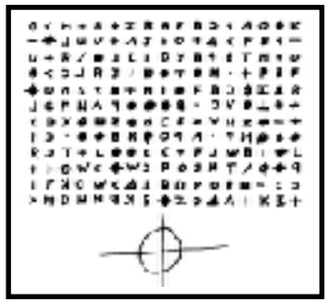
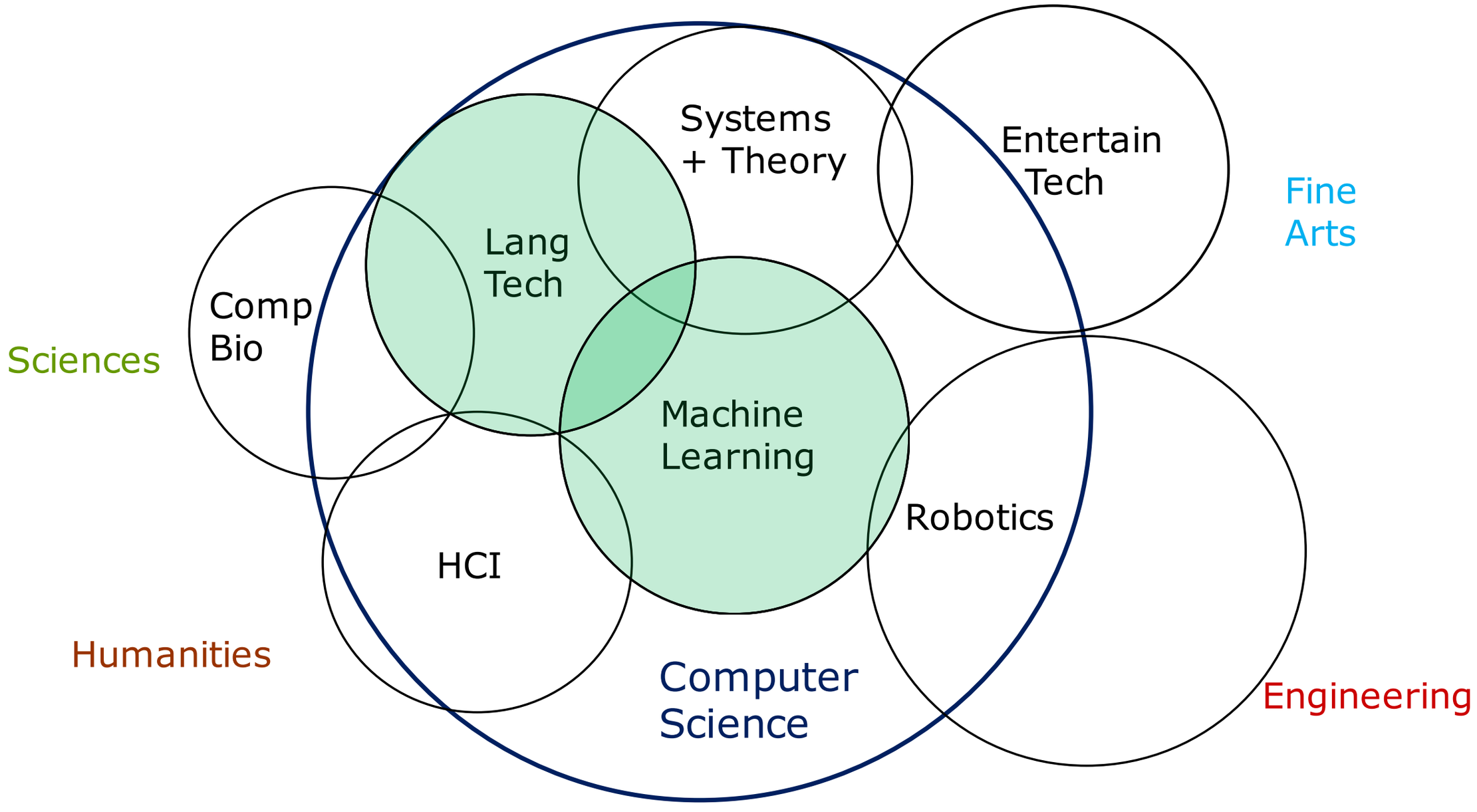
Modeling Continuous Human Artifacts: Music, Historical Documents, and Shakespeare



Taylor Berg-Kirkpatrick
Carnegie Mellon University

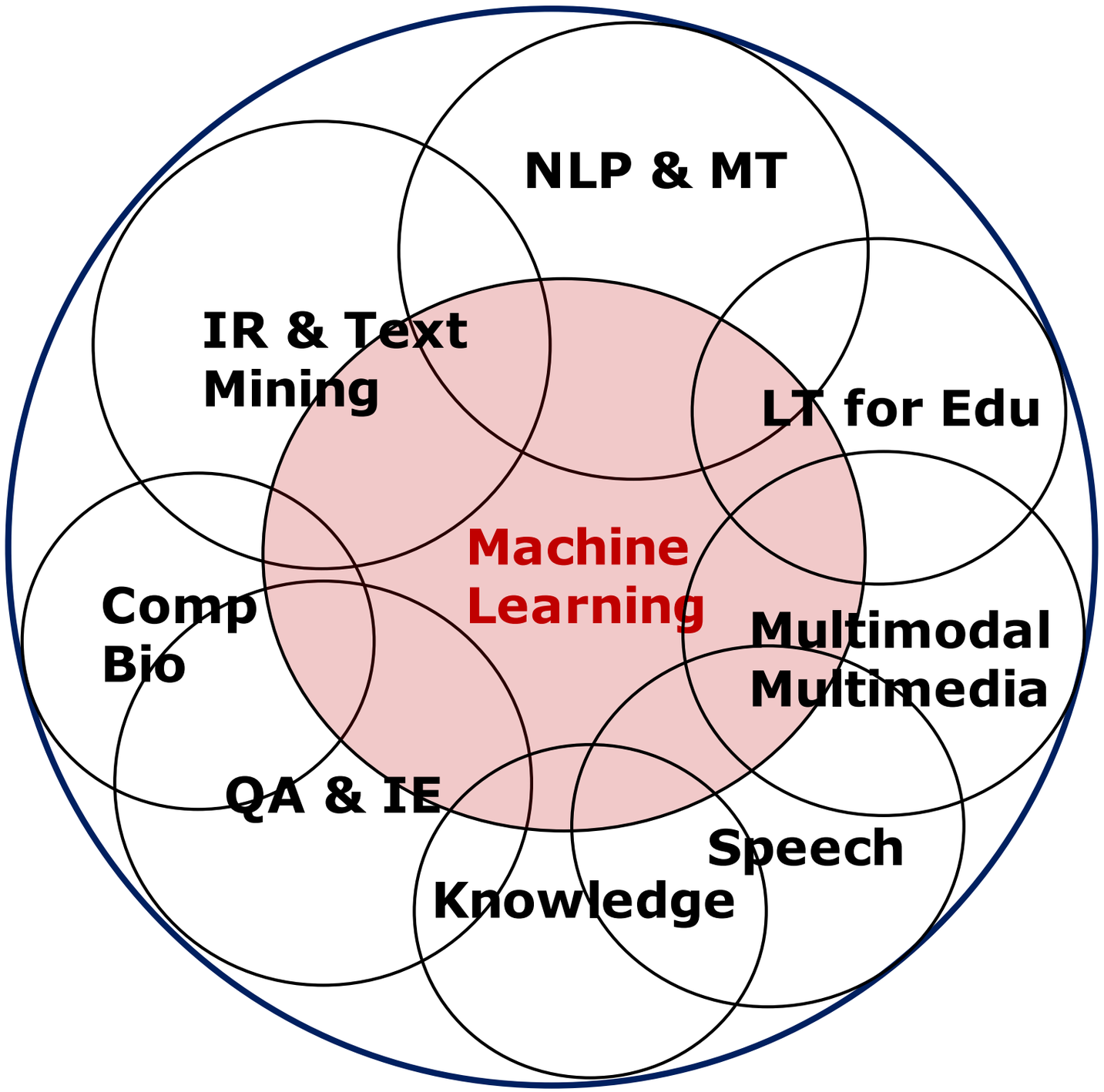
Joint work with Maria Ryskina, Kartik Goyal, Shruti Rijhwani,
Dan Garrette, Hannah Alpert-Abrams,
Greg Durrett, Jacob Andreas, and Dan Klein

Computing Departments at CMU



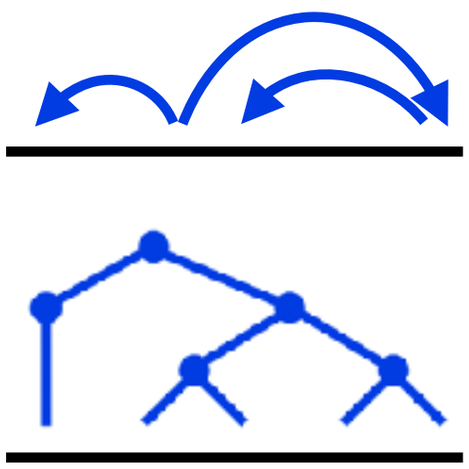
[slide credit: Yiming Yang]

Research in LTI

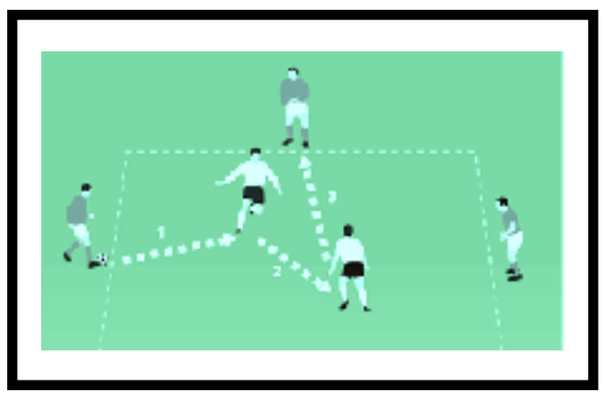


[slide credit: Yiming Yang]

Language Processing

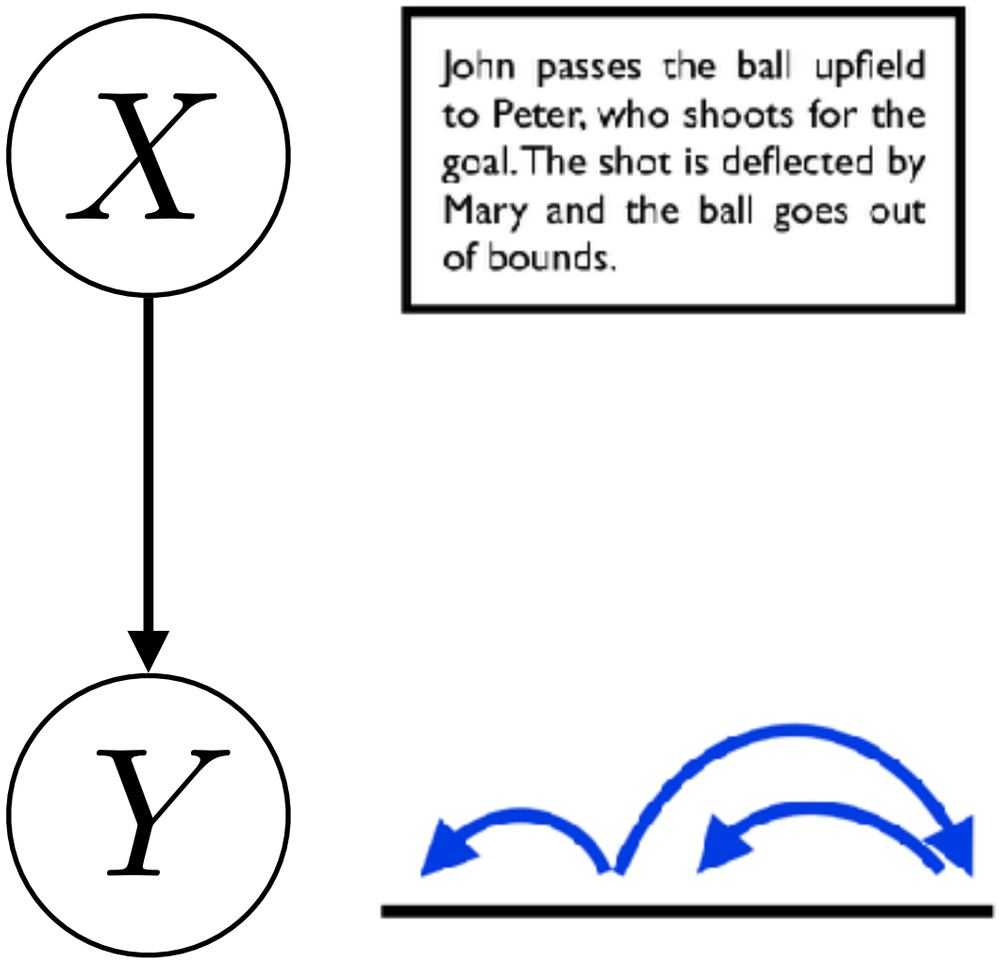


Mary prevents Peter from scoring a goal.

