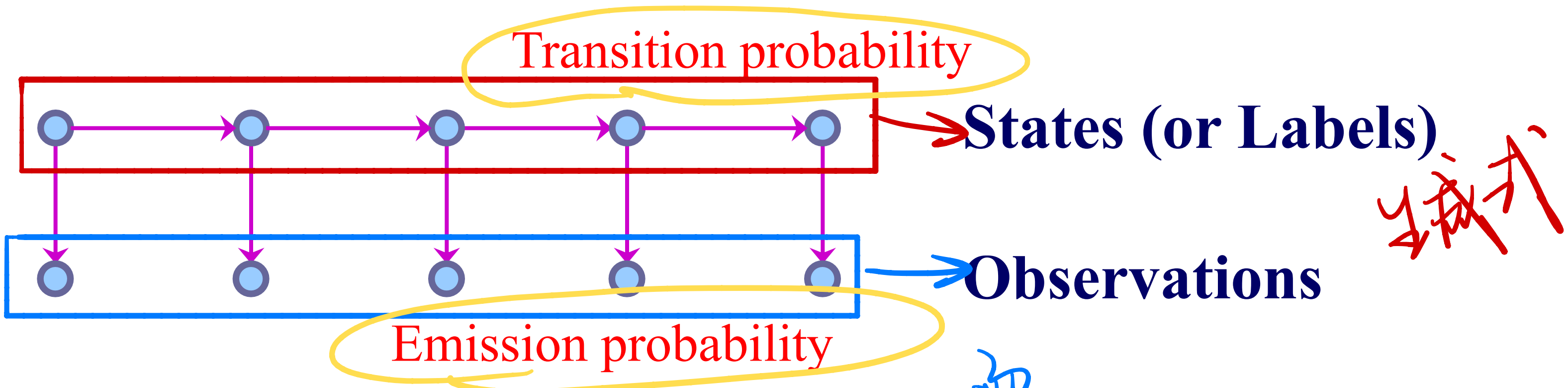
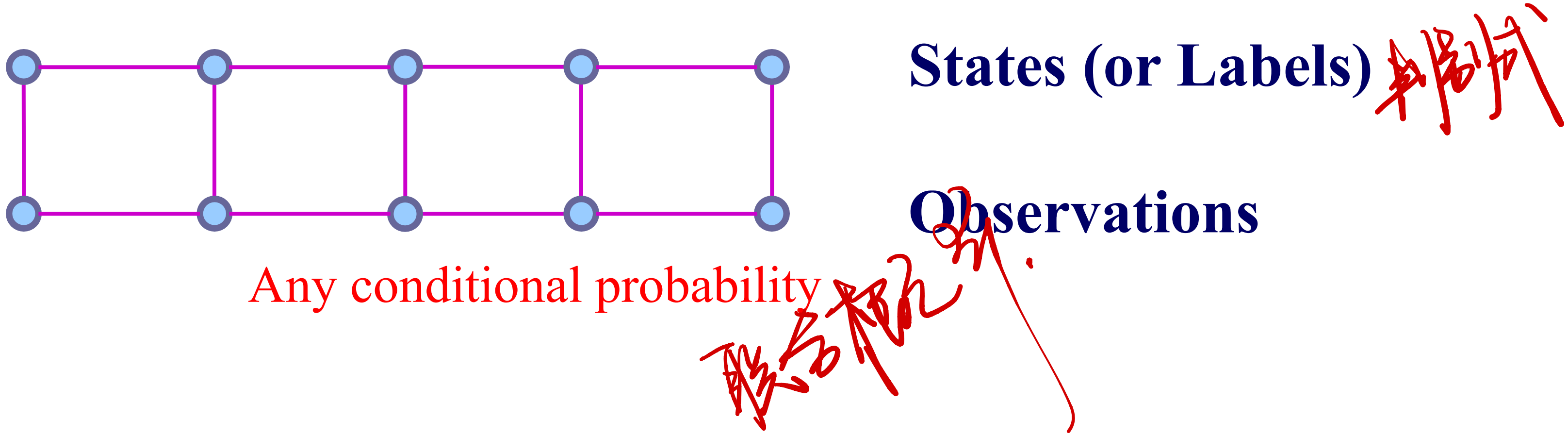
HMM vs. CRF

- Hidden Markov Model



- Conditional Random Fields 条件随机场模型



Conditional Random Fields

- CRFs model can incorporate ***arbitrary features*** and domain knowledge in a flexible and straightforward way.
- The choice of features is a ***critical***, but ***very hard*** problem.

任何

规则

窗口为5 → 5个特征的(feature)

从 小 学 计 算 机

C_{-2} C_{-1} C_0 C_1 C_2

Windows

用这个window 来预测 C_0 的标签 BEIS

Predict

The tag of C_0

Feature templates:

特征模板.

