Introduction to Cognitive Science

(4: Neural Networks)

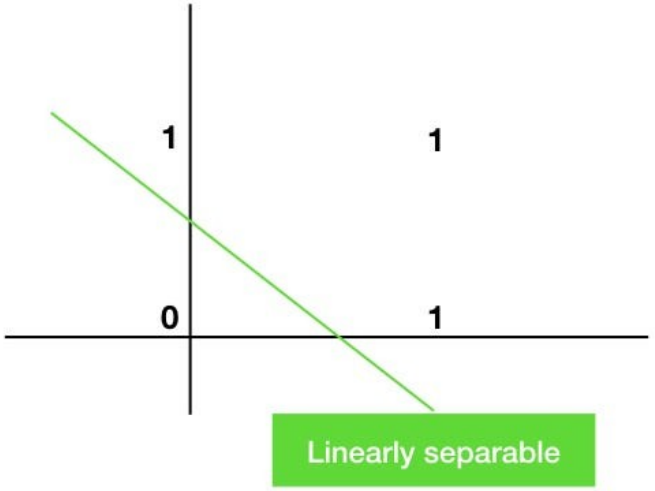
(Ch. 5 and 12)

Brief history

- 1943: W. Pitts, W. McCulloch
- 1954: Bar Hillel
- 1956: Dartmouth
- 1969: Minsky/Papert, XOR
- 1973: Lighthill report
- 1986: Rumelhart/Hinton, Backpropagation
- 1997: LeCun, CNN, Schmidhuber, RNN
- 2017: Attention is all you need

