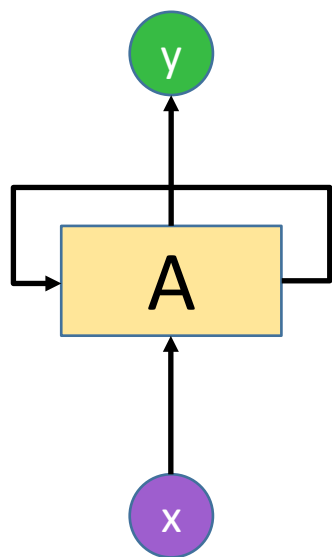


$$\begin{bmatrix} \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \end{bmatrix} = \tanh \left[ \begin{bmatrix} \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} \\ \text{blue} & \text{blue} & \text{blue} & \text{blue} \end{bmatrix} \cdot \begin{bmatrix} \text{purple} \\ \text{purple} \\ \text{purple} \end{bmatrix} \right]$$

# 再帰型ニューラルネットワーク (RNN)



# 再帰型ニューラルネットワーク(RNN)

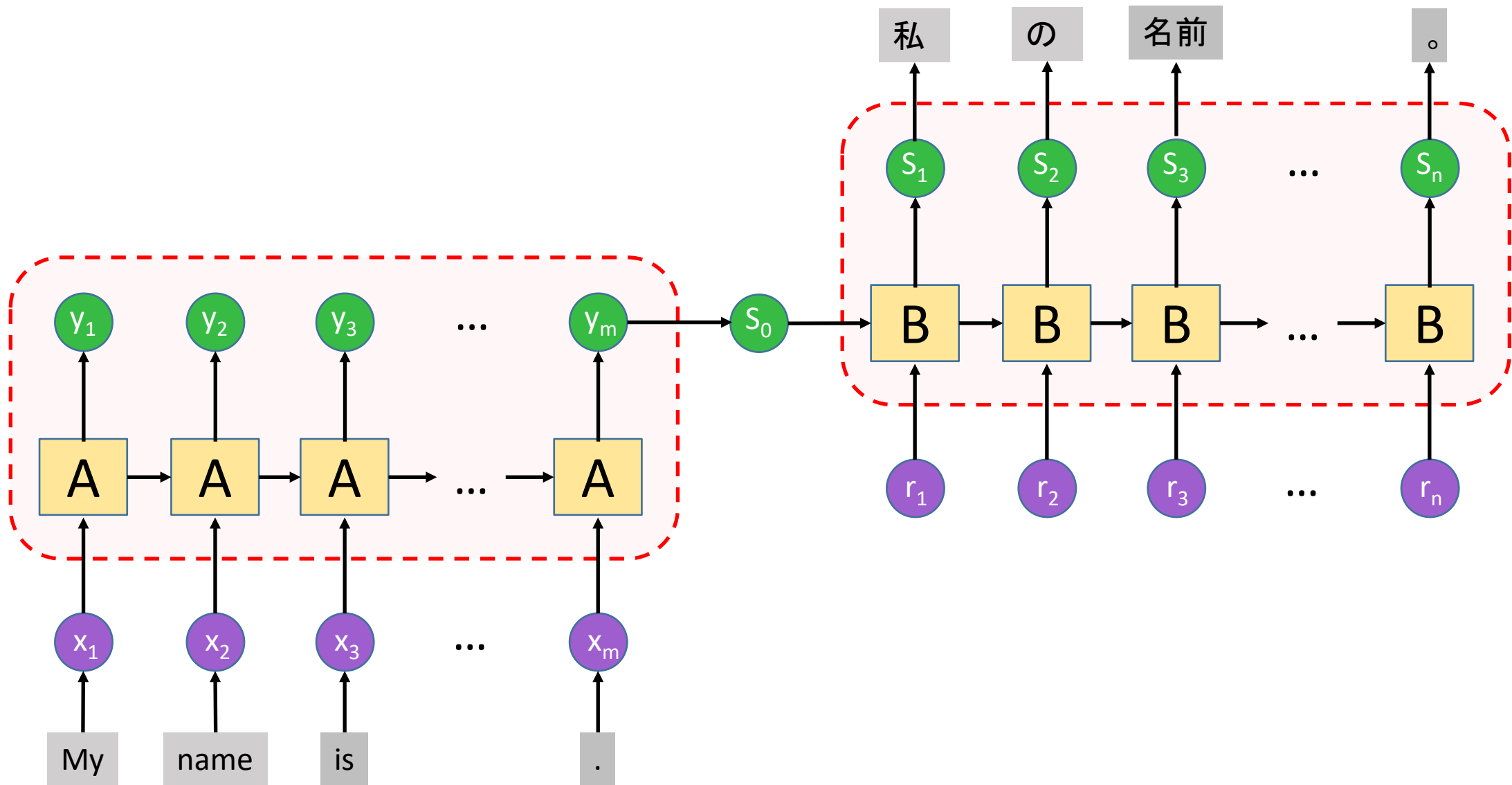


$$\begin{bmatrix} \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \end{bmatrix} = \tanh \left[ \begin{bmatrix} \text{green} & \text{blue} \end{bmatrix} \cdot \begin{bmatrix} \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \\ \text{green} \\ \text{purple} \\ \text{purple} \\ \text{purple} \end{bmatrix} \right]$$

The diagram illustrates the internal computation of the RNN unit. It shows a vertical vector of 5 green squares on the left, followed by an equals sign and the  $\tanh$  activation function. To the right of the  $\tanh$  function is a large bracketed expression. Inside this bracket is a 5x10 grid of squares, where the first 5 columns are green and the next 5 columns are blue. This grid is multiplied (indicated by a dot) by a vertical vector of 10 squares: the top 5 are green and the bottom 5 are purple. The entire expression is enclosed in large square brackets.

