

Introduction of Structured Learning

Hung-yi Lee

input/output包含structure的東西如tree,list...

Structured Learning

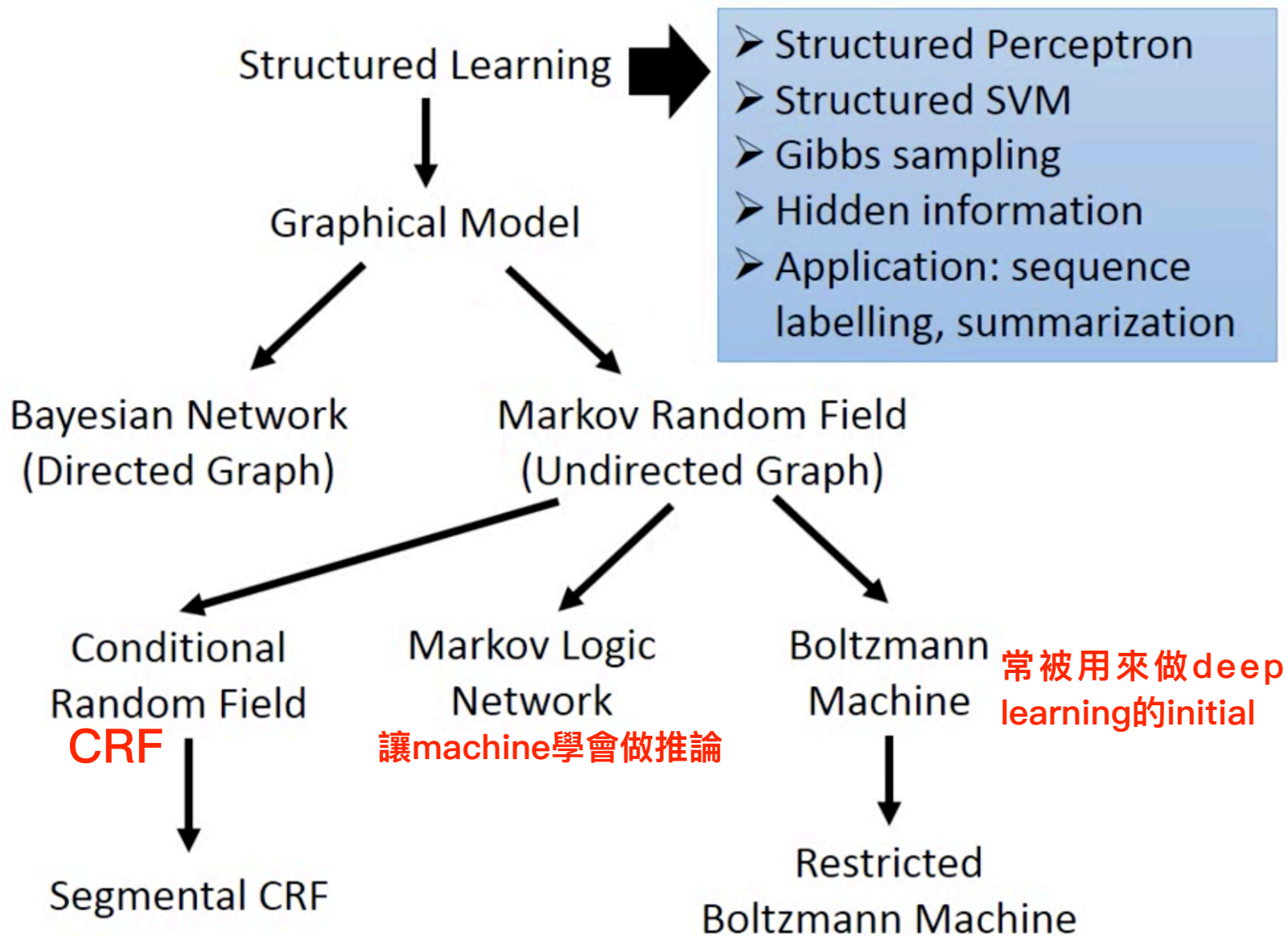
- We need a more powerful function f
 - Input and output are both objects with structures
 - *Object*: sequence, list, tree, bounding box ...

$$f : X \rightarrow Y$$

有structure的！
 X is the space of
one kind of object

有structure的！
 Y is the space of
another kind of object

In the previous lectures, the input and output are both vectors.