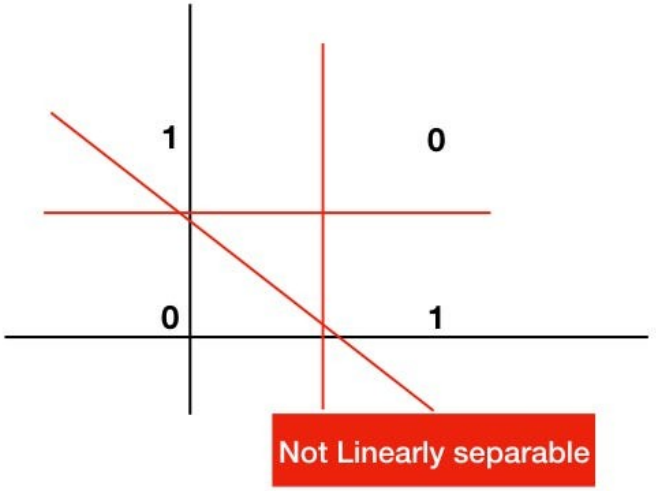
1969

Inclusive-OR



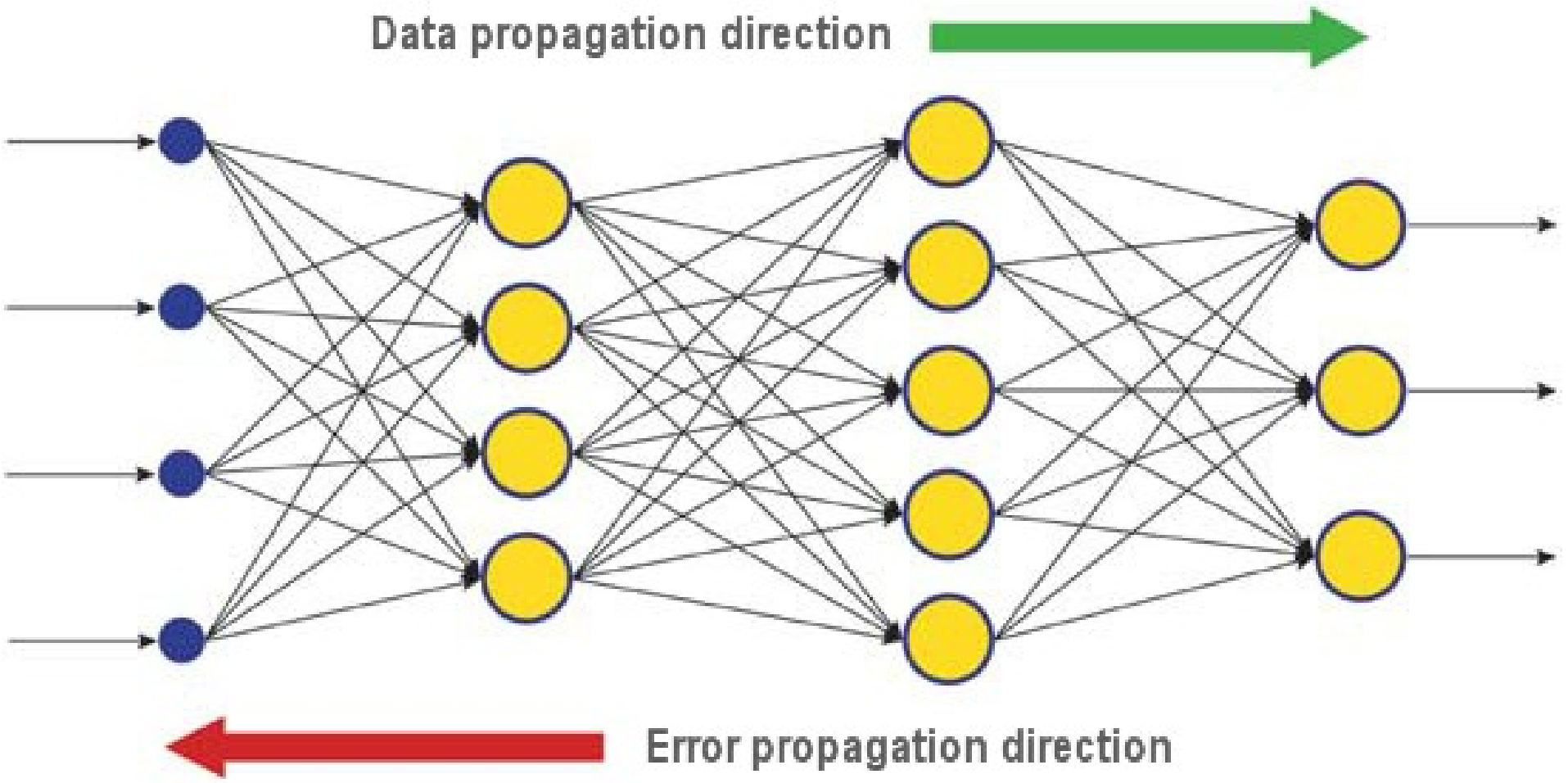
a	b	c