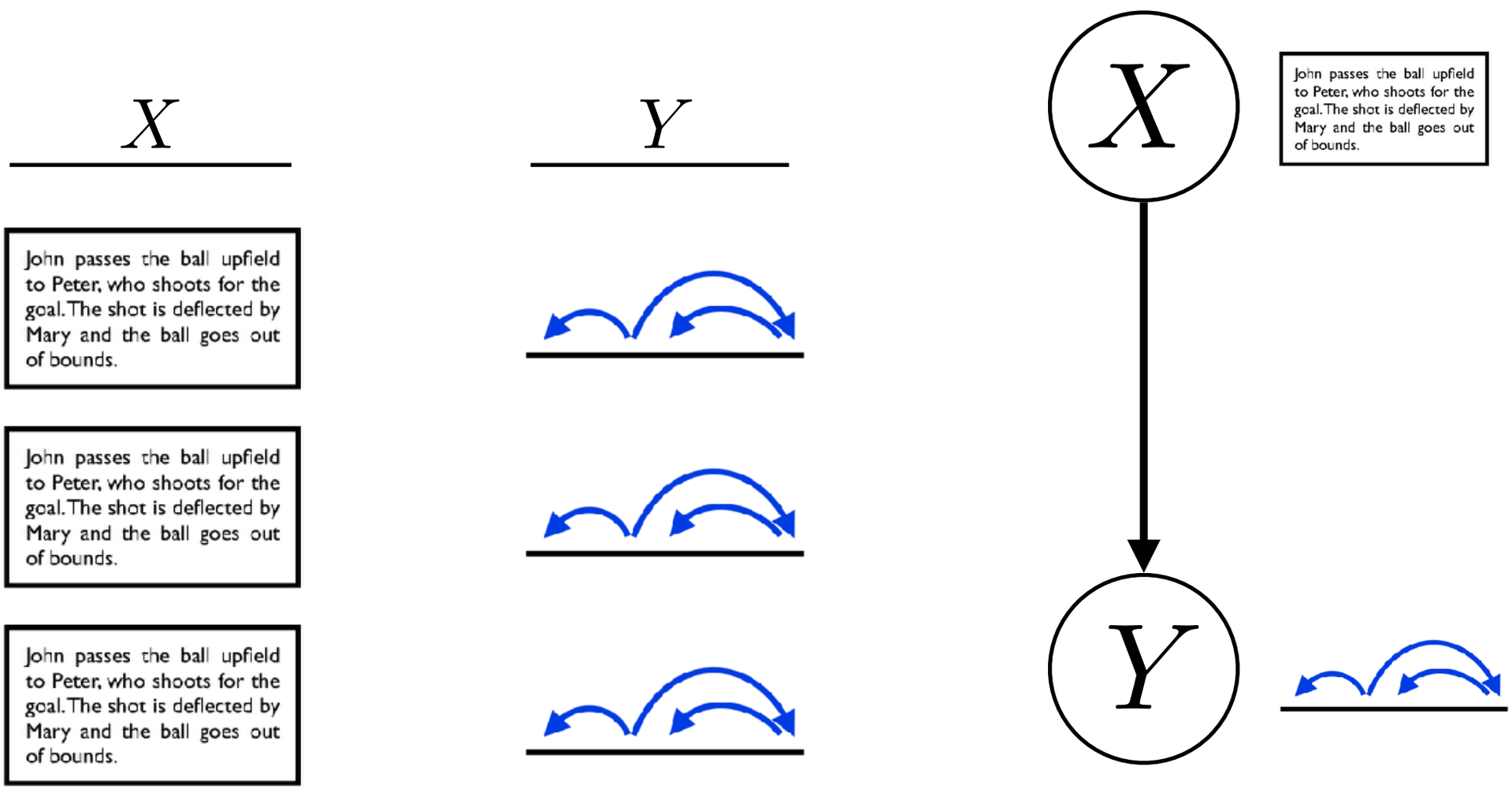


John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

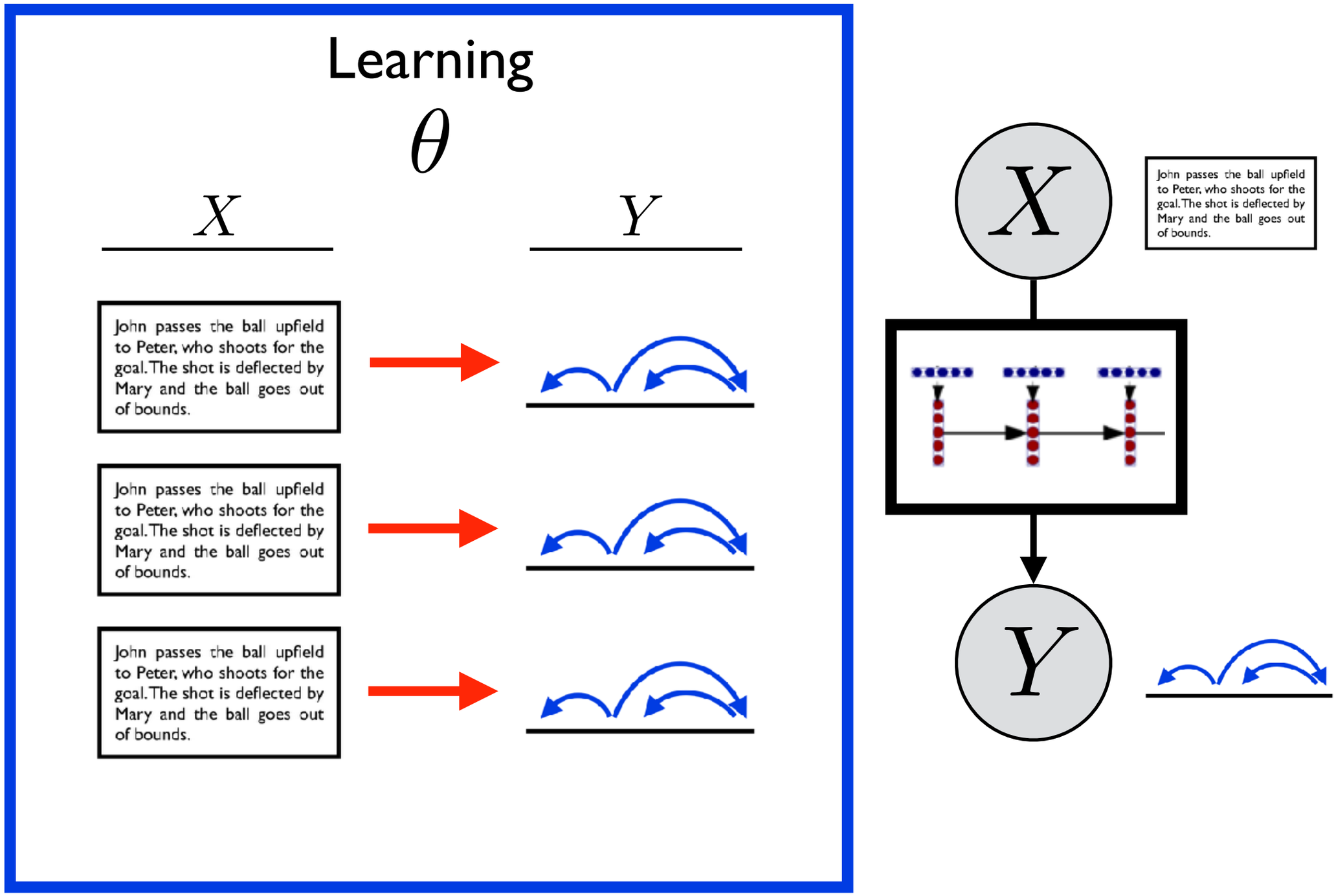
Supervised Learning



Supervised Learning



Supervised Learning

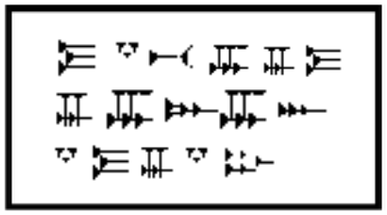


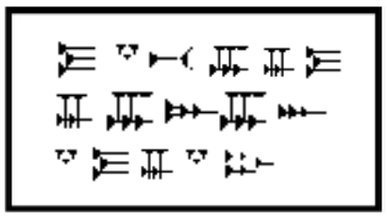
Supervised Learning

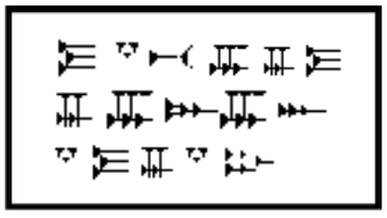
Learning

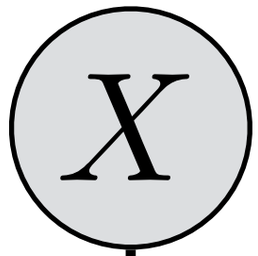
θ

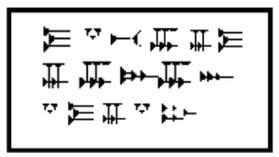
X | Y

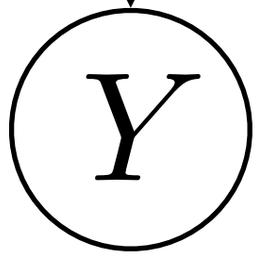




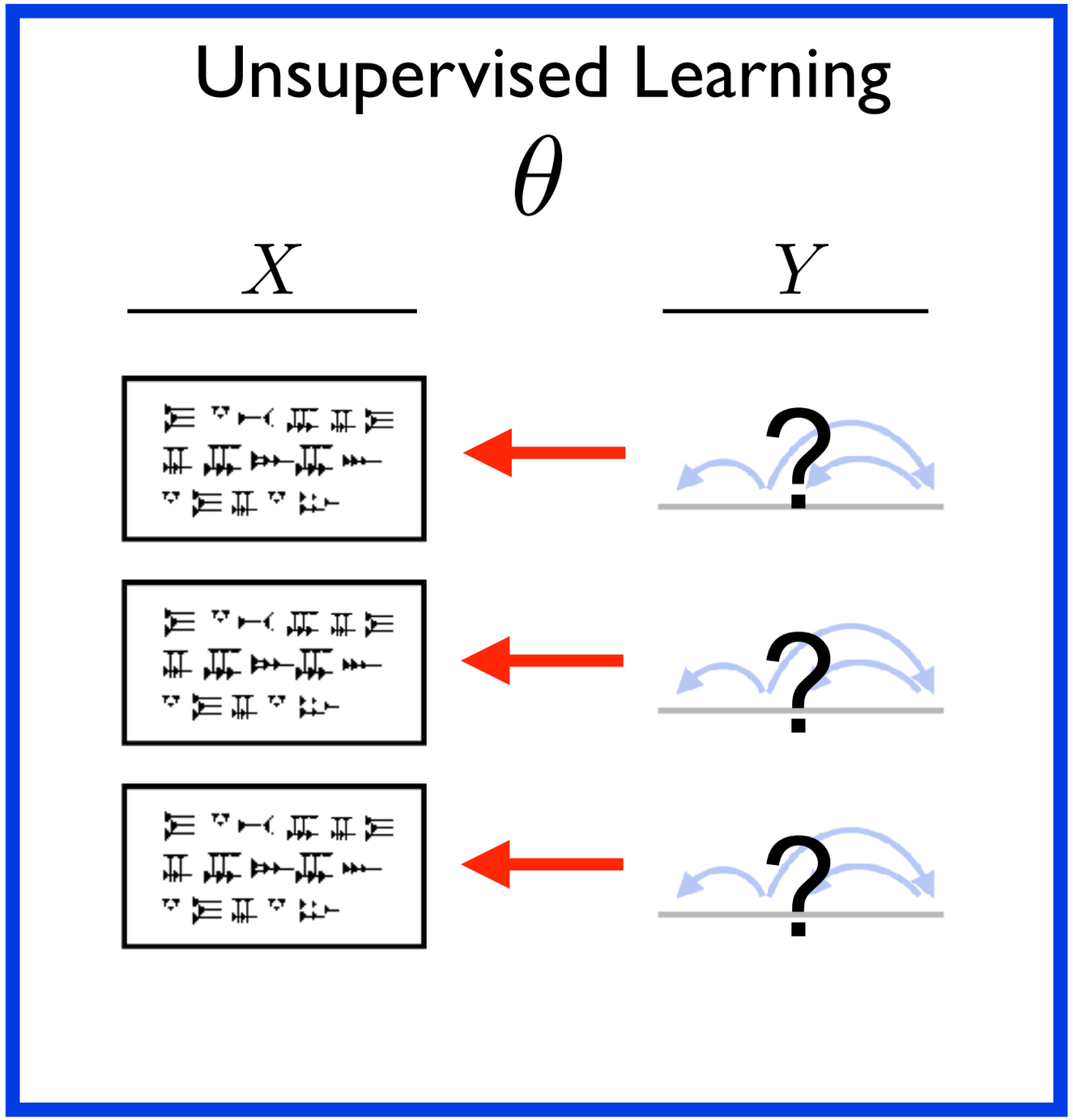


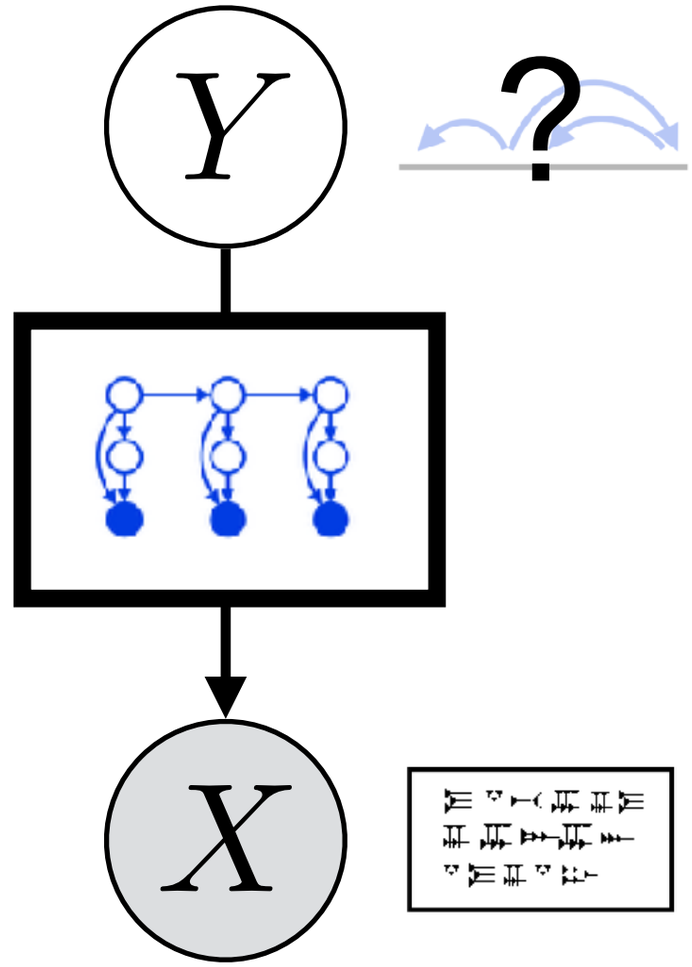






Unsupervised Learning





Latent Variable Models

X

Y

John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

Latent Variable Models

X

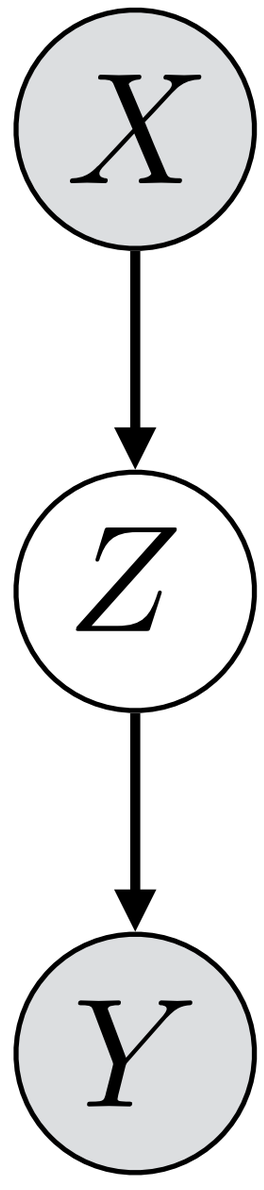
John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

John passes the ball upfield to Peter, who shoots for the goal. The shot is deflected by Mary and the ball goes out of bounds.