Introduction of Structured Learning Unified Framework

Unified Framework

Training

- Find a function F

$$F : X \times Y \rightarrow \mathbb{R}$$

R被用來evaluate X,Y
有多適合放在一起

- $F(x,y)$: evaluate how compatible the objects x and y is

Inference (Testing)

- Given an object x

$$\tilde{y} = \arg \max_{y \in Y} F(x, y)$$

因為在test的時候還要解一個optimization的problem

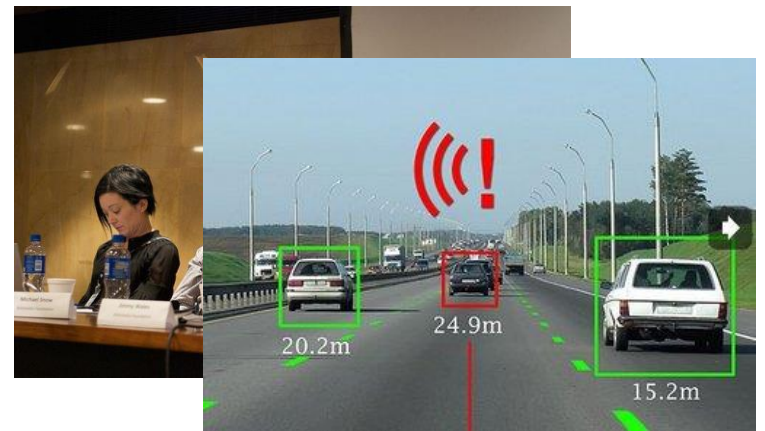
$$f : X \rightarrow Y \quad \Rightarrow \quad f(x) = \tilde{y} = \arg \max_{y \in Y} F(x, y)$$

窮舉所有 Y 找出哪個最適合output

Unified Framework – Object Detection

- Task description

- Using a bounding box to highlight the position of a certain object in an image
- E.g. A detector of Haruhi



X : Image \longrightarrow Y : Bounding Box



Haruhi

(the girl with
yellow ribbon)

Unified Framework – Object Detection

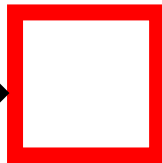
Training

- Find a function F
$$F: X \times Y \rightarrow \mathbb{R}$$
- $F(x,y)$: evaluate how compatible the objects x and y is

