Machine Translation 2
Neural Machine Translation

CS 287

Quiz:

Neural Machine Translation

$$p(w^t|w^s) = p(w_i^t|w^s, \theta)$$

Babbler

- $$p(w_i | w_1, \ldots, w_{i-1}) = \text{softmax}(RNN()) = \hat{y}_{w_i}$$

True Encoder-Decoder

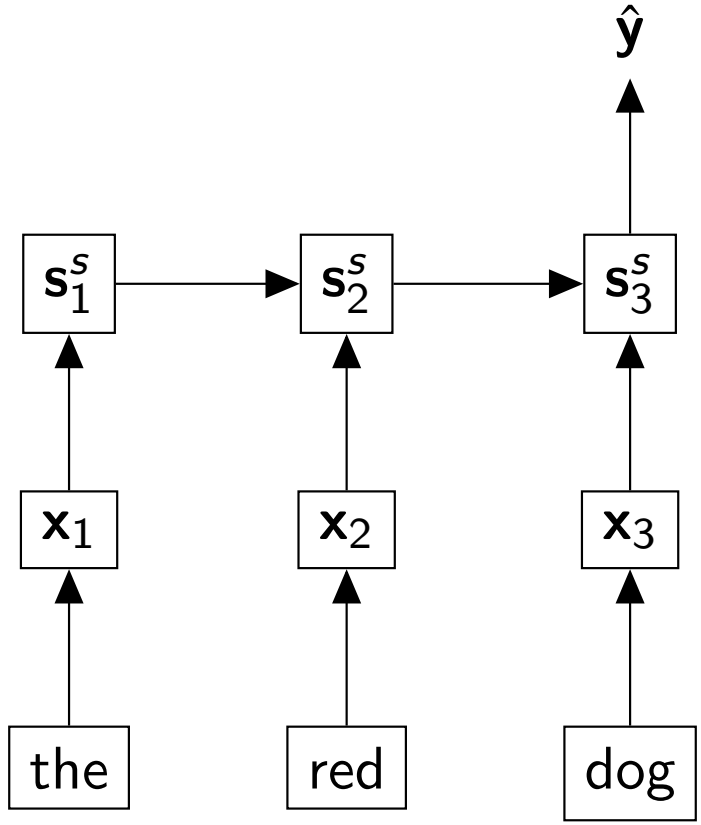
Compute a single vector $\mathbf{x}$ representing the source.