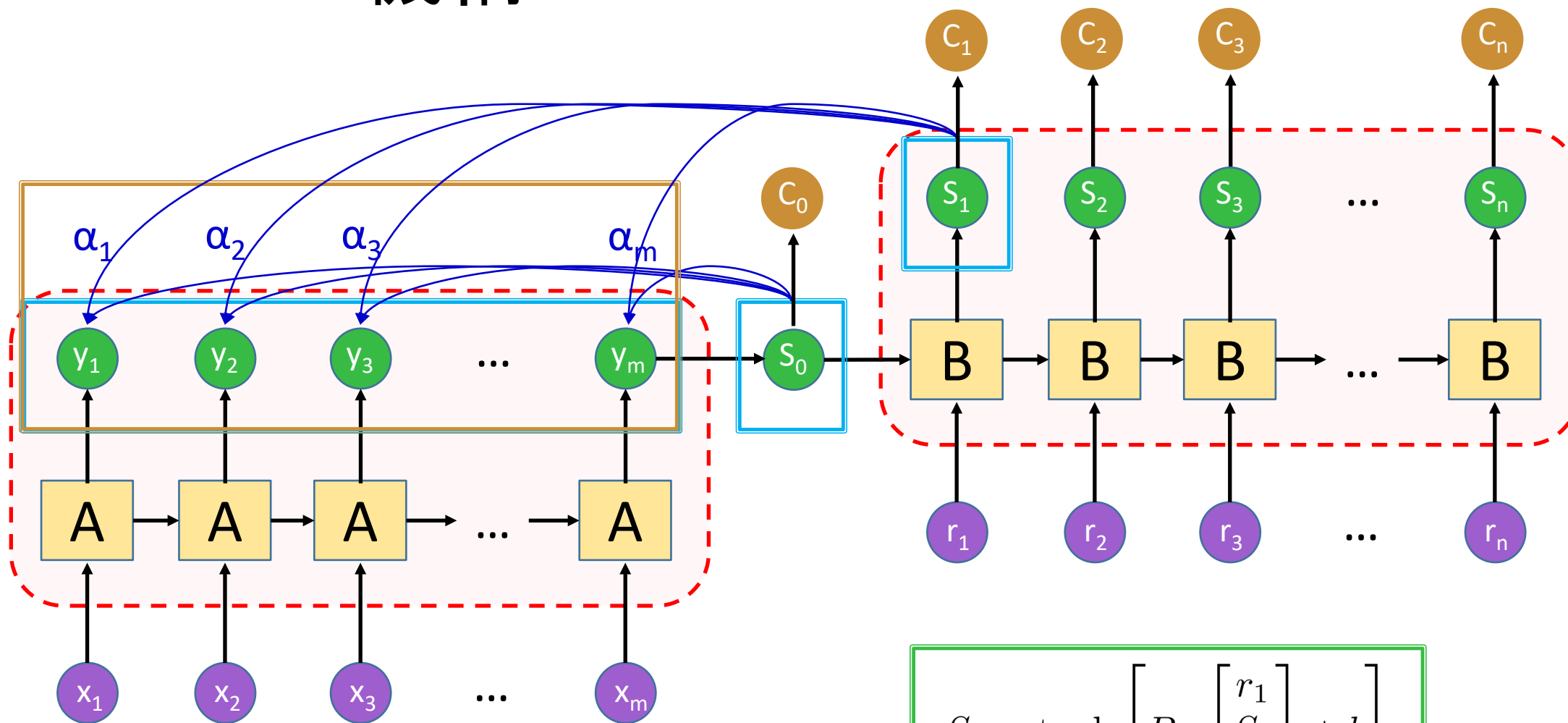


# Seq2Seqモデル(機械翻訳用)



# Attention機構

$$C_0 = \alpha_1 y_1 + \alpha_2 y_2 + \alpha_3 y_3 + \dots + \alpha_m y_m$$