x : Image \rightarrow



y : Bounding Box \rightarrow



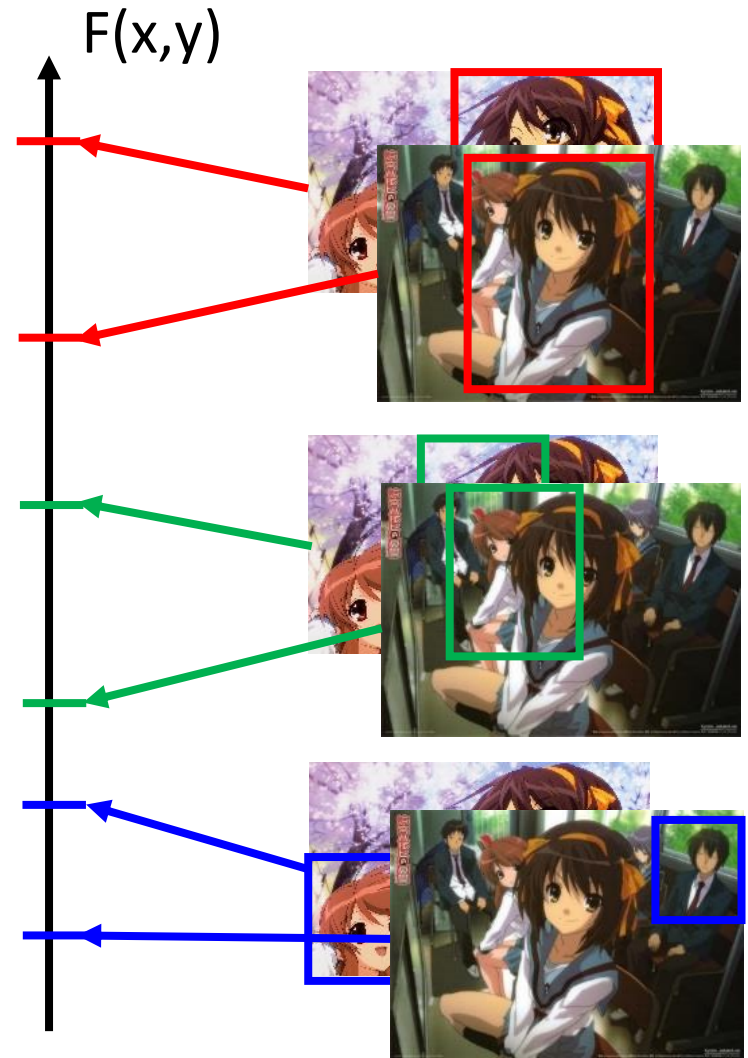
F 會評定這件事
情會有多正確

$F(x,y) \rightarrow F($



)

the correctness of taking
range of y in x as “Haruhi”



Unified Framework – Object Detection

Training

- Find a function F
$$F: X \times Y \rightarrow \mathbb{R}$$
- $F(x, y)$: evaluate how compatible the objects x and y is

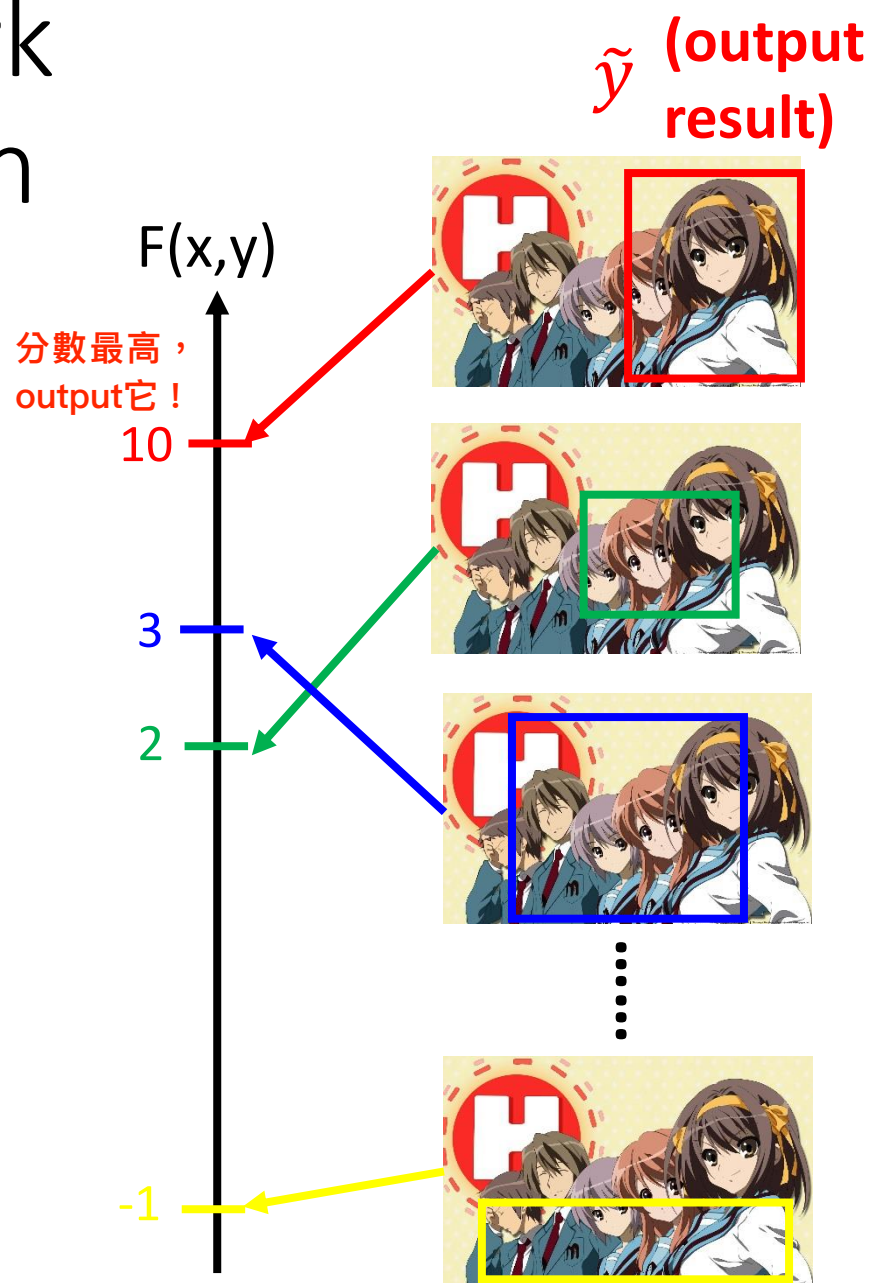
Inference (Testing)