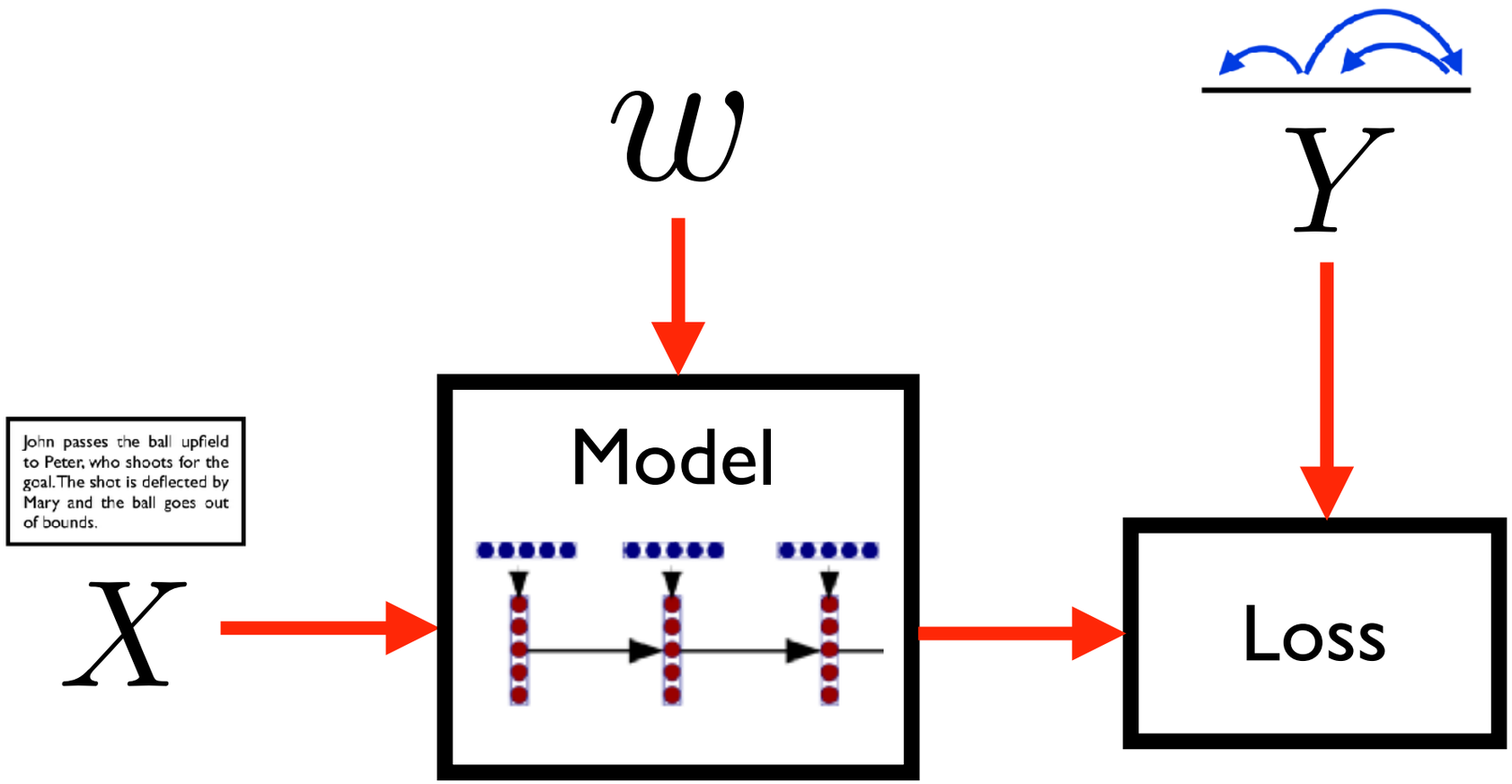
Z

Y

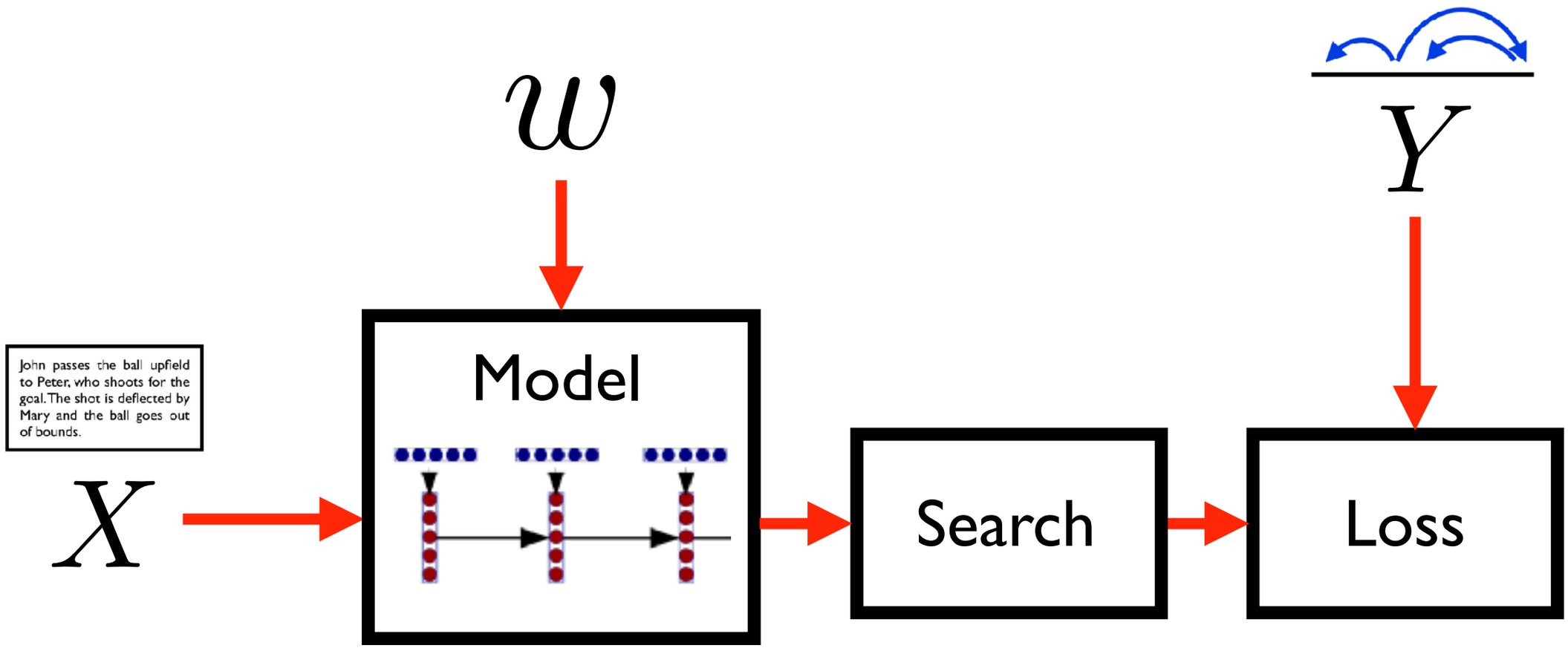


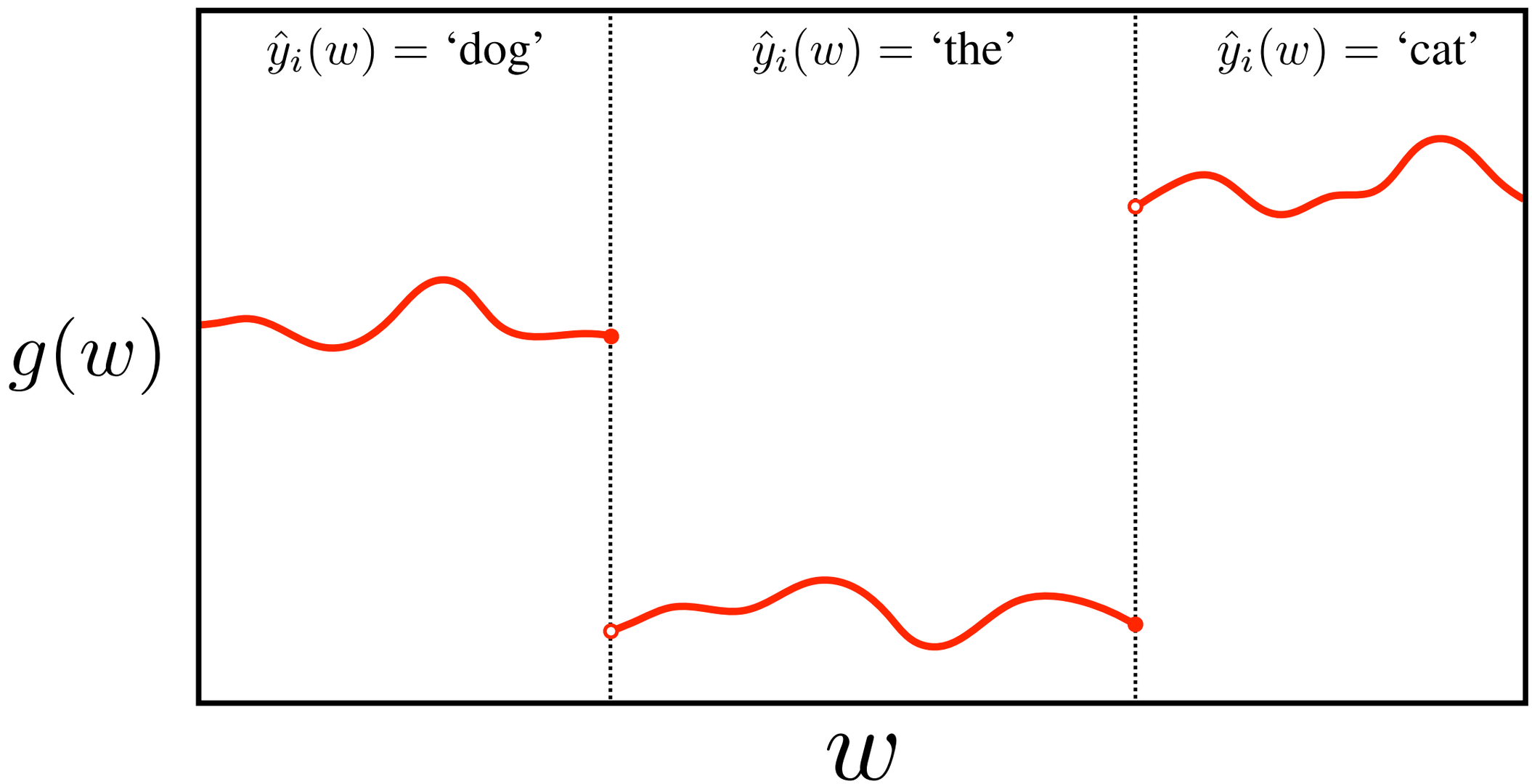
Supervised Learning

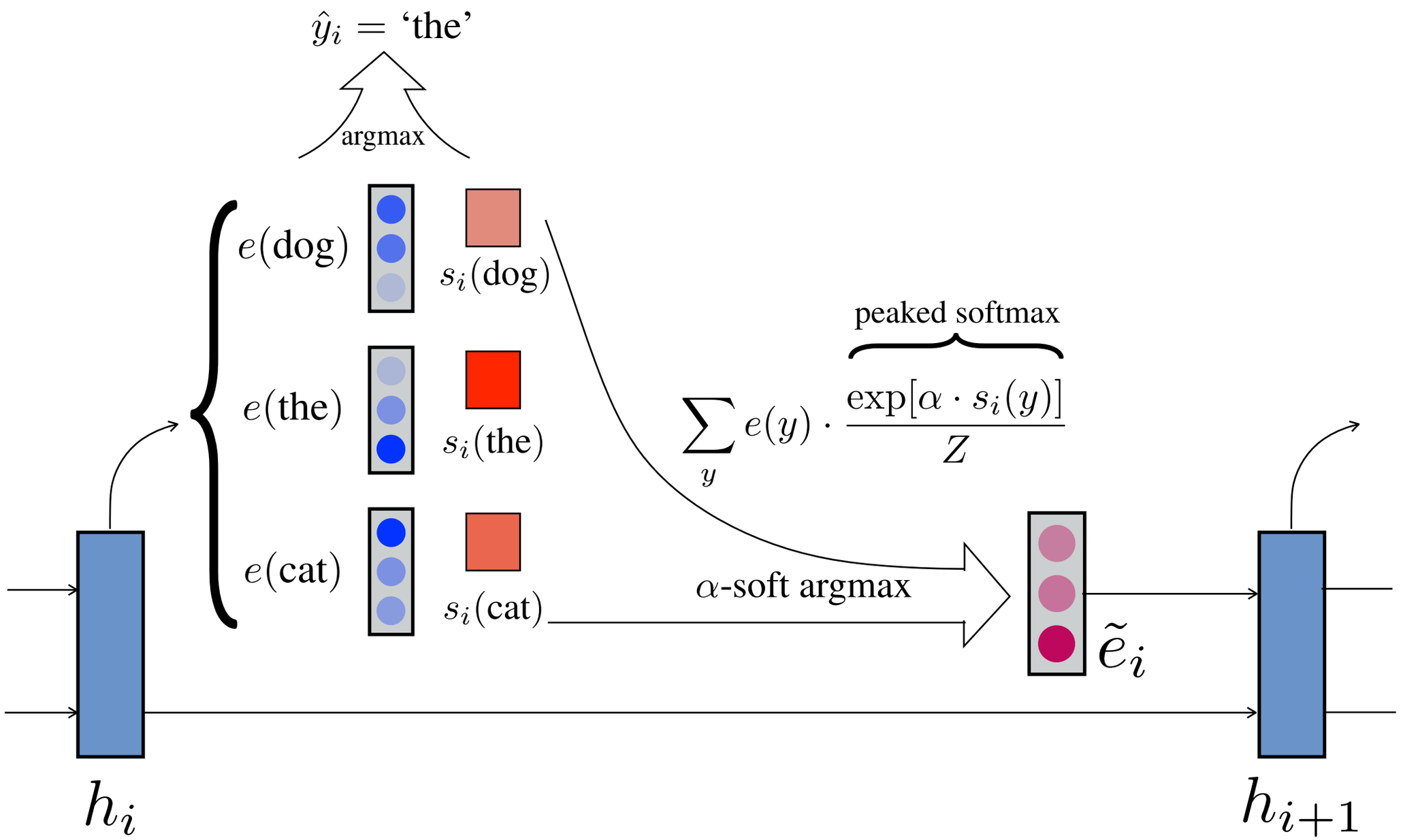
Neural Structured Prediction

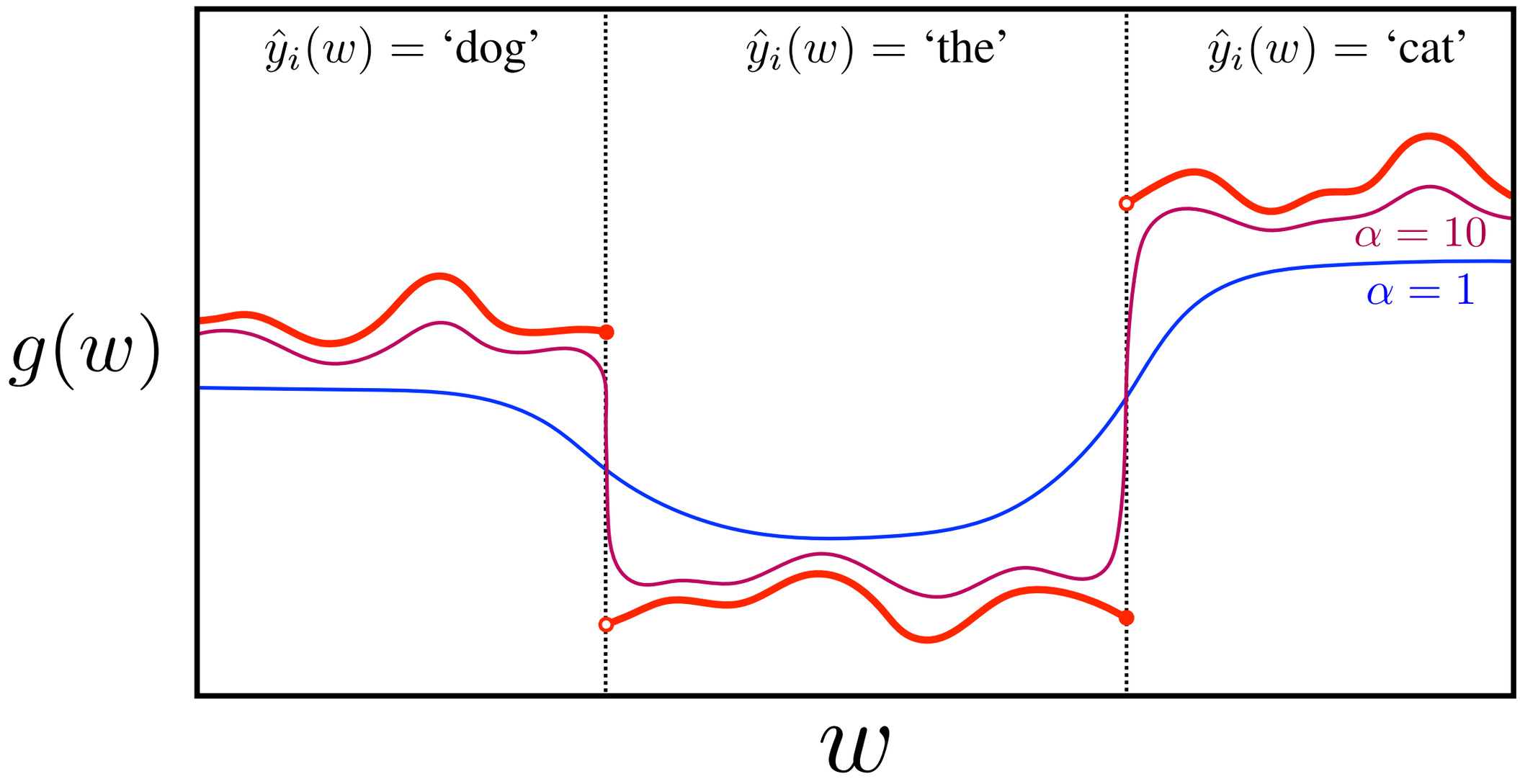


Neural Structured Prediction

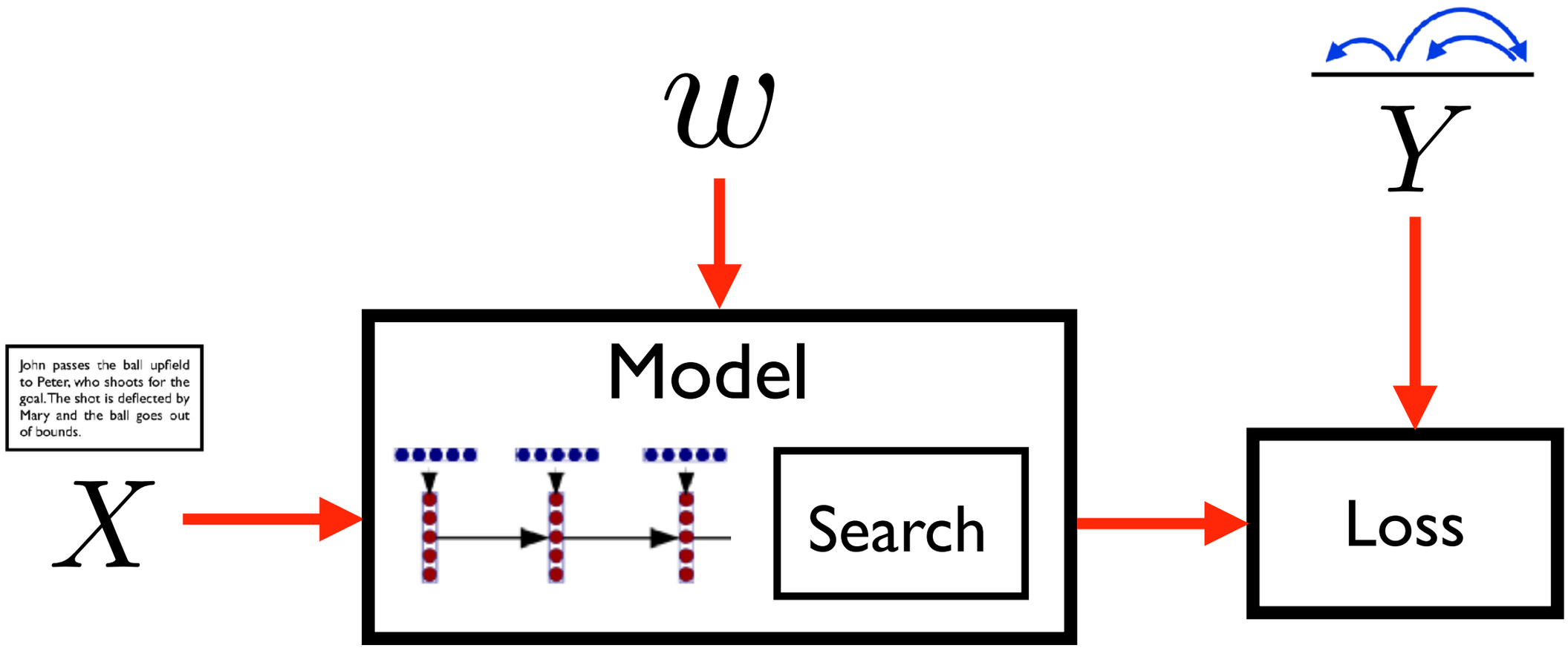



$\hat{y}_i(w) =$ ‘dog’
$\hat{y}_i(w) =$ ‘the’
$\hat{y}_i(w) =$ ‘cat’
$g(w)$
w


$\hat{y}_i$ = 'the'
argmax
e(dog)
s_i(dog)
e(the)
s_i(the)
e(cat)
s_i(cat)
peaked softmax
$\sum_y e(y) \cdot \frac{\exp[\alpha \cdot s_i(y)]}{Z}$
α-soft argmax
$\tilde{e}_i$
h_i
h_{i+1}

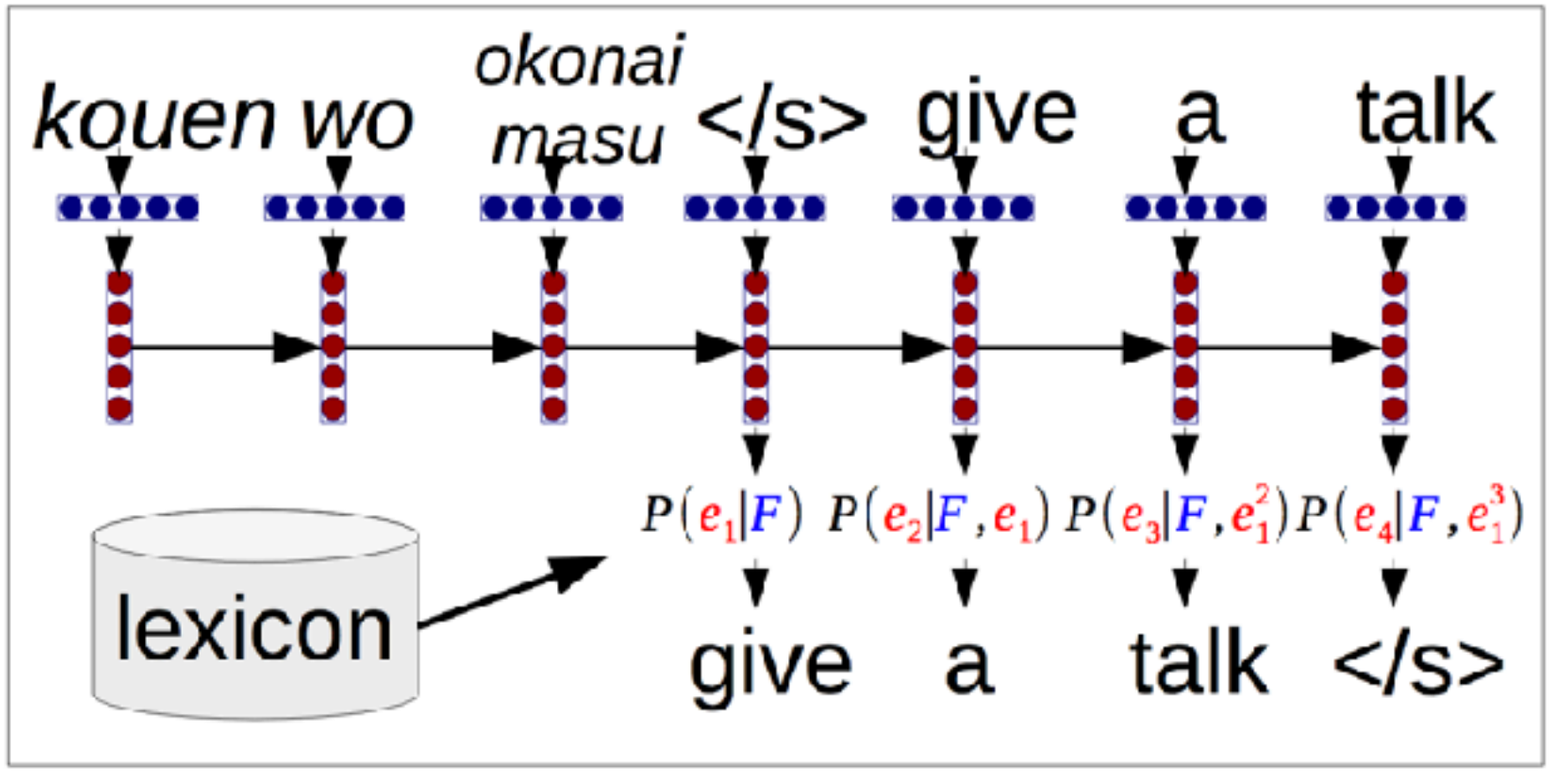

$\hat{y}_i(w) =$ ‘dog’
$\hat{y}_i(w) =$ ‘the’
$\hat{y}_i(w) =$ ‘cat’
$\alpha = 10$
$\alpha = 1$
$g(w)$
w

Neural Structured Prediction



Discrete Lexicons in Neural Seq2seq