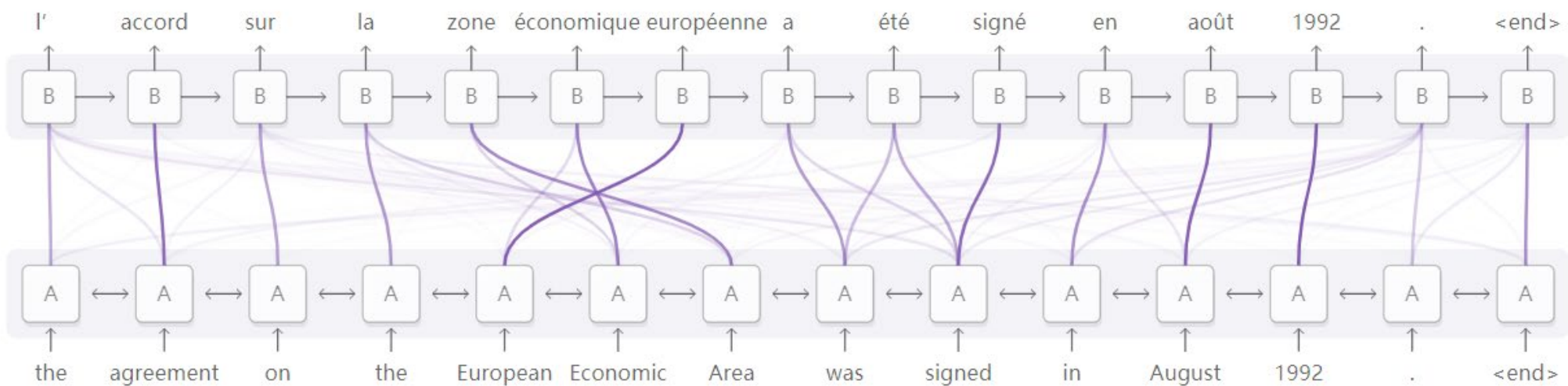


$$S_1 = \tanh \left[ B \cdot \begin{bmatrix} r_1 \\ S_0 \\ C_0 \end{bmatrix} + b \right]$$