- Given an object x
$$\tilde{y} = \arg \max_{y \in Y} F(x, y)$$

input $x =$



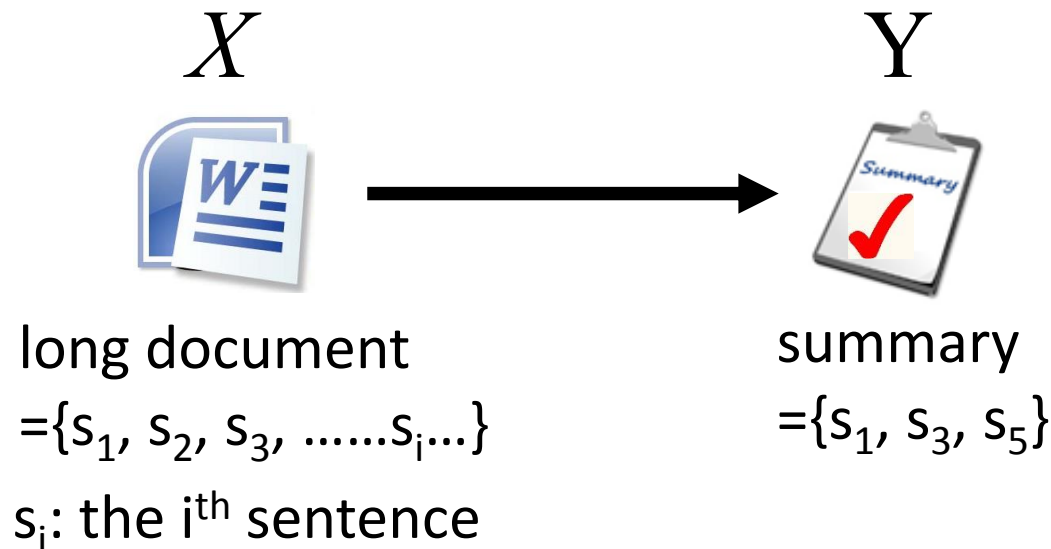
Enumerate all possible
bounding box y



Unified Framework

- Summarization

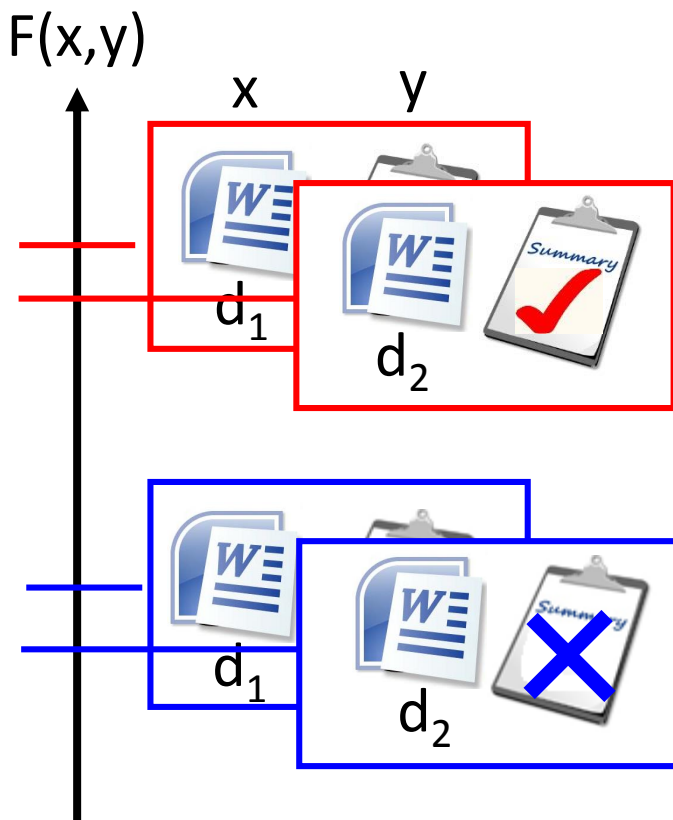
- Task description
 - Given a long document
 - Select a set of sentences from the document, and cascade the sentences to form a short paragraph