C_0	**学**
C_{-1}	**小**
C_1	**计**
C_{-2}	**从**
C_2	**算**
C_0C_1	**学计**
$C_{-1}C_0$	**小学**
$C_{-1}C_1$	**小计**

Feature templates for Chinese word segmentation

Example

Corpus 训练集

S ***S*** ***S***

他 恨 她

Test 测试集

S ***S*** ***S***

他 爱 她

Feature template

C_0

Features

$f(S_0 = S, C_0 = \text{他}) = 1$

$f(S_0 = S, C_0 = \text{恨}) = 1$

$f(S_0 = S, C_0 = \text{她}) = 1$

	他	爱	她
	C_0	C_0	C_0
B	**0**	**1/*S***	**1/*B***
E	**0**	**1/*S***	**1/*B***
I	**0**	**1/*S***	**1/*B***
S	**1**	**1/*S***	**2/*B***
	S	***B***	***S***

Example

Corpus

S ***S*** ***S***

他 恨 她

Feature template

C_0 $S_{-1}S_0$

Features

$f(S_0 = S, C_0 = 他) = 1$

$f(S_0 = S, C_0 = 恨) = 1$

$f(S_0 = S, C_0 = 她) = 1$

$f(S_{-1} = S, S_0 = S) = 2$

$f(S_{-1} = NIL, S_0 = S) = 1$

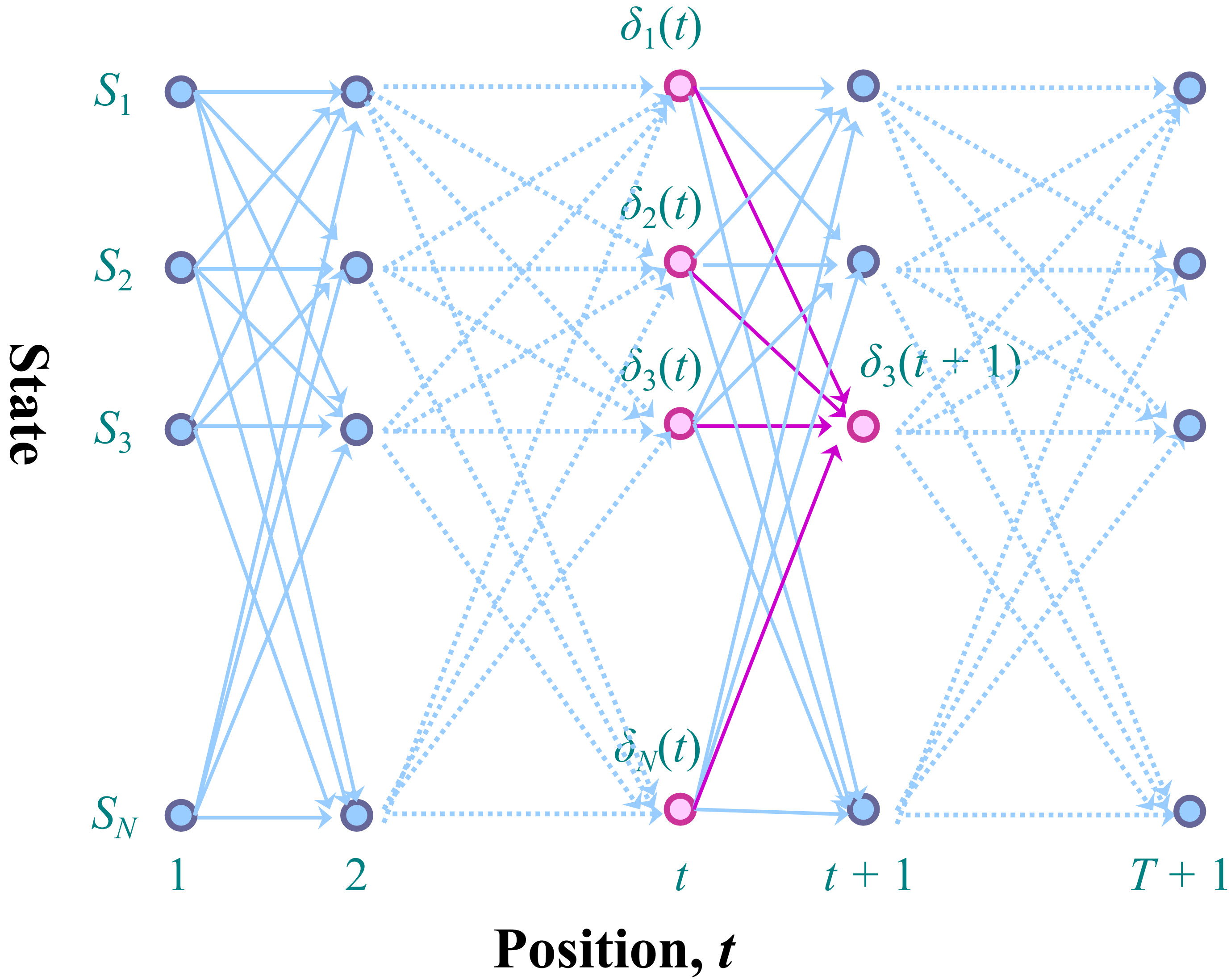
Test

S ***S*** ***S***

他 爱 她

	他	爱	她
	C_0	C_0	C_0
B	**0**	**2/*S***	**4/*S***
E	**0**	**2/*S***	**4/*S***
I	**0**	**2/*S***	**4/*S***
S	**2**	**4/*S***	**7/*S***
	S	*S*	*S*