0	0	0
0	1	1
1	0	1
1	1	1

Exclusive-OR

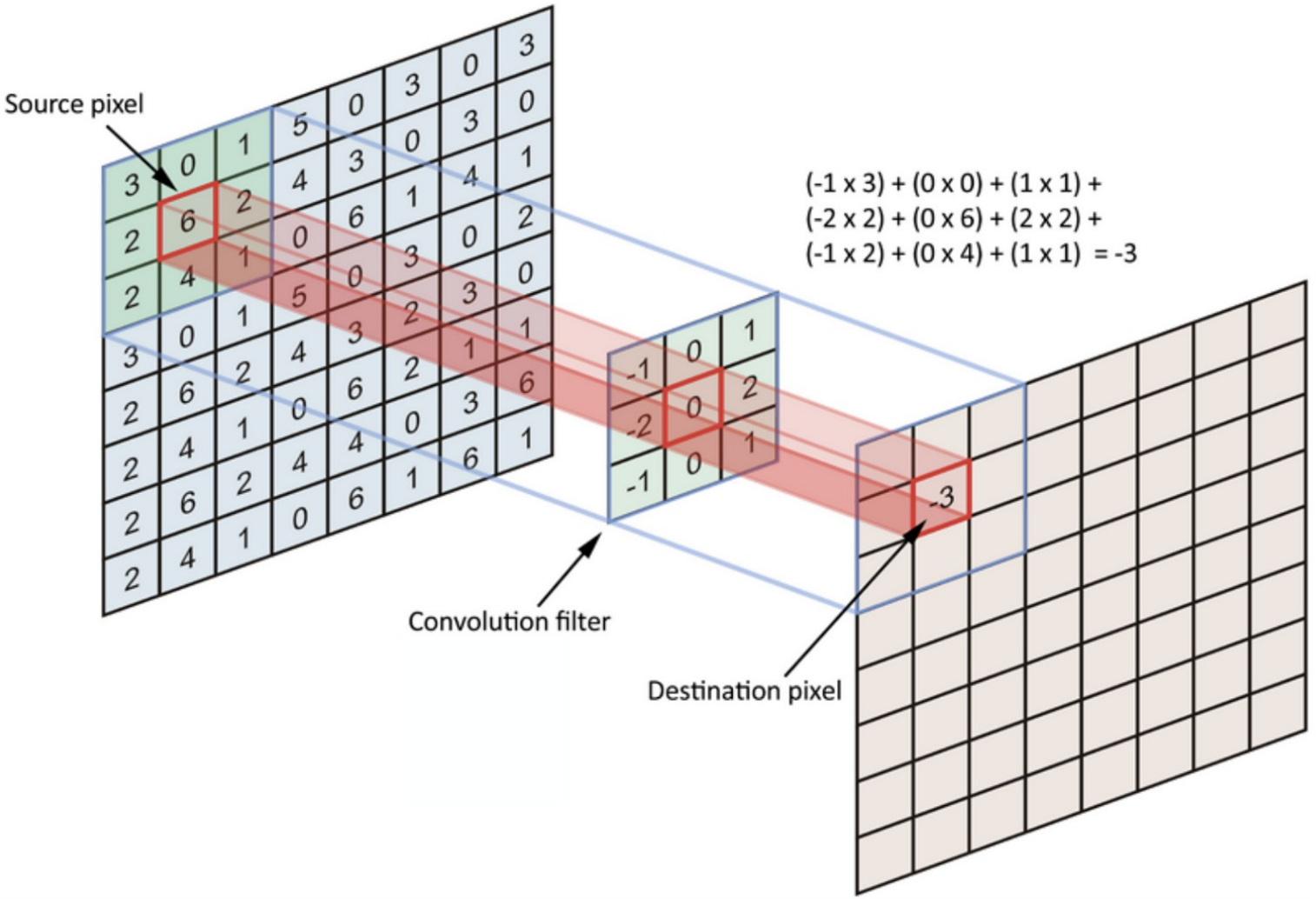


a	b	c
0	0	0
0	1	1
1	0	1
1	1	0