Unified Framework

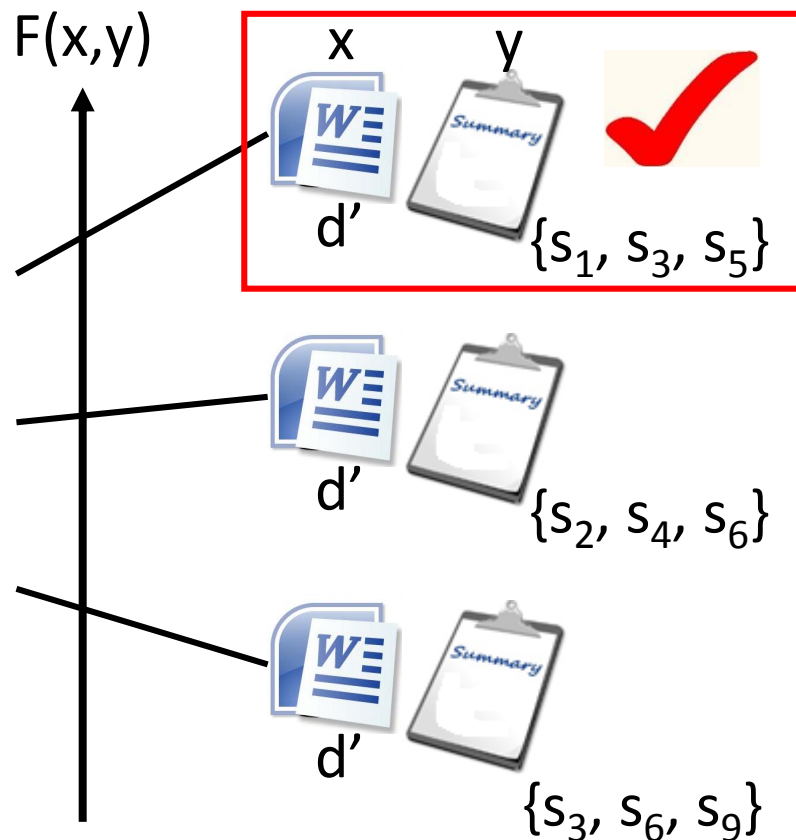
- Summarization

Training



窮舉所有可能的sentence set，看哪個set跟document和起來的分數最高就output它

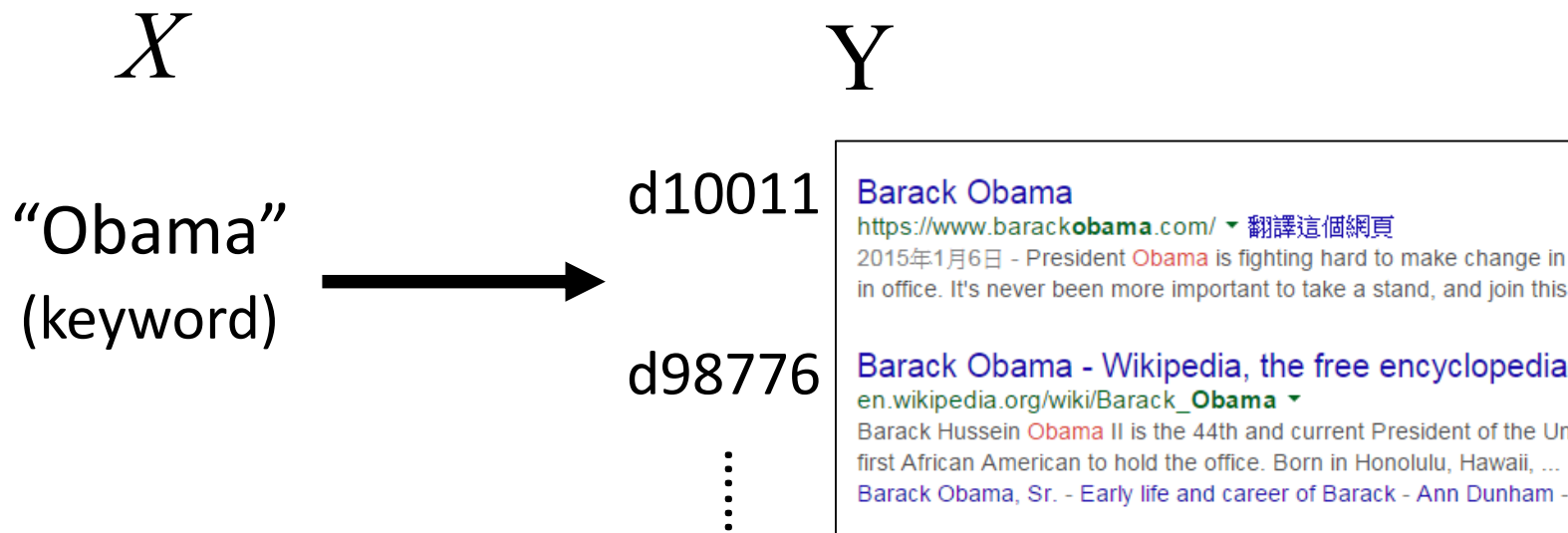
Inference



Unified Framework

- Retrieval