# Attention機構

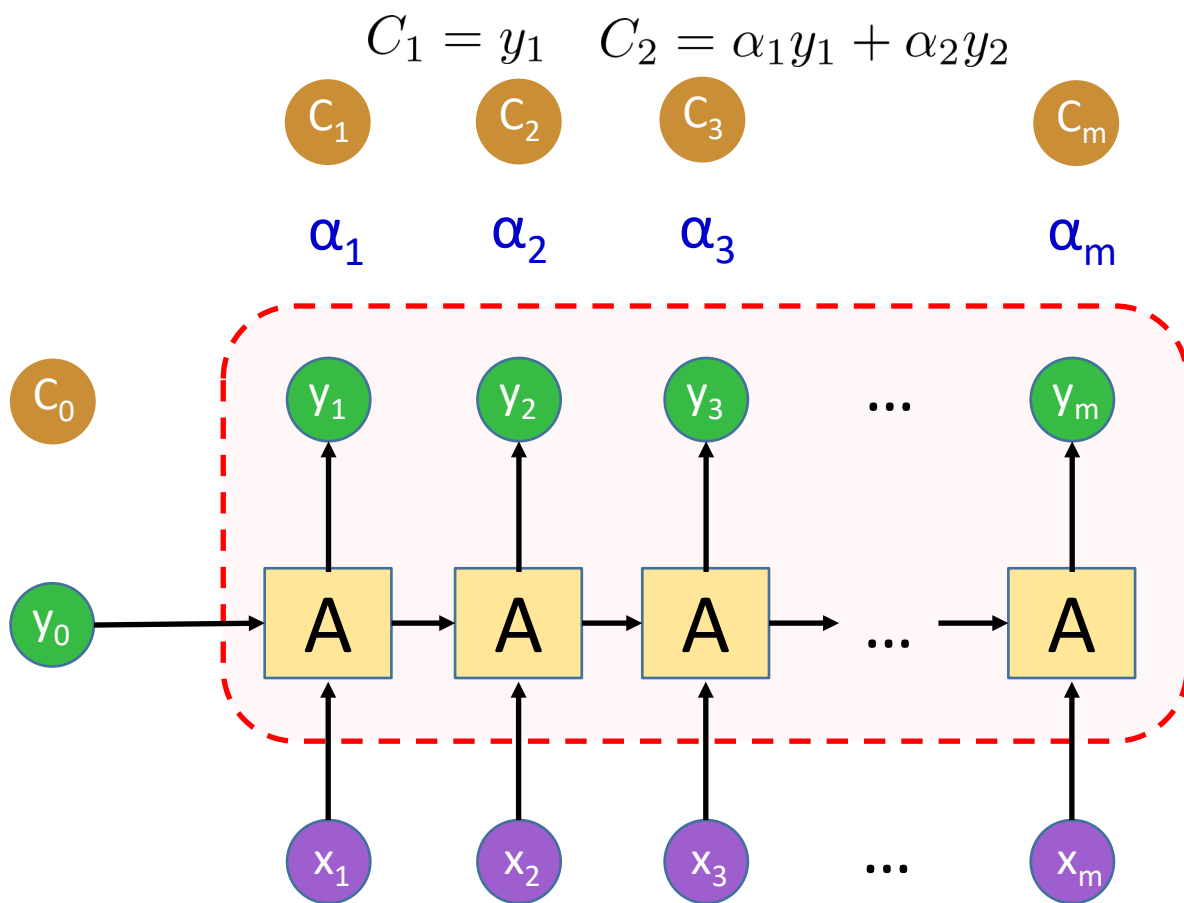
フランス語



英語

# Self-Attention機構

$$y_1 = \tanh \left[ B \cdot \begin{bmatrix} x_1 \\ y_0 \\ C_0 \end{bmatrix} + b \right]$$



ご清聴ありがとうございました