- Problem: Neural translation models fail on rare words
- Solution: Addition of translation lexicon



- Result: Improvement of translation accuracy, particularly in low-data scenarios

[slide credit: Graham Neubig]

DyNet: Dynamic Neural Toolkit

- A toolkit for neural networks and deep learning that allows for more flexible **dynamic declaration of nets**

clab / dynet

Unwatch 139 | Star 1,221 | Fork 331

Code | Issues 57 | Pull requests 4 | Projects 1 | Wiki | Pulse | Graphs | Settings

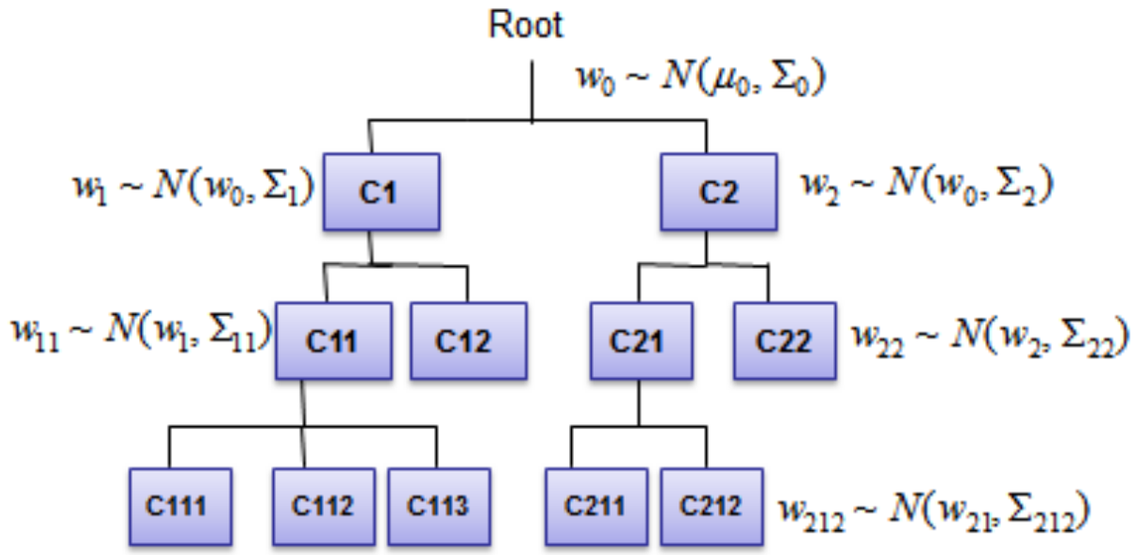
DyNet: The Dynamic Neural Network Toolkit — Edit

- Particularly suited for **NLP applications**
  - Declare complicated structures such as parse trees
  - Relatively fast, even on CPU
- **Lots of stuff to do!** (Dynamic mini-batching, distributed computing, etc.)