- Task description
 - User input a keyword Q
 - System returns a ***list*** of web pages



A list of web pages (Search Result)

Unified Framework

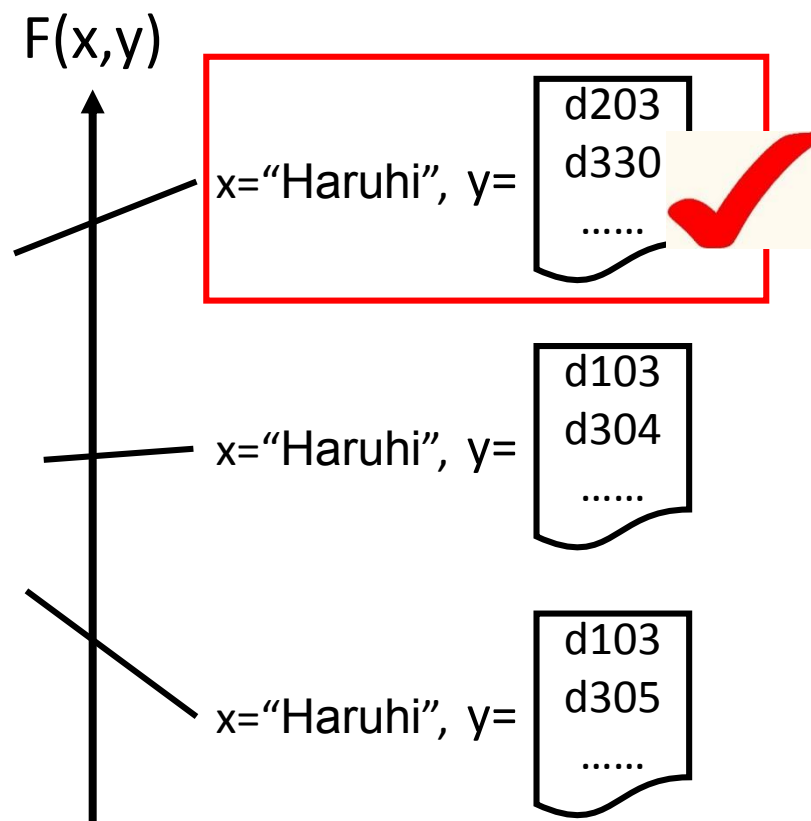
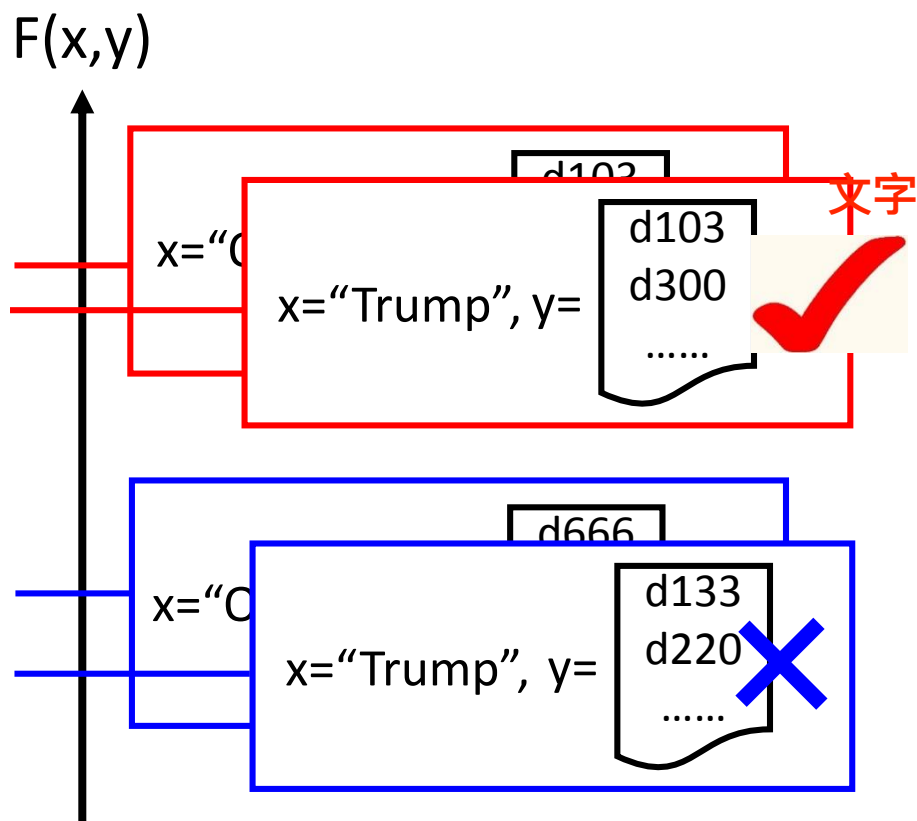
- Retrieval