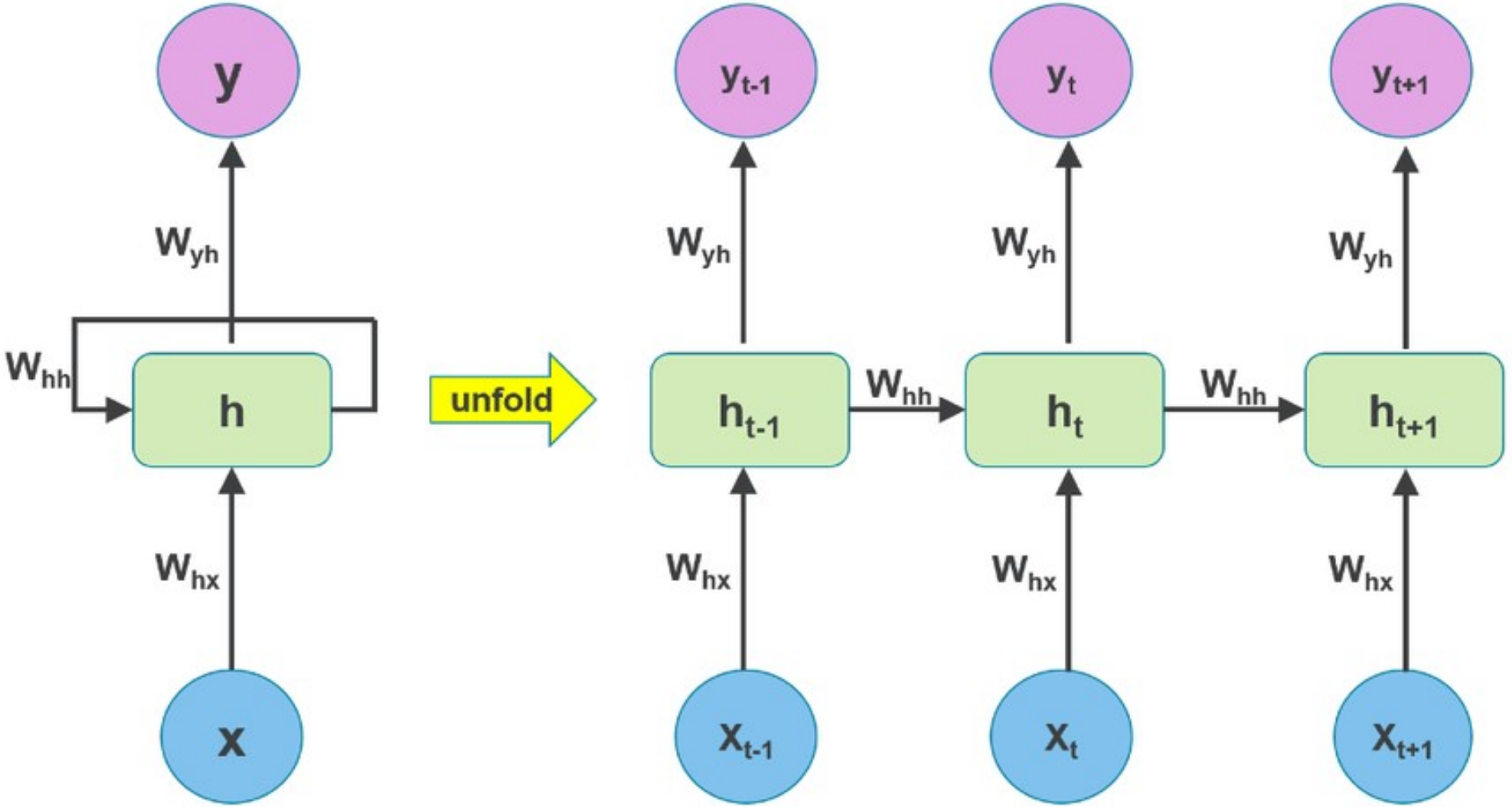
1986

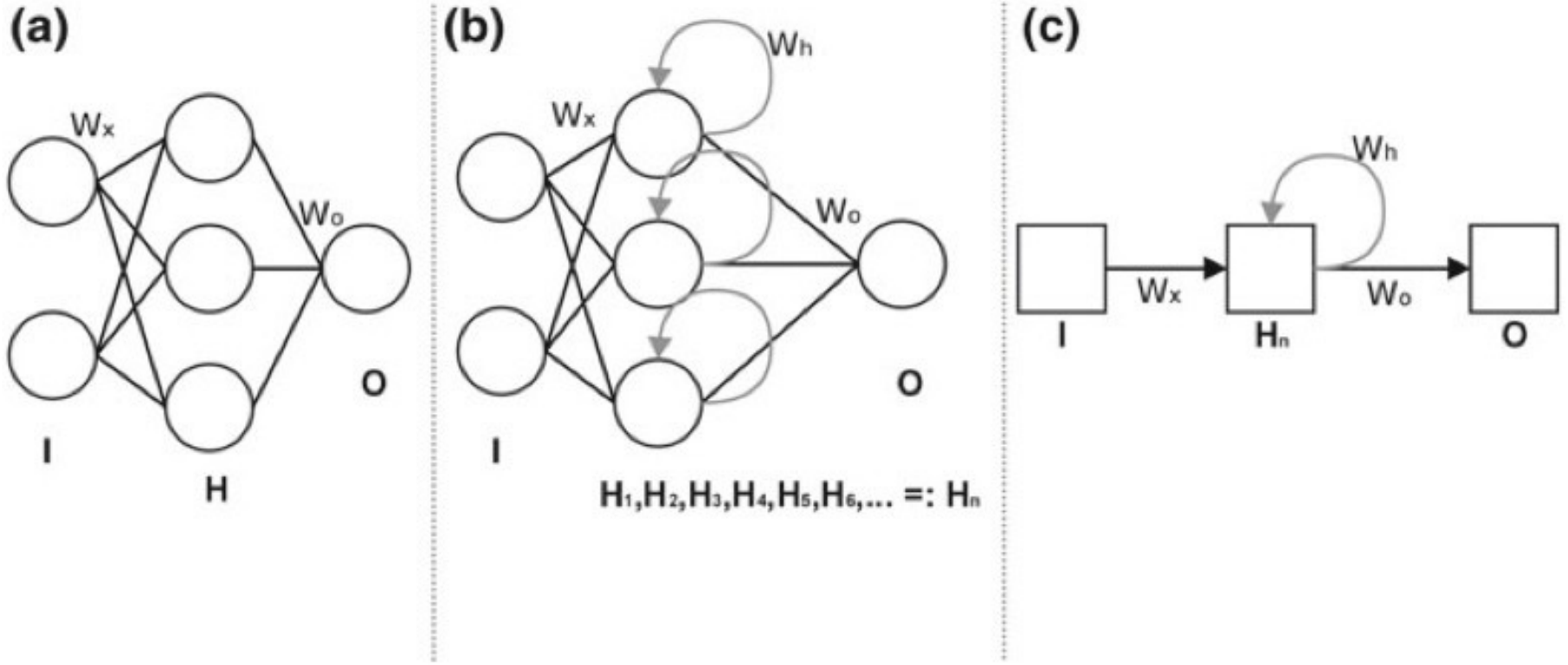


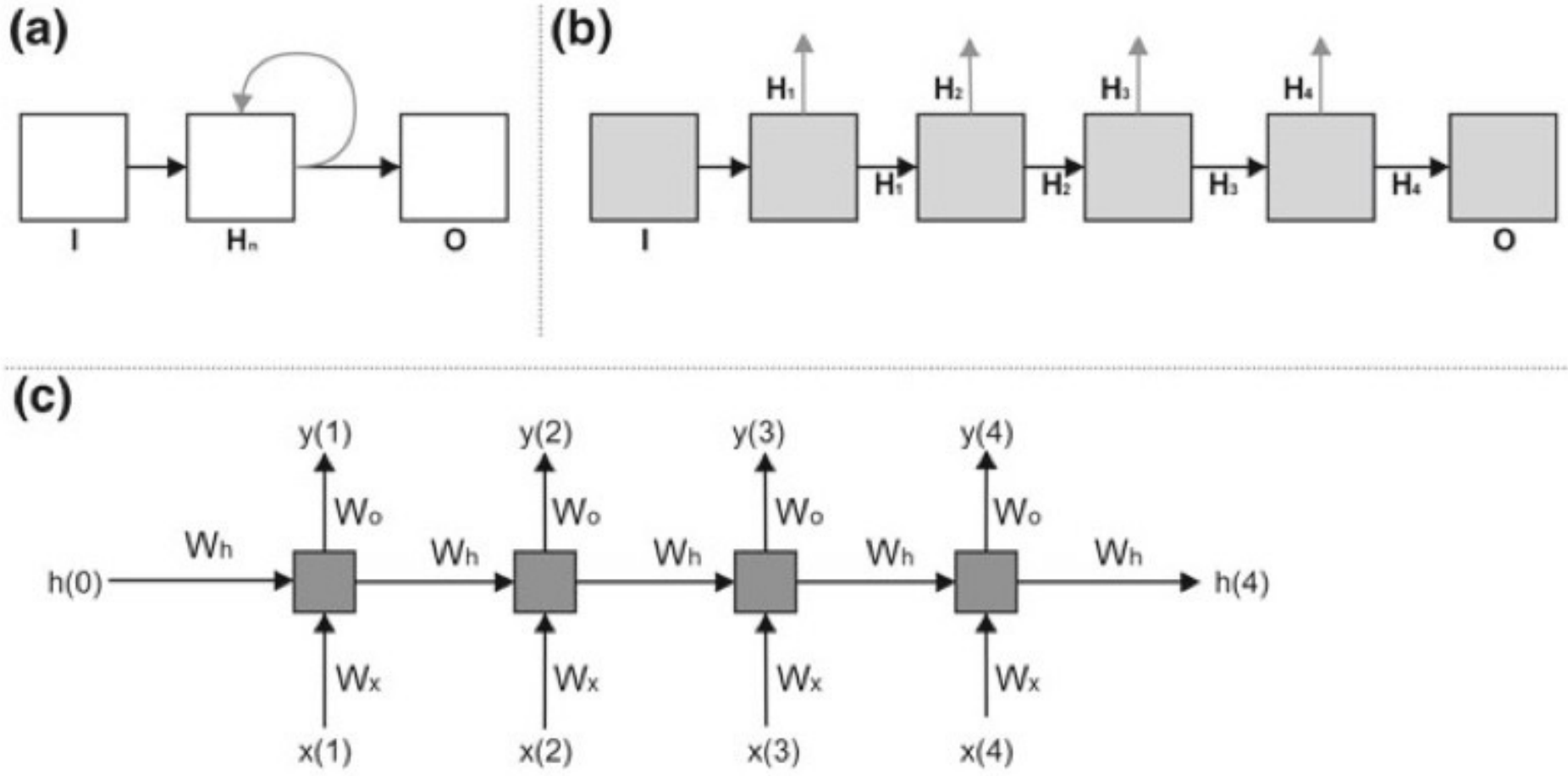
