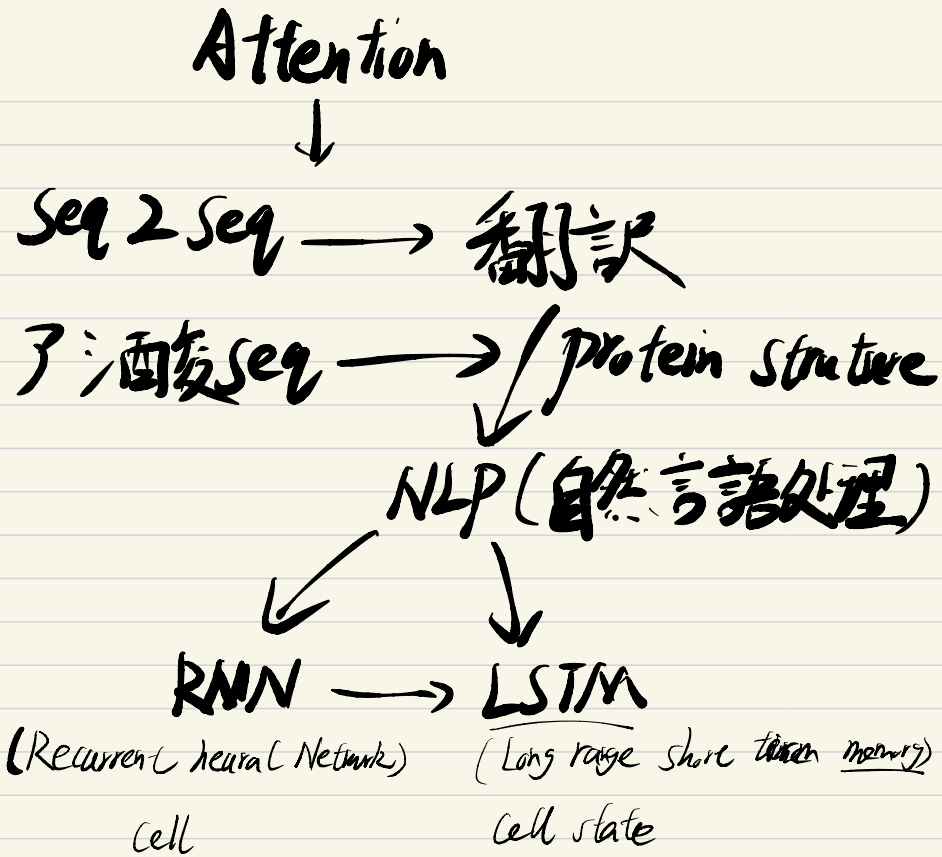
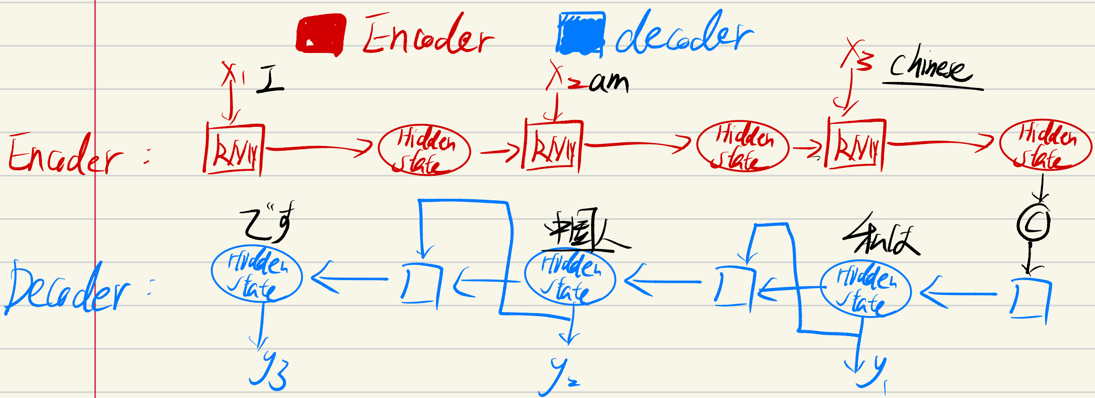
Alphafold 2

Rose TTAfold 2


Attention
seq2seq → 翻訳
アミノ酸seq → protein structure
NLP(自然言語処理)
RNN → LSTM
(Recurrent Neural Network)
(Long range short term memory)
cell
cell state

Seq 2 Seq RNN



f( ) Decoder 関数 (非线性变换)

$y_1 = f(c)$

$y_2 = f(c, y_1)$

$y_3 = f(c, y_1, y_2)$

Attention 注目方

中国人 { I - 0.15, am - 0.05, chinese - 0.8

時刻t、y_i を出力する確率

$p(y_i | y_1, y_2 \ldots y_{i-1}) = g(y_{i-1}, s_i, c_i)$

s_i → decoder hidden state

c_i → Attention weight Matrix

$s_i = f(s_{i-1}, y_{i-1}, c_i)$

$c_i = \sum_{j=1}^{T} a_{ij} h_j$ (Attention vector)

α_{ij}: y_i と x_j の Attention weight

$\alpha_{ij} = \mathrm{Softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k=1}^{T} \exp(e_{ik})}$

e_{ij} = x_i と y_j の alignment 値

$e_{ij} = a(s_{i-1}, h_j)$

Self-attention

Google — <<Attention is all you need>>

↓

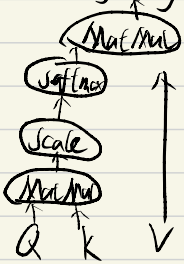
Transformer — self-attention instead of conv.

self attention — 入为=出为

Q, K, V

↓ ↓ ↓

Query key value.



$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V$$

	Think	Machine
initial embedding	X_1	X_2
	$X_1 \cdot W^Q$	$X_2 \cdot W^Q$
Query	q_1	q_2
	$X_1 \cdot W^K$	$X_2 \cdot W^K$
key	k_1	k_2

W^Q = Query Matrix

W^K = key Matrix

W^V = value Matrix.

Attention score (scale dot produce)

$q_1 \cdot k_1^T = s_{11}$ — $X_1 W^V$.

$q_1 \cdot k_2^T = s_{12}$

$Softmax\left(\frac{s}{\sqrt{d_k}}\right) \cdot V$ = self attention

$\sqrt{d_k}$ = dimension of query / key