web mining (retrieval) : learning to rank

list-wise gradient descent就是這個架構

Training

Inference



Statistics

Unified Framework

Training

- Find a function F

$$F : X \times Y \rightarrow \mathbb{R}$$

- $F(x,y)$: evaluate how compatible the objects x and y is

描述成機率

Training

- Estimate the probability

$$P(x,y)$$

joint probability

$$P : X \times Y \rightarrow [0,1]$$

Inference

- Given an object x

$$\tilde{y} = \arg \max_{y \in Y} P(y | x)$$

$$= \arg \max_{y \in Y} \frac{P(x, y)}{P(x)}$$

$P(x)$ 是固定的因此不需要比較

$$= \arg \max_{y \in Y} P(x, y)$$

Inference

- Given an object x

$$\tilde{y} = \arg \max_{y \in Y} F(x, y)$$

描述成機率

$$F(x, y) = P(x, y)?$$

Statistics

Unified Framework

$$F(x, y) = P(x, y)?$$

Drawback for probability

- Probability cannot explain everything
 x:關鍵字
y:搜尋結果
x,y一起出現很怪
- 0-1 constraint is not necessary
 機率和=1，但這是不必要的

Strength for probability

- Meaningful
 輸出有意義

Energy-based Model:
<http://www.cs.nyu.edu/~yann/research/ebm/>

Training

- Estimate the probability $P(x, y)$
聲音 文字
 $P: X \times Y \rightarrow [0, 1]$

Inference

- Given an object x

$$\begin{aligned}\tilde{y} &= \arg \max_{y \in Y} P(y | x) \\ &= \arg \max_{y \in Y} \frac{P(x, y)}{P(x)} \\ &= \arg \max_{y \in Y} P(x, y)\end{aligned}$$

Unified Framework

That's it!?

Training

- Find a function F

$$F : X \times Y \rightarrow \mathbb{R}$$

- $F(x,y)$: evaluate how compatible the objects x and y is

Inference (Testing)

- Given an object x

$$\tilde{y} = \arg \max_{y \in Y} F(x, y)$$

There are three problems in this framework.