1997 CNN

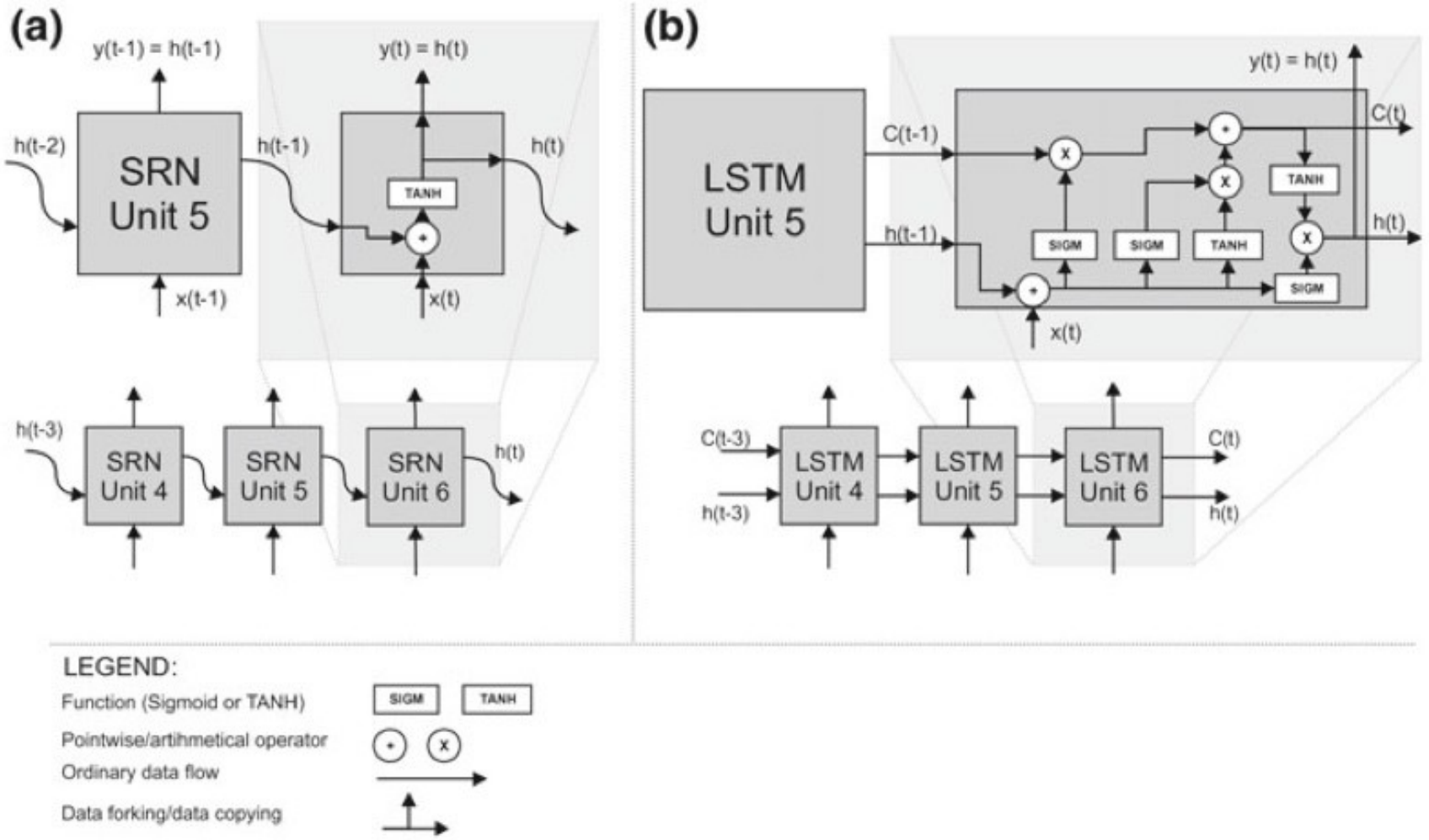

Source pixel
(-1 x 3) + (0 x 0) + (1 x 1) +
(-2 x 2) + (0 x 6) + (2 x 2) +
(-1 x 2) + (0 x 4) + (1 x 1) = -3
-3
Convolution filter
Destination pixel