[slide credit: Graham Neubig]

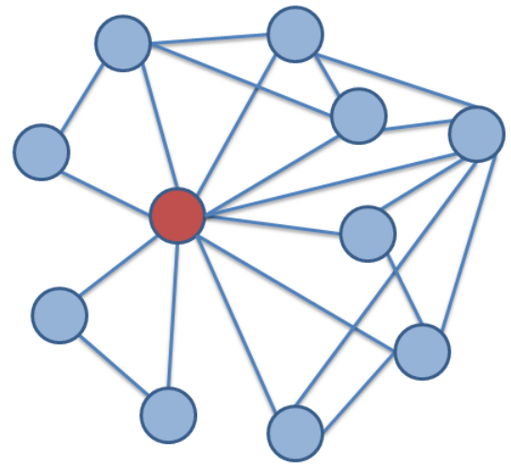
(Massive) Structured Regularizers

Hierarchical Bayesian Logistic Regression (HBLR)



- Each node has a vector ($\boldsymbol{w_i}$) of model parameters
- Assume $\boldsymbol{w_i}$ depending on its parent via a Gaussian prior (via its mean and covariance)
- All the parameters need to be jointly optimized.

Graphical Dependencies among Classifiers



$$\hat{\mathbf{W}} = \arg\min_{\mathbf{w}} \lambda_G(\mathbf{w}) + C \times R_{emp}(\mathbf{w}, D_{train})$$

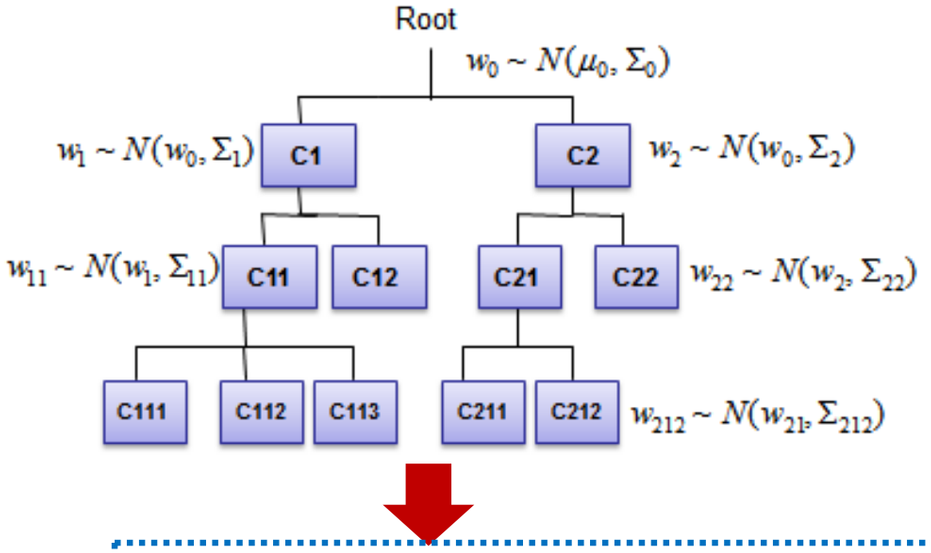
$$where \quad \lambda_G(\mathbf{W}) = \sum_{(i.j)\in E} \| w_i - w_j \|^2$$

graph-based regularization term

[slide credit: Yiming Yang]

Adapting Models to New Languages

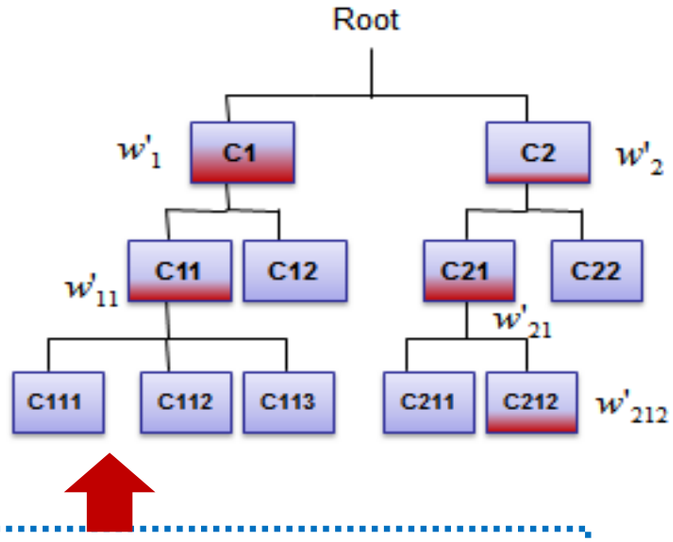
A hierarchy of classification models in L1

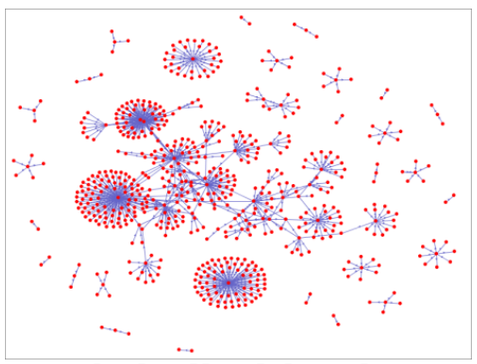


?

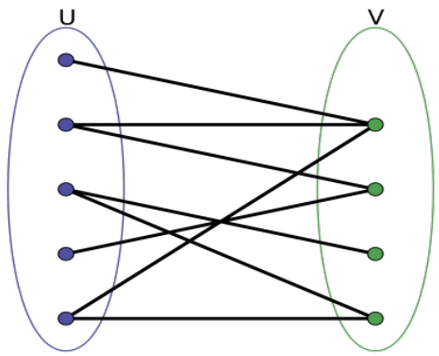
Model Translation

Partially translated classification models in L2

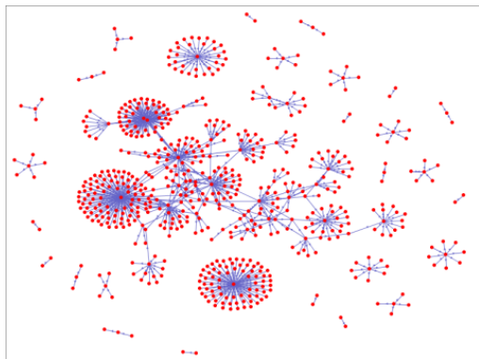




Word similarities in L1 (induced via word2vec)



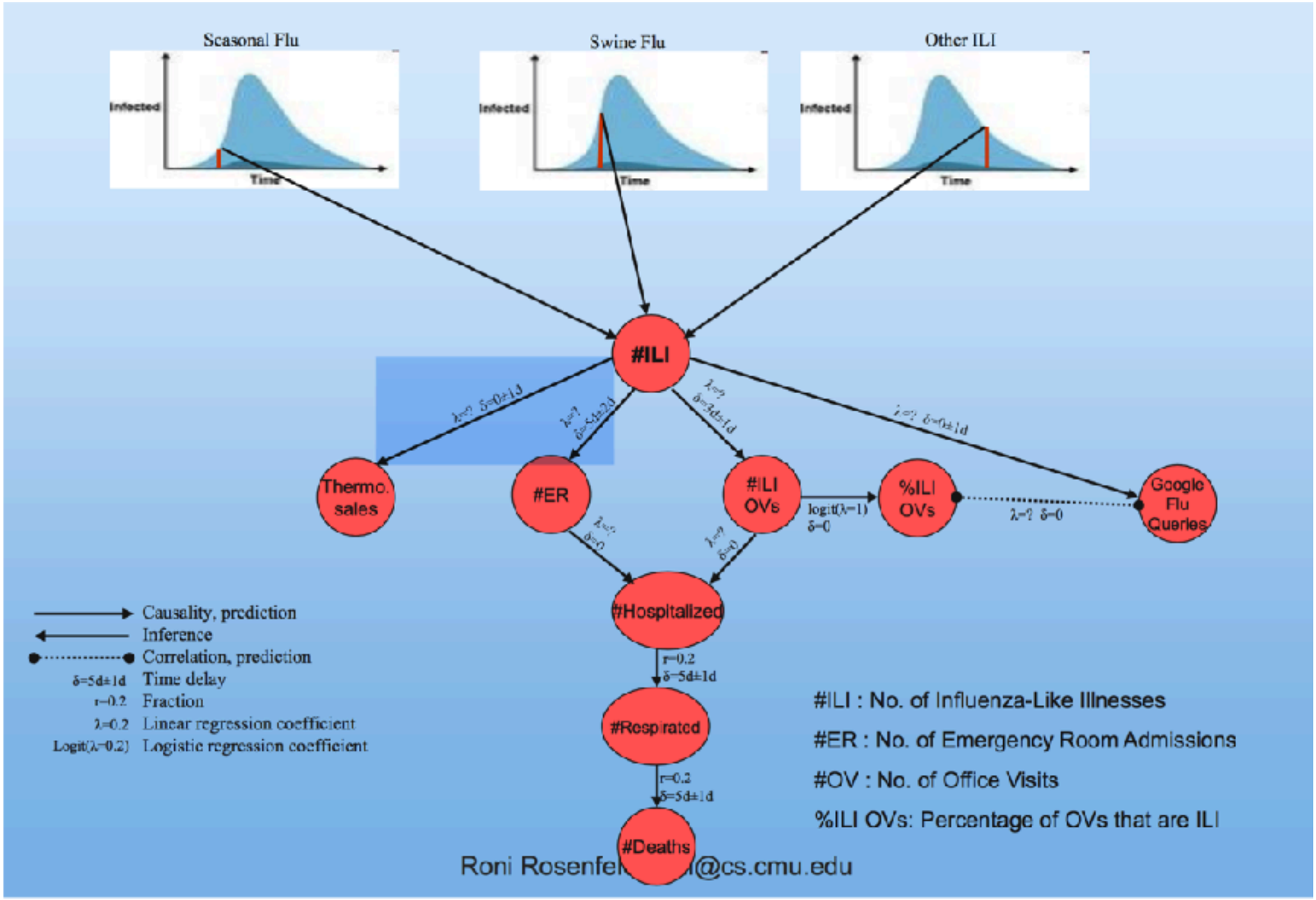
Bilingual Dictionary + Transductive Link prediction



Word similarities in L2 (induced via word2vec)

[slide credit: Yiming Yang]

Computational Epidemiology



[slide credit: Roni Rosenfeld]

Unsupervised Learning

Historical Document

Old Bailey Court Proceedings 1775

the prisoner at the bar. Jacob Lazarus and his wife, the prisoner, were both together when I received them. I sold eleven pair of them for three guineas, and delivered the remainder back to the prisoner. I sold seven pair of silk to Mark Simper: one pair of mixed, and two pair of thread to the footman, and one pair of thread to the barber.

Q. What is the footman's name?

Frances Moses. I don't know.