- $$p(w_i | w_1, \ldots, w_{i-1}, \mathbf{x}) = \hat{y}_{w_i}$$

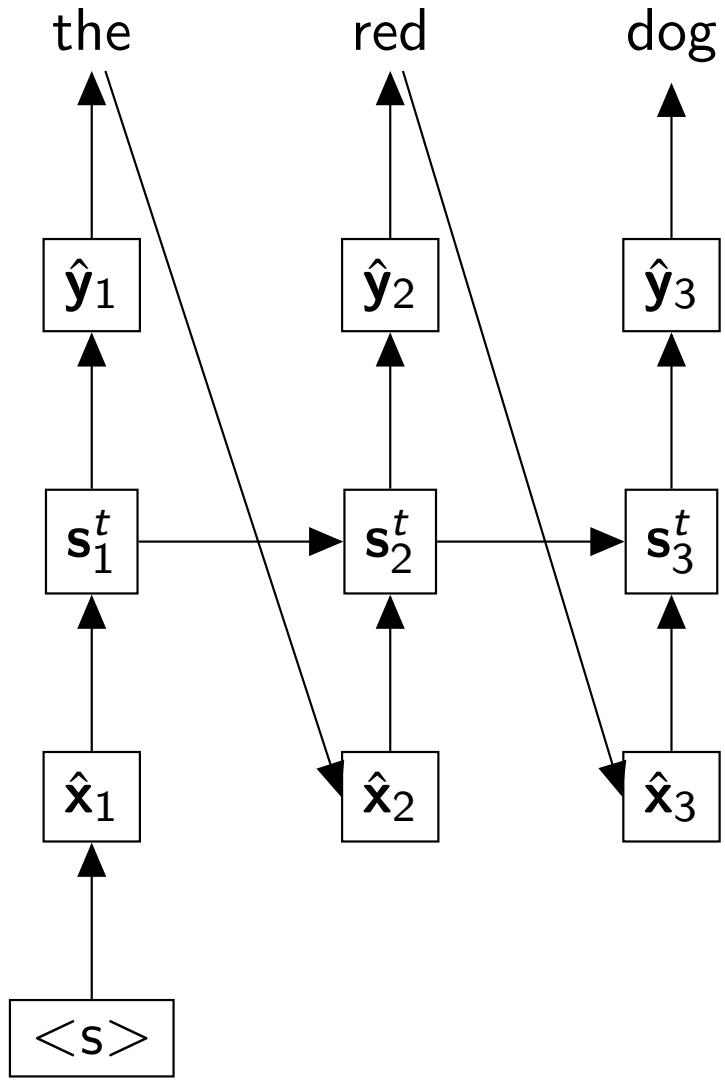
Compute output $\mathbf{y}$

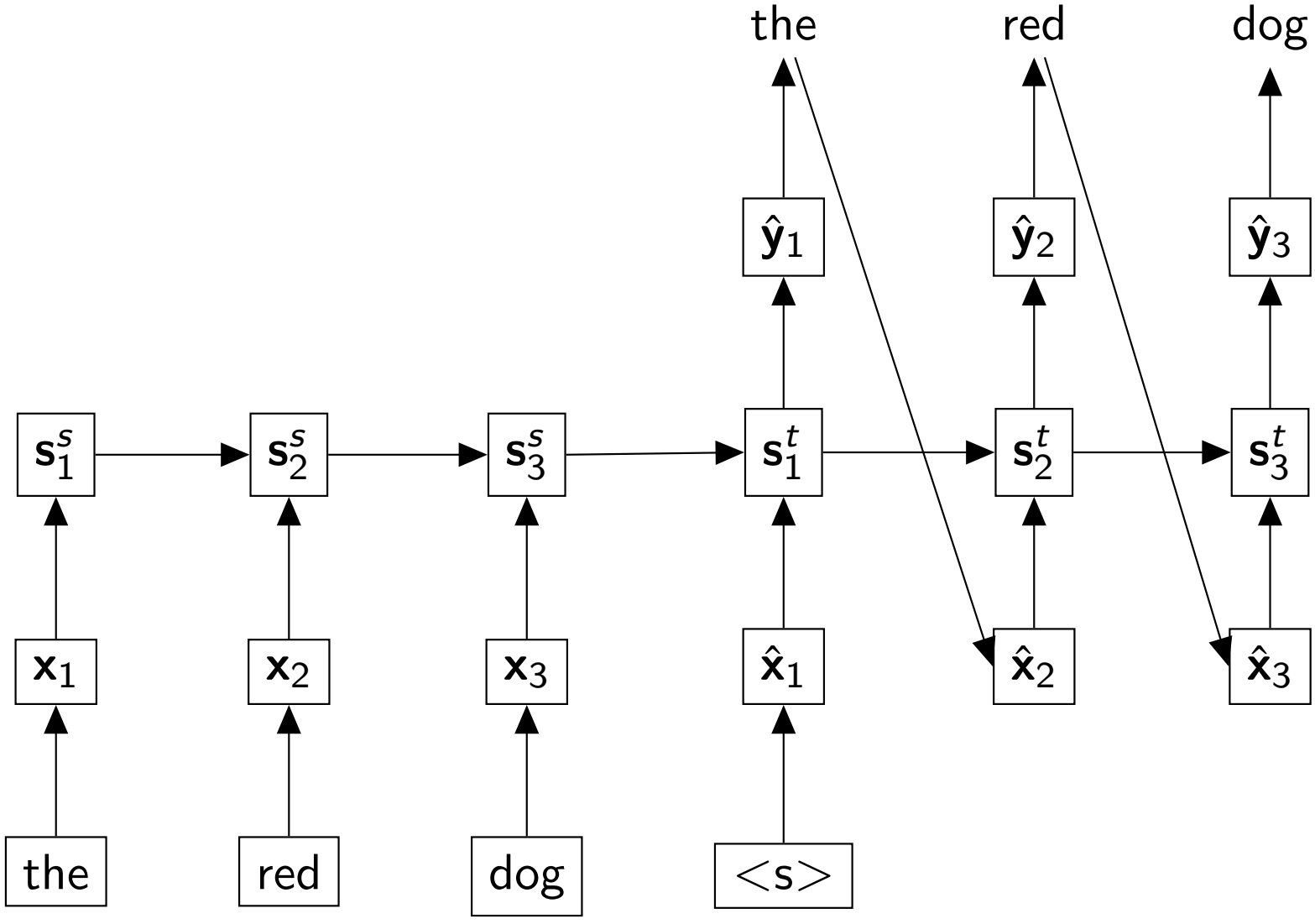
Bottleneck here $\mathbf{x}$ full representation of source.