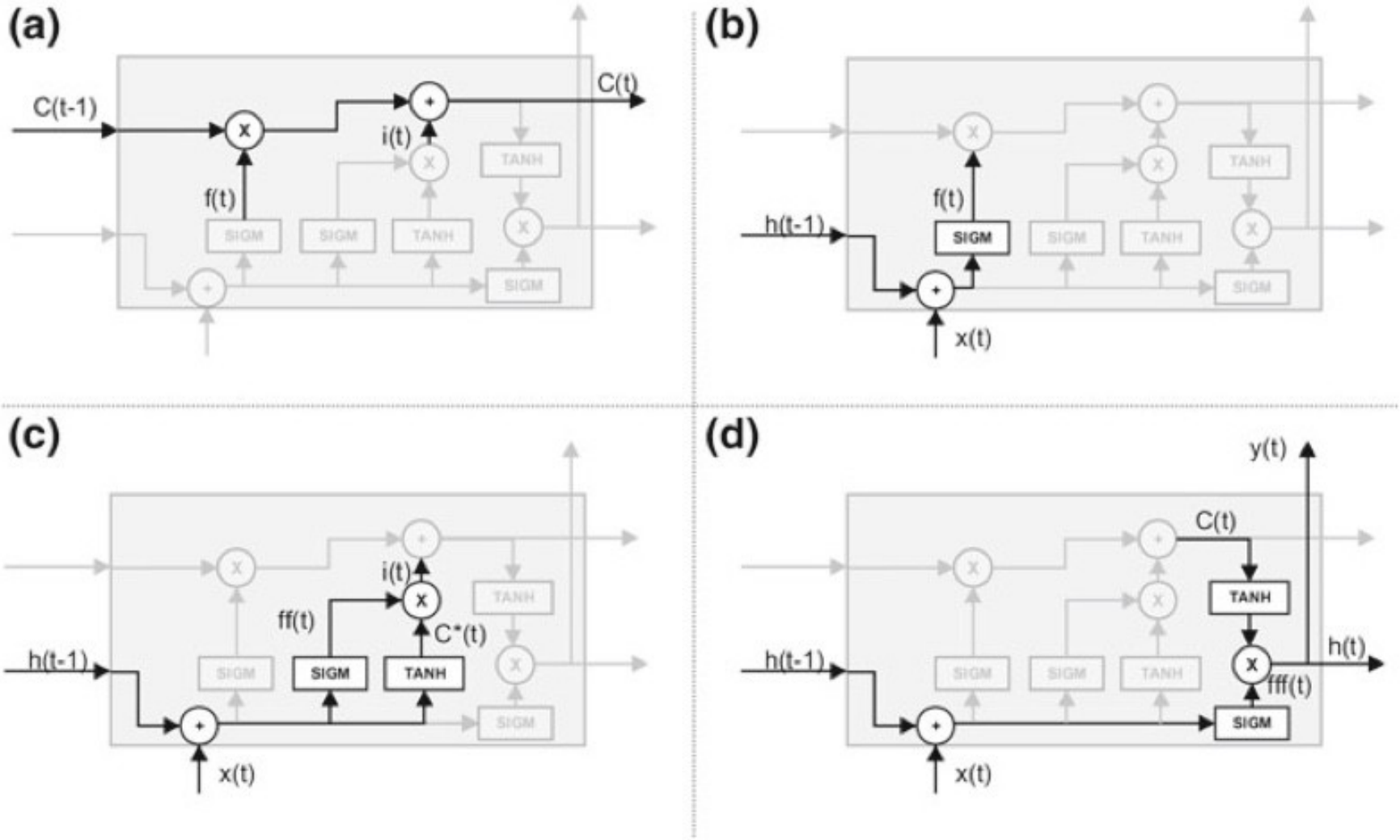
1997 RNN


y
W_yh
W_hh
h
W_hx
x
unfold
y_t-1
y_t
y_t+1
W_yh
W_yh
W_yh
W_hh
W_hh
h_t-1
h_t
h_t+1
W_hx
W_hx
W_hx
x_t-1
x_t
x_t+1