Viterbi algorithm



Conditional random fields

observation

S	**B**	**E**	**S**	**B**	**E**	**S**	**S**	**B**	**E**	**S**
他	**从**	**小**	**就**	**开**	**始**	**学**	**画**	**动**	**物**	**。**

$f(S_0 = S, O_0 = 就)$

$f(S_0 = S, O_{-1} = 小)$

$f(S_0 = S, O_{-2} = 从)$

$f(S_0 = S, O_1 = 开)$

$f(S_0 = S, O_2 = 始)$

$f(S_0 = S, O_{-1} = 小, O_0 = 就)$

$f(S_0 = S, O_{-2} = 从, O_0 = 就)$

$f(S_0 = S, O_0 = 就, O_1 = 开)$

$f(S_0 = S, O_0 = 就, O_2 = 始)$

$f(S_0 = S, O_{-1} = 小, O_0 = 就, O_1 = 开)$

Conditional random fields

S	**B**	**E**	**S**	**B**	**E**	**S**	**S**	**B**	**E**	**S**
他	**从**	**小**	**就**	**开**	**始**	**学**	**画**	**动**	**物**	**。**

$$f(S_0 = S, S_{-1} = E, O_0 = \text{就})$$

$$f(S_0 = S, S_1 = B, O_0 = \text{就})$$

$$f(S_0 = S, S_{-2} = B, O_0 = \text{就})$$

$$f(S_0 = S, S_2 = E, O_0 = \text{就})$$

$$f(S_0 = S, S_{-1} = E, O_0 = \text{就}, O_{-1} = \text{小})$$

CRF training

一定会收敛

调整

所I特征--

S	**B**	**E**	**I**	**B**	**E**	**S**	**S**	**B**	**E**	**S**
S	**B**	**E**	**S**	**B**	**E**	**S**	**S**	**B**	**E**	**S**
他	**从**	**小**	**就**	**开**	**始**	**学**	**画**	**动**	**物**	**。**

所有的特征++

$f(S_0, O_0)$ $\quad f(S_0 = S, O_0 = \text{就}) + +$ $\quad f(S_0 = I, O_0 = \text{就}) - -$

$f(S_0, O_{-1}, O_0)$ $\quad f(S_0 = S, O_{-1} = \text{小}, O_0 = \text{就}) + +$ $\quad f(S_0 = I, O_{-1} = \text{小}, O_0 = \text{就}) - -$

$f(S_0, O_0, O_1)$ $\quad f(S_0 = S, O_0 = \text{就}, O_1 = \text{开}) + +$ $\quad f(S_0 = I, O_0 = \text{就}, O_1 = \text{开}) - -$

$f(S_0, O_{-1}, O_1)$ $\quad f(S_0 = S, O_{-1} = \text{小}, O_1 = \text{开}) + +$ $\quad f(S_0 = I, O_{-1} = \text{小}, O_1 = \text{开}) - -$

$f(S_{-1}, S_0)$ $\quad f(S_{-1} = E, \ S_0 = S) + +$ $\quad f(S_{-1} = E, \ S_0 = I) - -$

$\quad f(S_{-1} = S, \ S_0 = B) + +$ $\quad f(S_{-1} = I, \ S_0 = B) - -$

HashTable

一种数组[illegible]。

Feature templates

用这多个模板的原因：训练集中的 unknown 判断。

通过有限的特征 → 得到……

当N个pattern……

	Unigram	Bigram
C_0	C_0S_0	$C_0S_{-1}S_0$
C_{-1}	$C_{-1}S_0$	$C_{-1}S_{-1}S_0$
C_1	C_1S_0	$C_1S_{-1}S_0$
$C_{-1}C_0$	$C_{-1}C_0S_0$	$C_{-1}C_0S_{-1}S_0$
C_0C_1	$C_0C_1S_0$	$C_0C_1S_{-1}S_0$
$C_{-1}C_1$	$C_{-1}C_1S_0$	$C_{-1}C_1S_{-1}S_0$
$C_{-2}C_{-1}$	$C_{-2}C_{-1}S_0$	$C_{-2}C_{-1}S_{-1}S_0$
$S_{-1}S_0$	$S_{-1}S_0$	$S_{-1}S_0$
$\vdots$	$\vdots$	$\vdots$

无

有标签转移.

可以考虑前一个日期和下一个日期之间的关系???

可以定义很远的日期间的关系