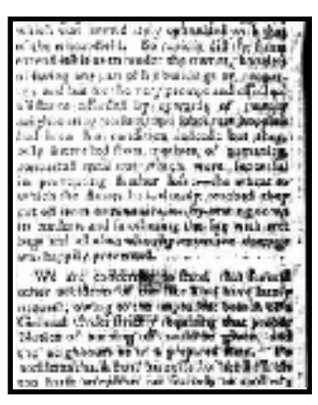
Henry Harris. I was standing at the Compter waiting for the sherriff's officers to employ me: Moses's daughter came for me to go and take the prisoner. I went to the Old Bailey

Unknown Fonts

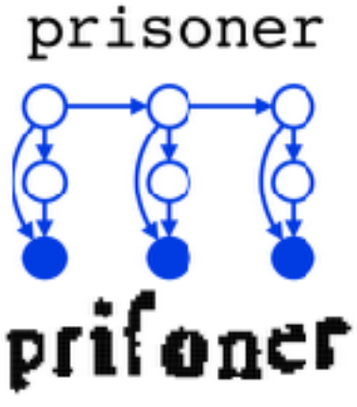
Unsupervised Transcription

Learning

Doc Image



Generative Model

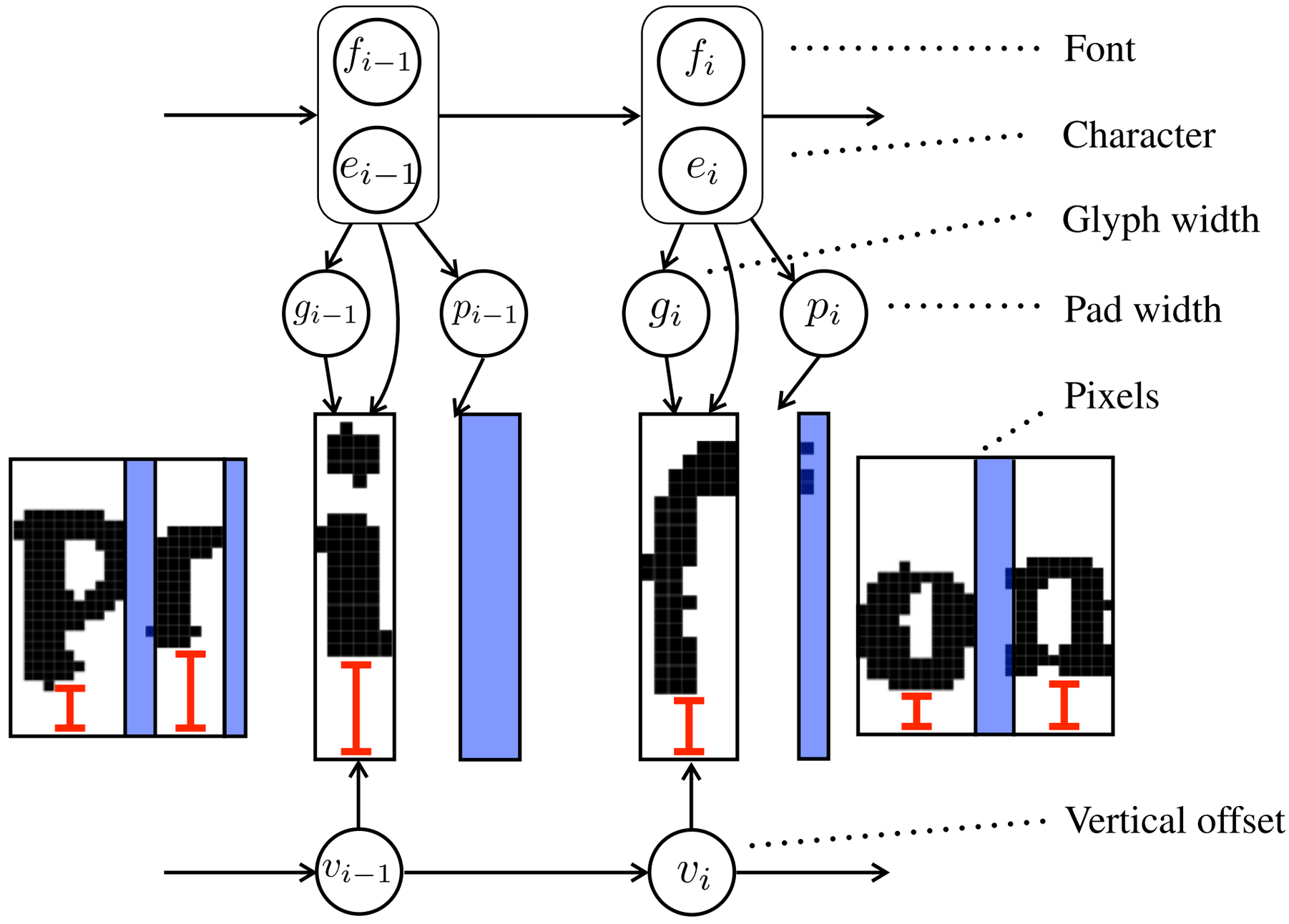


Font Parameters

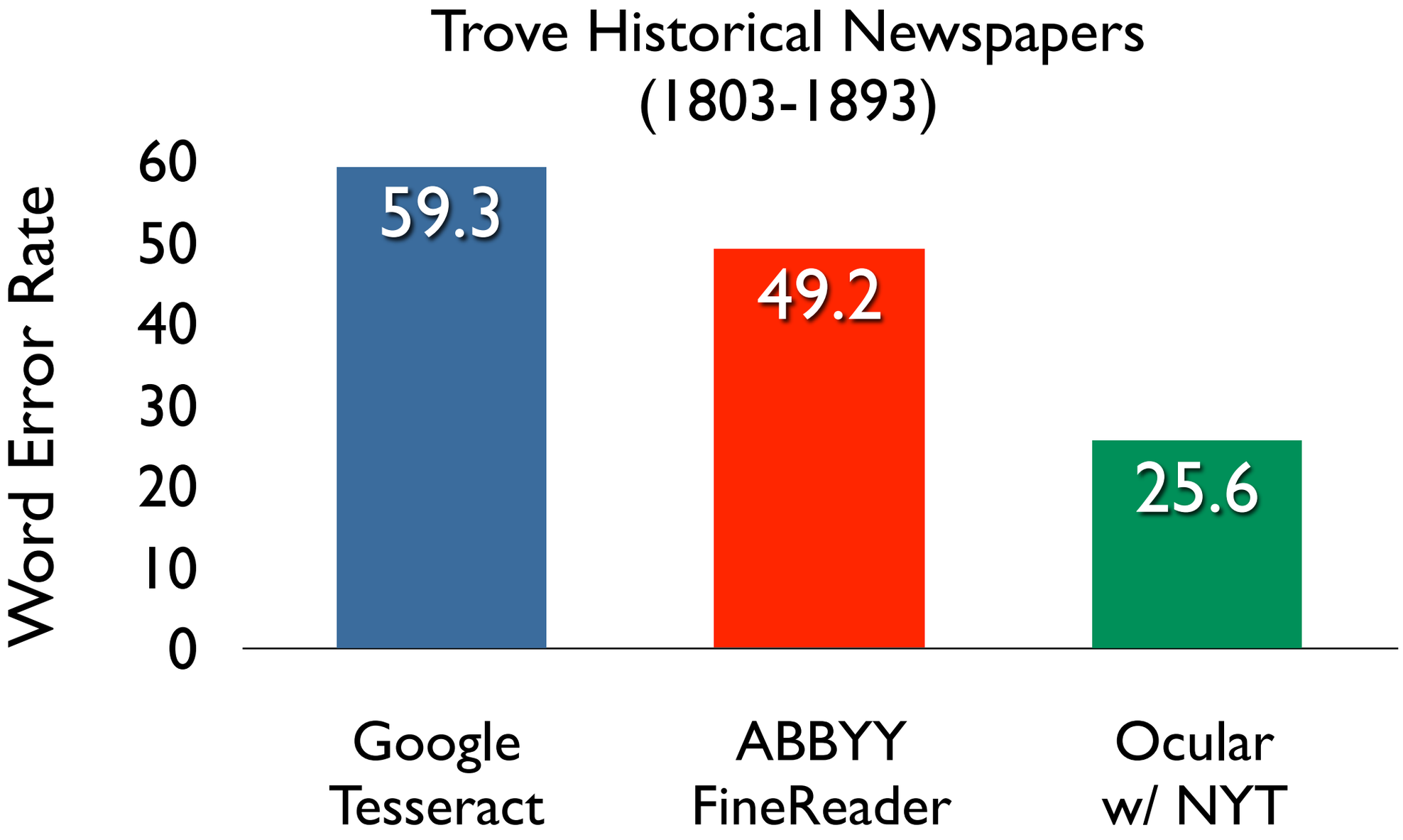
Text

which was immediately connected with that of the wheat-field. So Pollamounter came into the Kitchen. The little Girl went away, and the Prisoner and Cornish went out together to see for Peg, but while they were gone, Peg came in again by her self and asked Mrs. Provost, the Bar-keeper, where

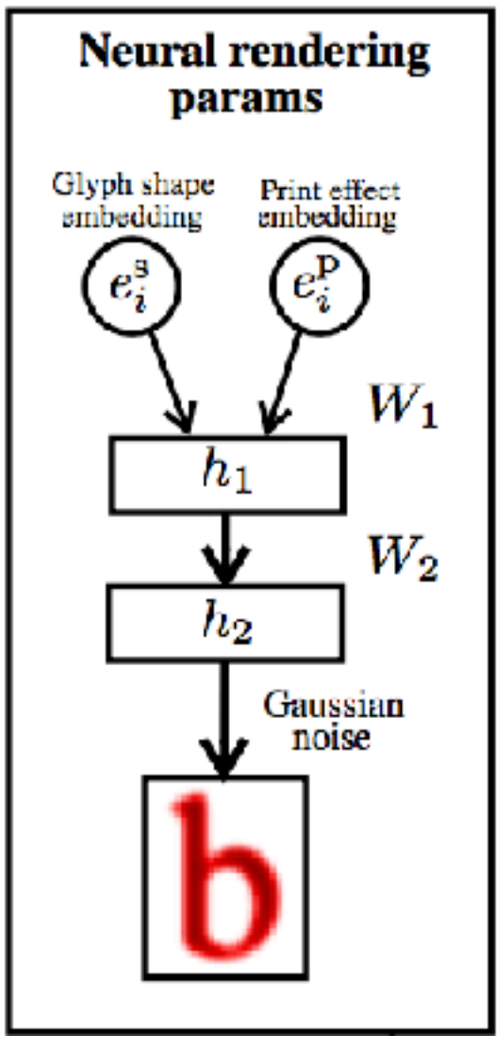
Multiple Fonts / Markov Offsets

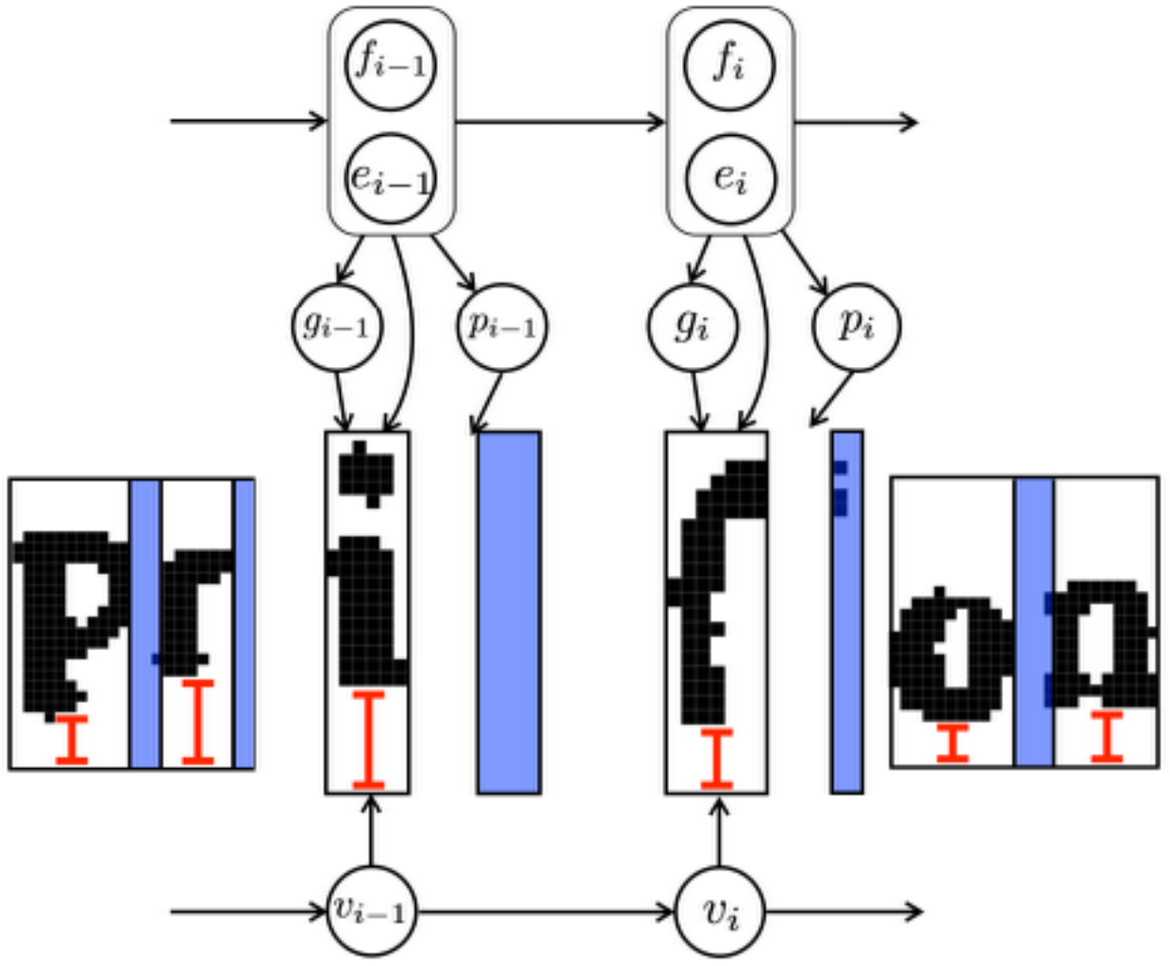


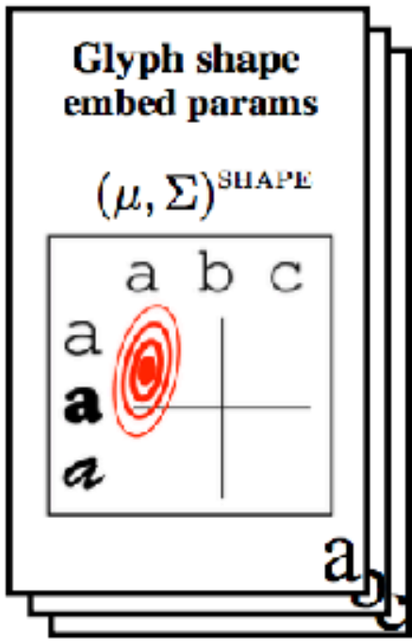
Results

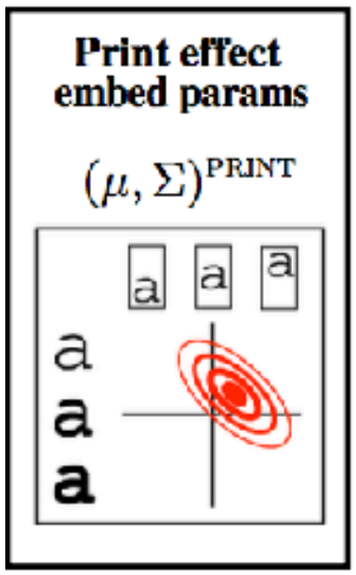

Trove Historical Newspapers
(1803-1893)
Word Error Rate
60
50
40
30
20
10
0
59.3
49.2
25.6
Google
Tesseract
ABBYY
FineReader
Ocular
w/ NYT