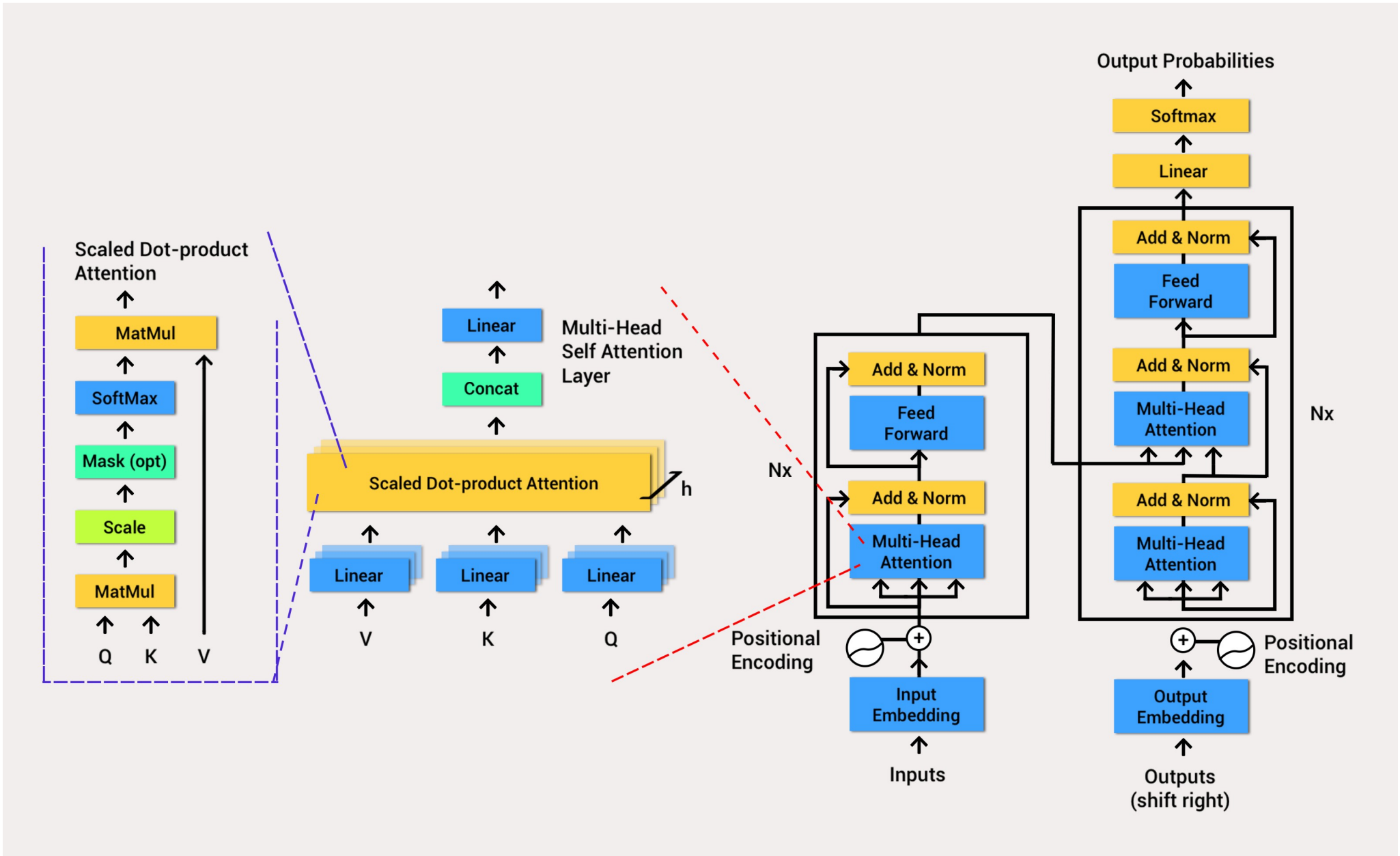

(a)
Wx
Wo
I
H
O
(b)
Wh
Wx
Wo
I
O
H1,H2,H3,H4,H5,H6,... =: Hn
(c)
Wh
Wx
Wo
I
Hn
O


(a)
I
Hn
O
(b)
H1
H2
H3
H4
I
H1
H2
H3
H4
O
(c)
y(1)
y(2)
y(3)
y(4)
Wo
Wh
h(0)
h(4)
Wx
x(1)
x(2)
x(3)
x(4)


(a)
y(t-1) = h(t-1)
y(t) = h(t)
h(t-2)
SRN Unit 5
h(t-1)
TANH
h(t)
x(t-1)
x(t)
h(t-3)
SRN Unit 4
SRN Unit 5
SRN Unit 6
h(t)
(b)
y(t) = h(t)
C(t-1)
LSTM Unit 5
C(t)
X
+
TANH
h(t-1)
SIGM
SIGM
TANH
h(t)
SIGM
x(t)
C(t-3)
h(t-3)
LSTM Unit 4
LSTM Unit 5
LSTM Unit 6
C(t)
h(t)
LEGEND:
Function (Sigmoid or TANH)
SIGM
TANH
Pointwise/artihmetical operator
Ordinary data flow
Data forking/data copying


(a)
C(t-1)
C(t)
i(t)
f(t)
SIGM
SIGM
TANH
SIGM
TANH
(b)
f(t)
h(t-1)
SIGM
SIGM
TANH
TANH
SIGM
x(t)
(c)
i(t)
ff(t)
C*(t)
h(t-1)
SIGM
TANH
TANH
SIGM
SIGM
x(t)
(d)
y(t)
C(t)
TANH
h(t)
h(t-1)
fff(t)
SIGM
SIGM
SIGM
TANH
SIGM
x(t)

2017: Transformers



Key features for CogSci

- Distributed representations:
  - According to the PhySysHyp representations are distinct and identifiable components in a physical system
- No clear distinction between information storage and processing
- Learnability