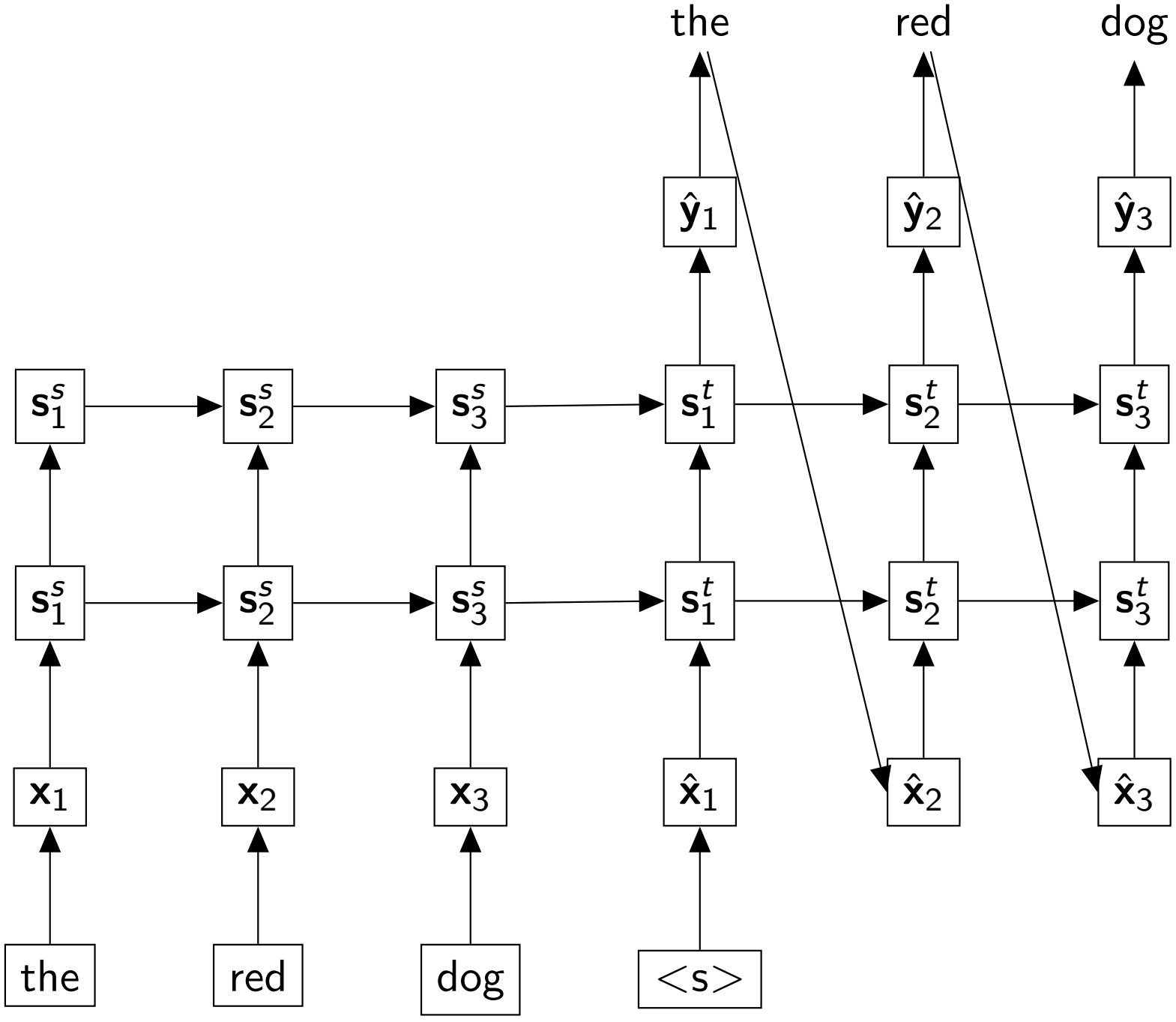
Acceptor


$\hat{\mathbf{y}}$
$\mathbf{s}_1^s$
$\mathbf{s}_2^s$
$\mathbf{s}_3^s$
$\mathbf{x}_1$
$\mathbf{x}_2$
$\mathbf{x}_3$
the
red
dog

Babbler


the
red
dog
$\hat{\mathbf{y}}_1$
$\hat{\mathbf{y}}_2$
$\hat{\mathbf{y}}_3$
$\mathbf{s}_1^t$
$\mathbf{s}_2^t$
$\mathbf{s}_3^t$
$\hat{\mathbf{x}}_1$
$\hat{\mathbf{x}}_2$
$\hat{\mathbf{x}}_3$
<s>


the
red
dog
$\hat{\mathbf{y}}_1$
$\hat{\mathbf{y}}_2$
$\hat{\mathbf{y}}_3$
$\mathbf{s}_1^s$
$\mathbf{s}_2^s$
$\mathbf{s}_3^s$
$\mathbf{s}_1^t$
$\mathbf{s}_2^t$
$\mathbf{s}_3^t$
$\mathbf{x}_1$
$\mathbf{x}_2$
$\mathbf{x}_3$
$\hat{\mathbf{x}}_1$
$\hat{\mathbf{x}}_2$
$\hat{\mathbf{x}}_3$
the
red
dog
<s>


the
red
dog
$\hat{\mathbf{y}}_1$
$\hat{\mathbf{y}}_2$
$\hat{\mathbf{y}}_3$
$\mathbf{s}_1^s$
$\mathbf{s}_2^s$
$\mathbf{s}_3^s$
$\mathbf{s}_1^t$
$\mathbf{s}_2^t$
$\mathbf{s}_3^t$
$\mathbf{s}_1^s$
$\mathbf{s}_2^s$
$\mathbf{s}_3^s$
$\mathbf{s}_1^t$
$\mathbf{s}_2^t$
$\mathbf{s}_3^t$
$\mathbf{x}_1$
$\mathbf{x}_2$
$\mathbf{x}_3$
$\hat{\mathbf{x}}_1$
$\hat{\mathbf{x}}_2$
$\hat{\mathbf{x}}_3$
the
red
dog
<s>

Which word should we translate next?

Alignment

- Pick a source word.
- Pick a translation.
- Score placing that as next word.

$$i = \arg\max p(;\theta)$$

$$p(w_i^s)$$

- What is the issue here?

Argmax

Can't backprop over the argmax

The Attention Mechanism

Idea: Replace internal argmax with softmax

- Similar to the

Attention-Based Encoder-Decoder

At timestep i,

$$\hat{\mathbf{a}} = \text{softmax}(\mathbf{S}^s \mathbf{s}_i^t)$$

Soft-alignment over source.

$$p(a = |\mathbf{x}) = \hat{\mathbf{a}}_i$$

Neural Machine Translation