Shape Embeddings

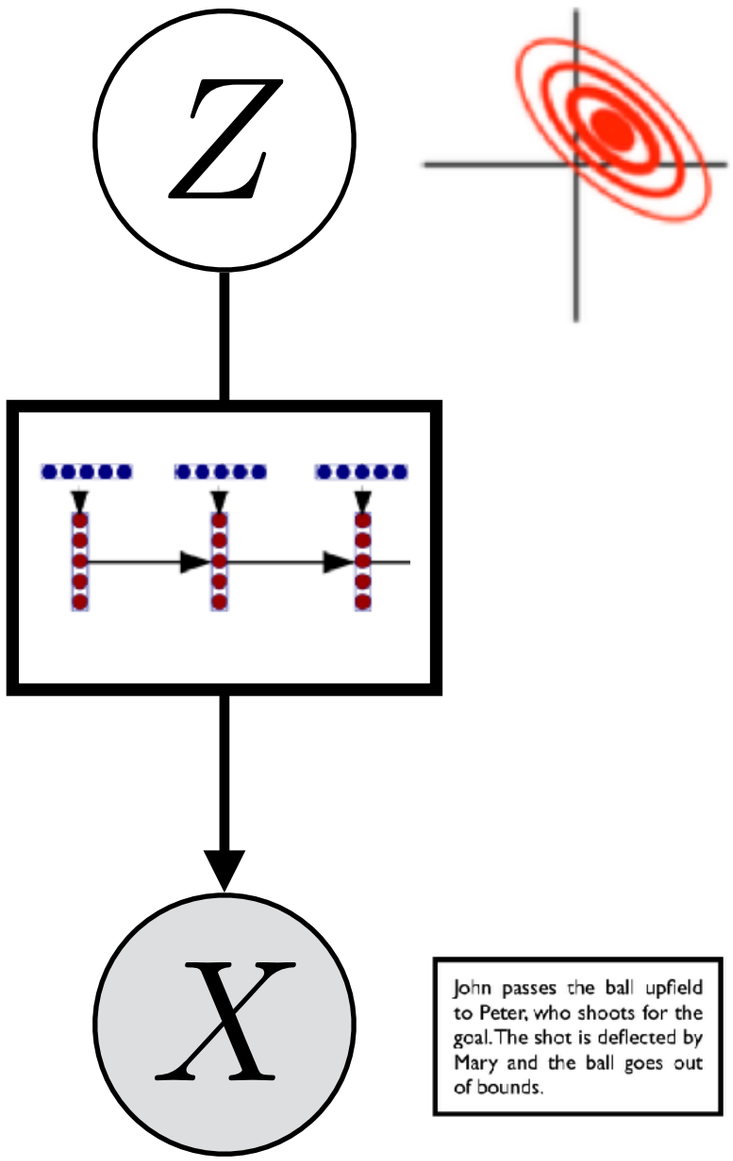




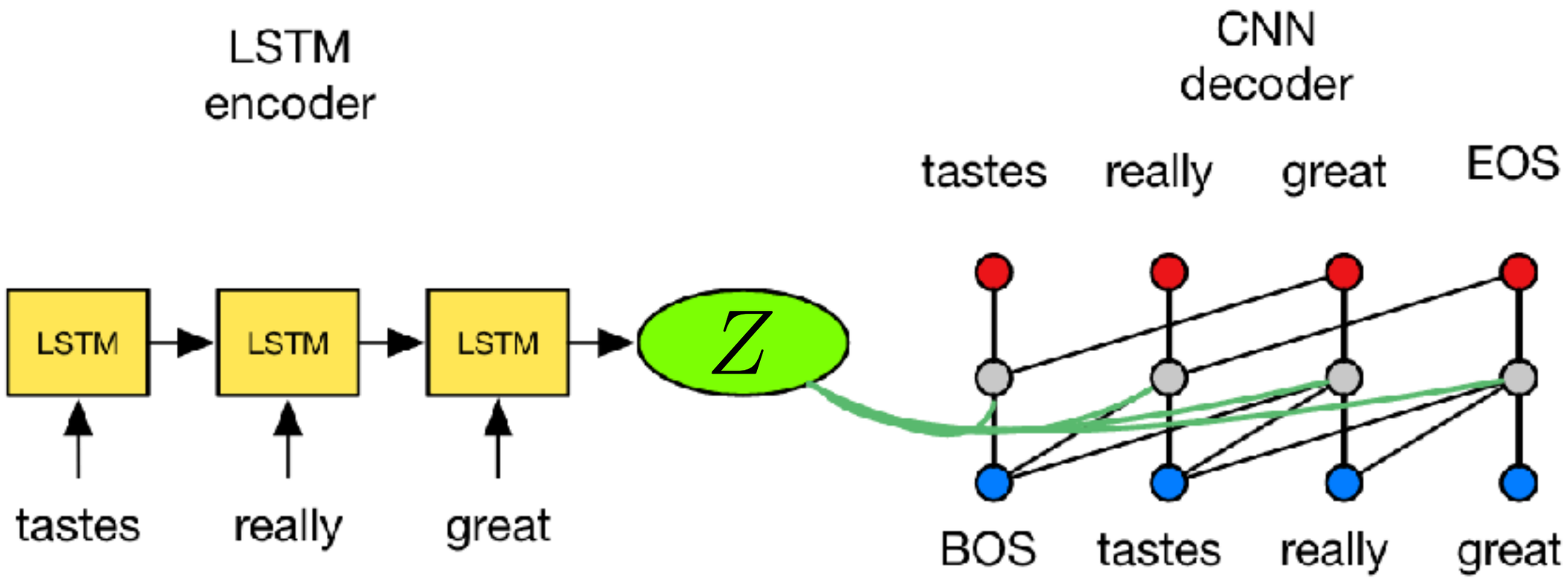




VAEs for Text

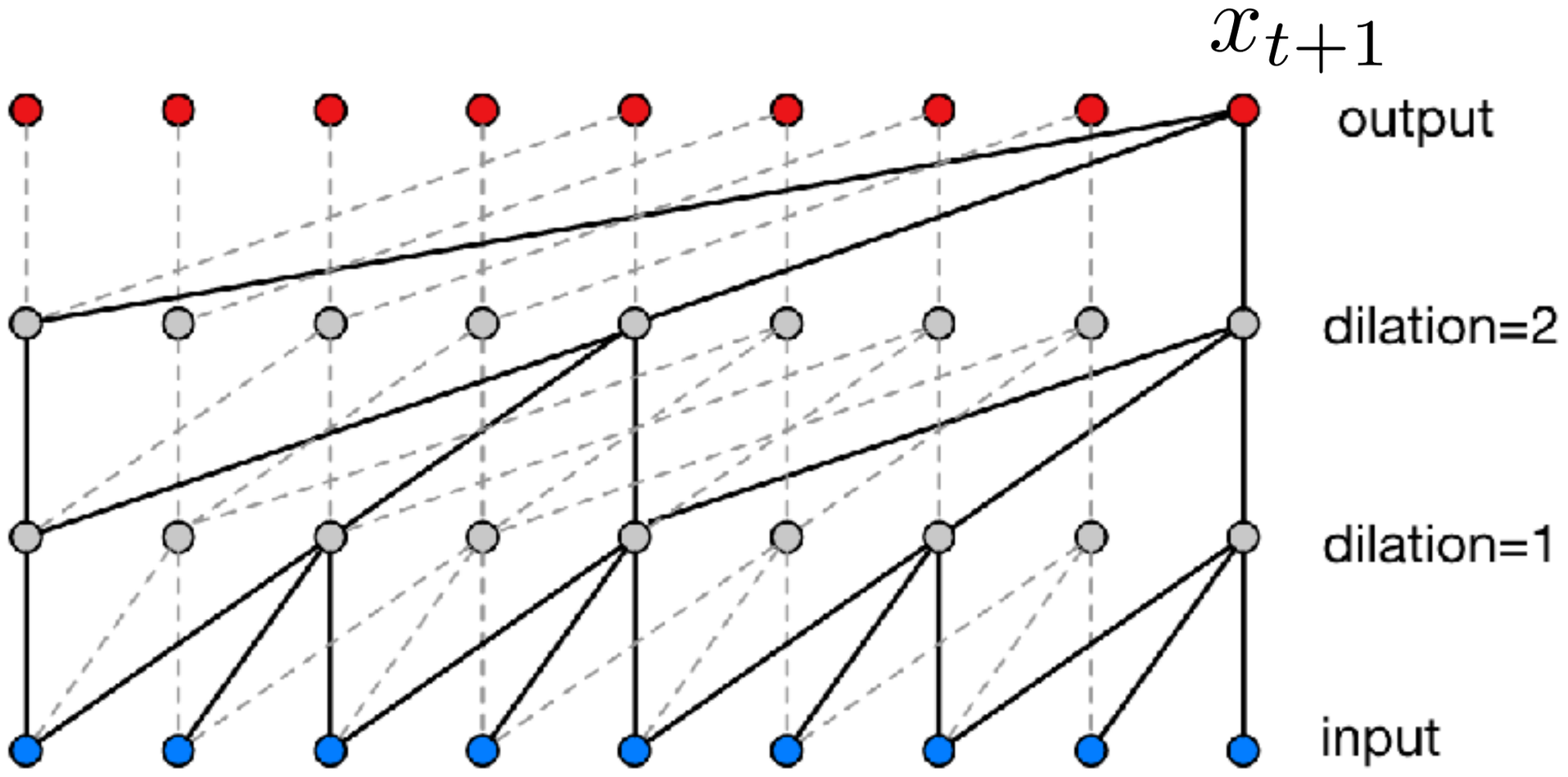


VAEs for Text



X X

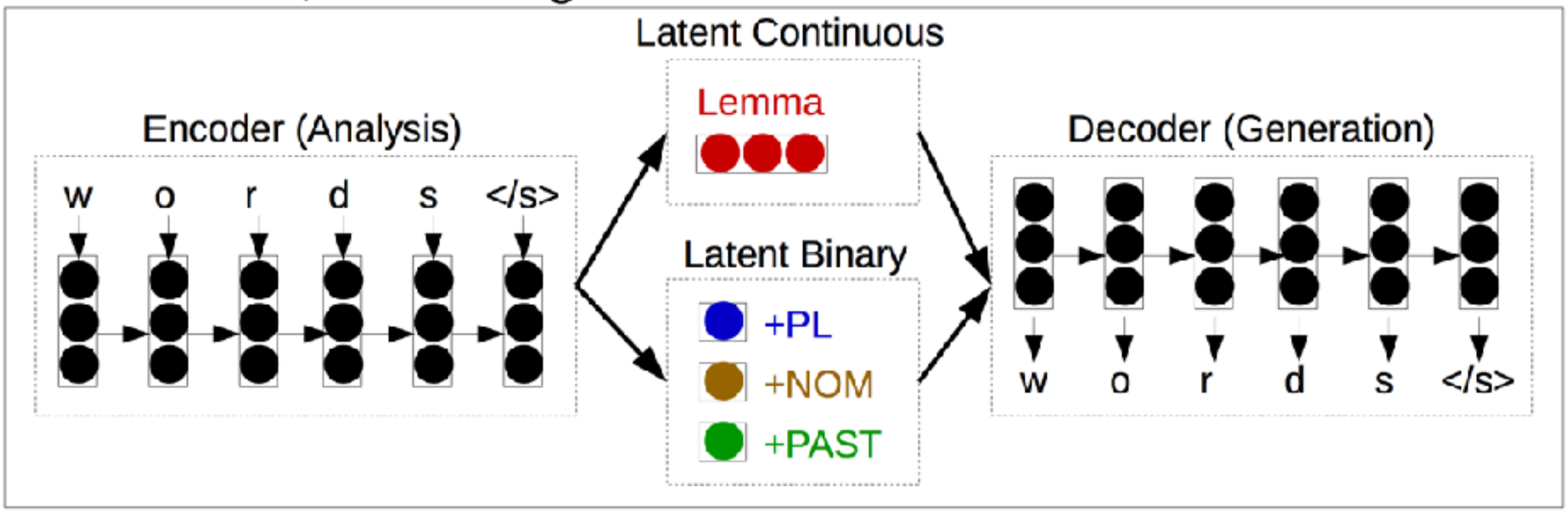
Dilational CNN Decoder



$x_1, \ldots, x_t$

Labeled Sequence Transduction

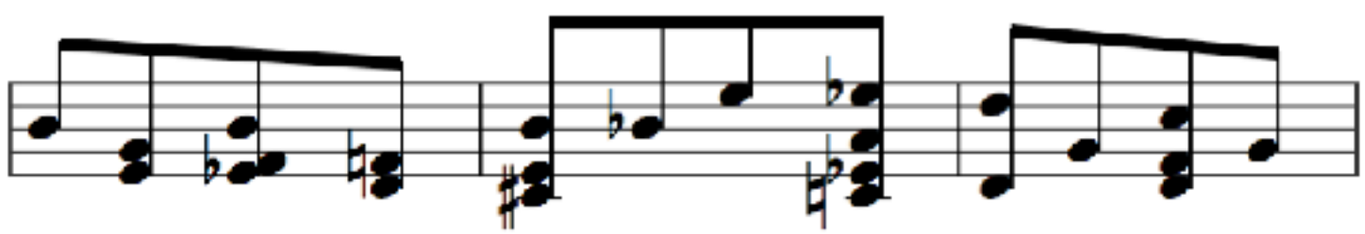
- Problem: We want to transform an input into another input, where output is controlled by tags
- Solution: Tags/content as discrete/continuous latent variables, train using method called variational auto-encoder