Problem 1

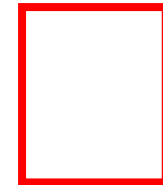
- **Evaluation:** What does $F(x,y)$ look like?
 - How $F(x,y)$ compute the “compatibility” of objects x and y

Object Detection:

$F(x=$



, $y=$



)

Summarization:

$F(x=$



(a long document)

, $y=$



(a short paragraph)

Retrieval:

$F(x= \text{“Obama”}$
(keyword)

, $y=$

Barack Obama
<https://www.barackobama.com/> - 翻譯這個網頁
2015年1月6日 - President Obama is fighting hard to make change in his final two years in office. It's never been more important to take a stand, and join this ...
Barack Obama - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Barack_Obama -
Barack Hussein Obama II is the 44th and current President of the United States, and the first African American to hold the office. Born in Honolulu, Hawaii, ...
Barack Obama, Sr. - Early life and career of Barack - Ann Dunham - Michelle Obama

(Search Result)

Problem 2

- **Inference:** How to solve the “arg max” problem

$$y = \arg \max_{y \in Y} F(x, y)$$

The space Y can be extremely large!

Object Detection: Y =All possible bounding box (maybe tractable)

Summarization: Y =All combination of sentence set in a document ...

Retrieval: Y =All possible webpage ranking

Problem 3

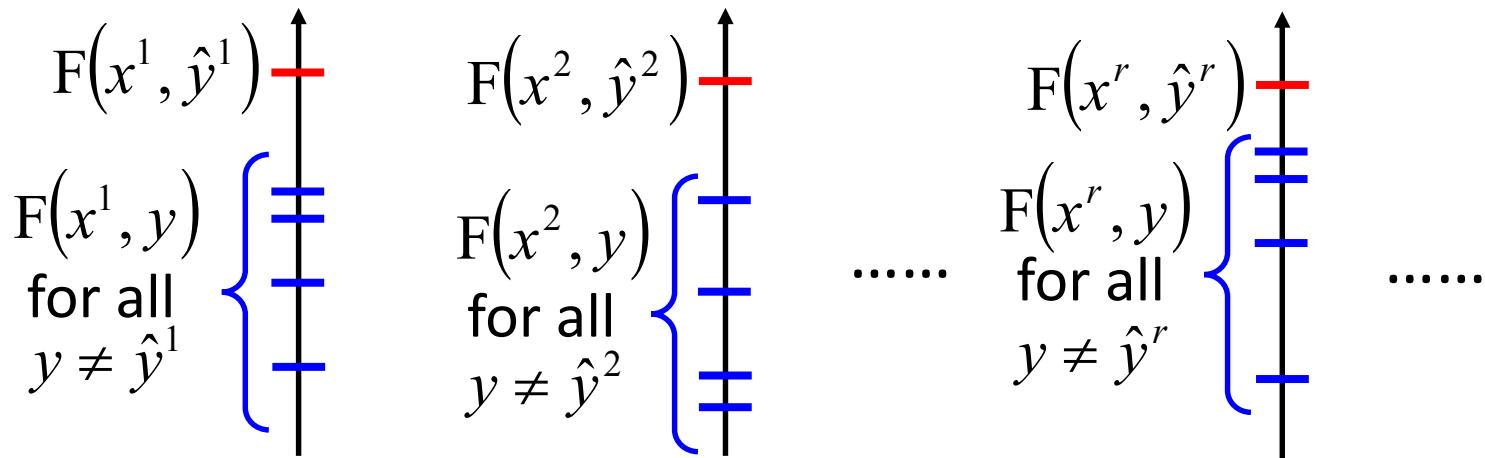
- **Training**: Given training data, how to find $F(x,y)$

Principle

Training data: $\{(x^1, \hat{y}^1), (x^2, \hat{y}^2), \dots, (x^r, \hat{y}^r), \dots\}$

We should find $F(x,y)$ such that

滿足以下條件



Three Problems

Problem 1: Evaluation

- What does $F(x,y)$ look like?



Problem 2: Inference

- How to solve the “arg max” problem

$$y = \arg \max_{y \in Y} F(x, y)$$



Problem 3: Training

- Given training data, how to find $F(x,y)$