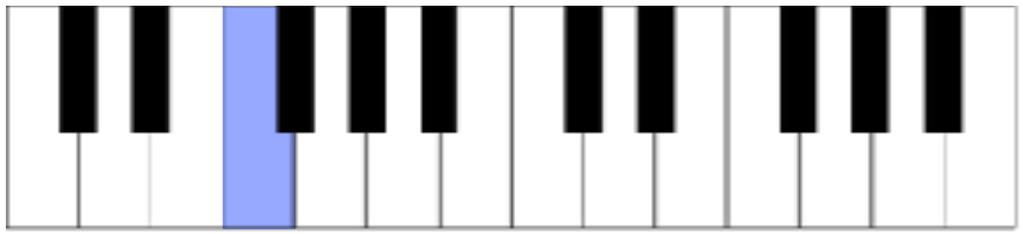
- Result: Improvement of translation accuracy, particularly in low-data scenarios

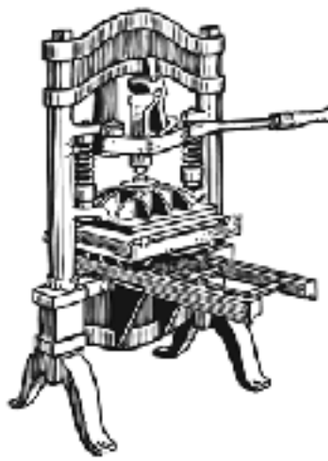
[slide credit: Graham Neubig]

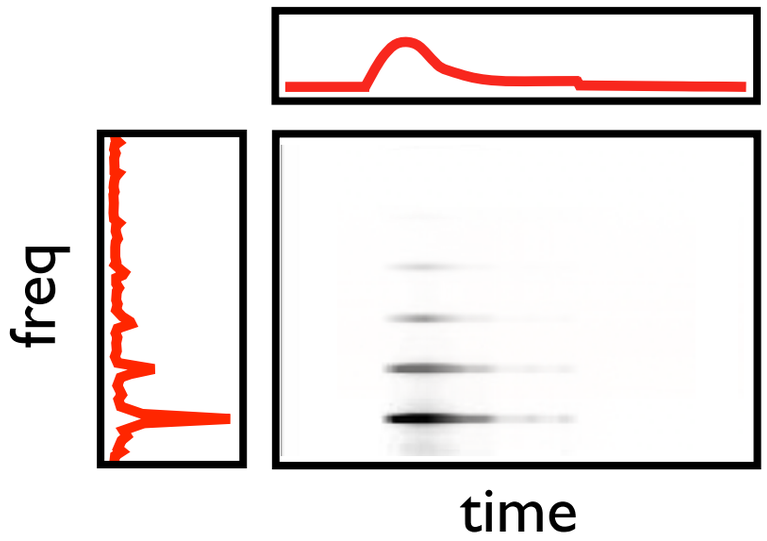
Piano Music Transcription

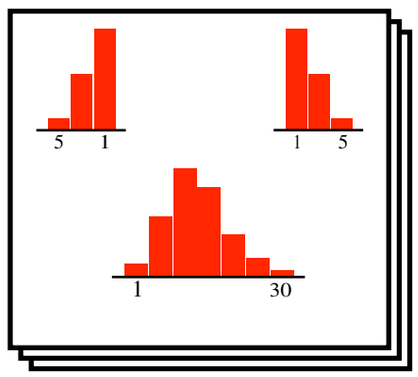


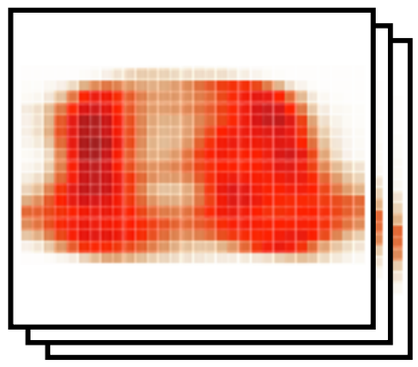
Parameters



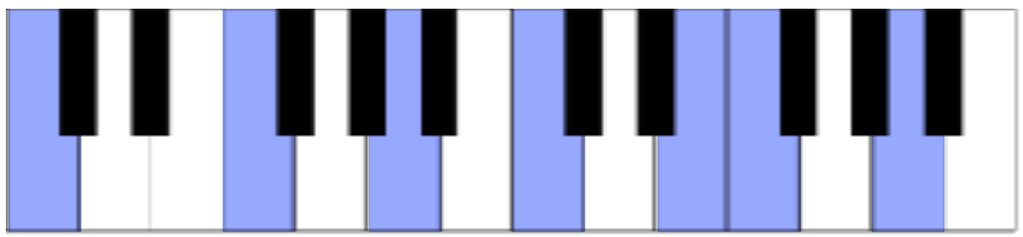


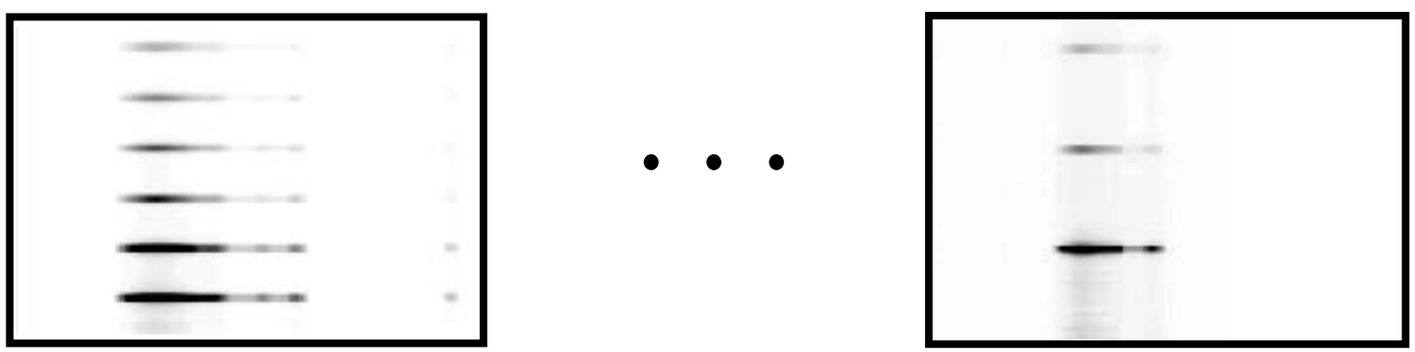




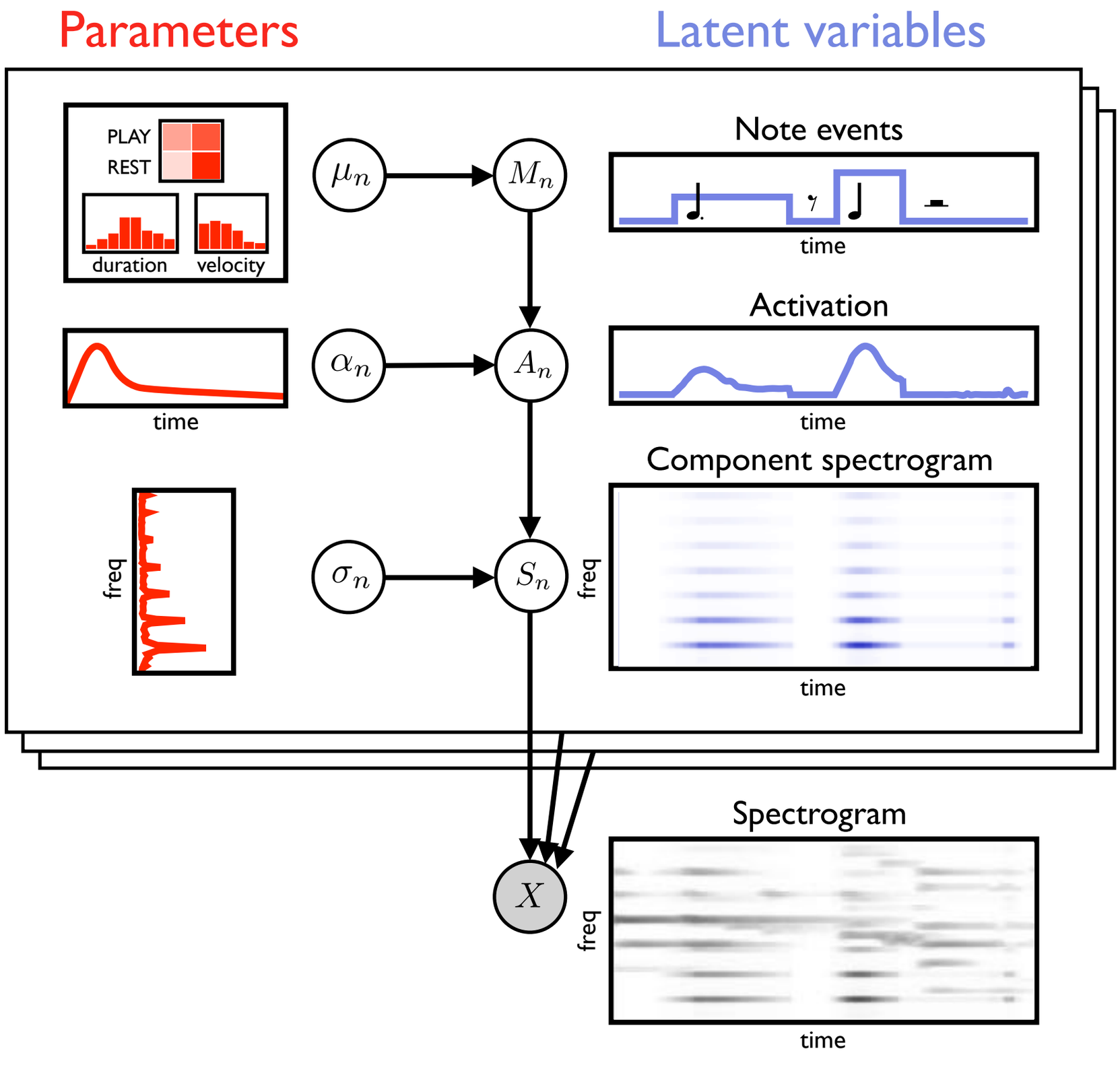


Polyphony

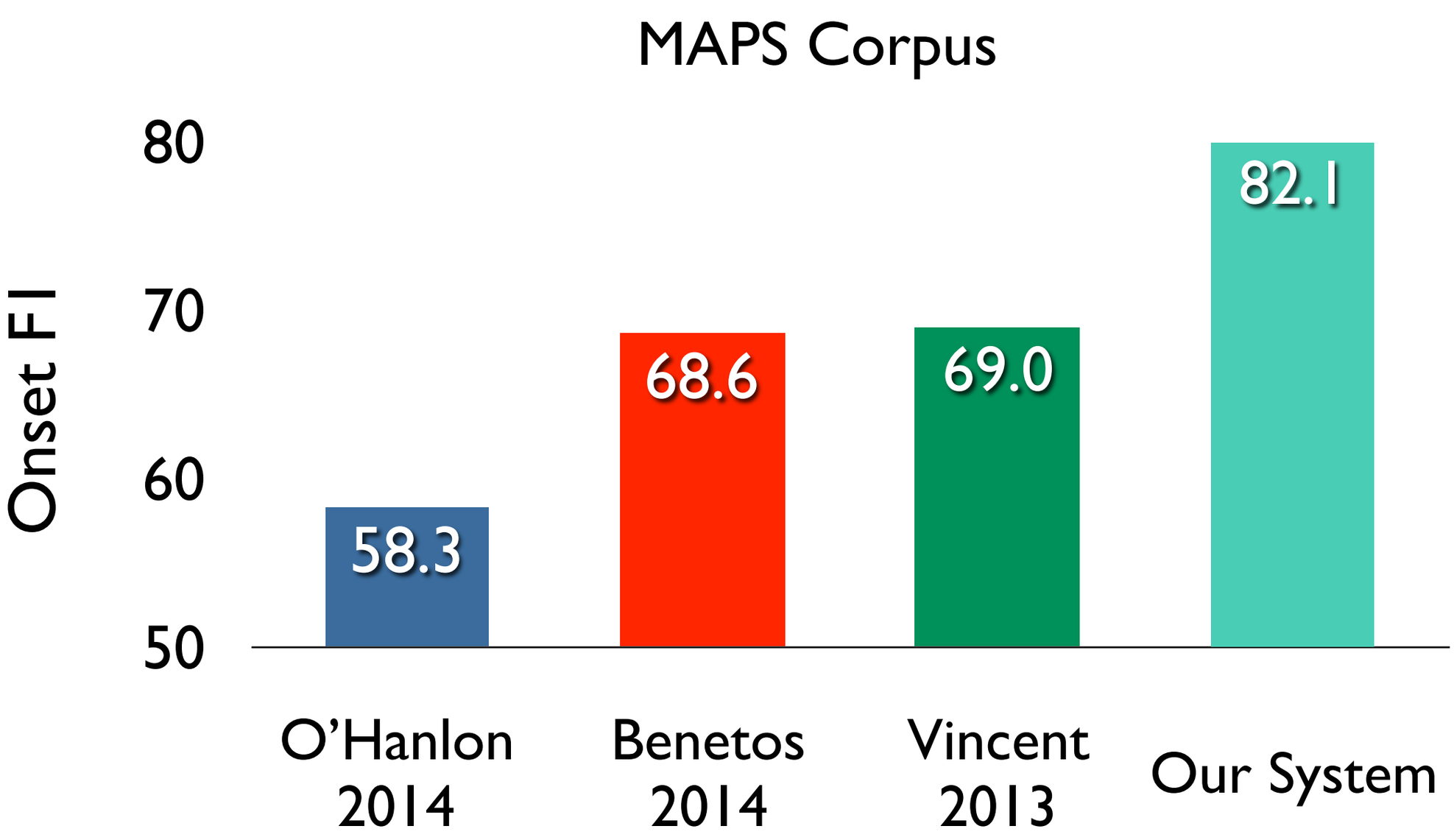




Generative Model



Results


MAPS Corpus
Onset F1
80
70
60
50
58.3
68.6
69.0
82.1
O'Hanlon 2014
Benetos 2014
Vincent 2013
Our System

Demo!

https://github.com/tberg12/klavier