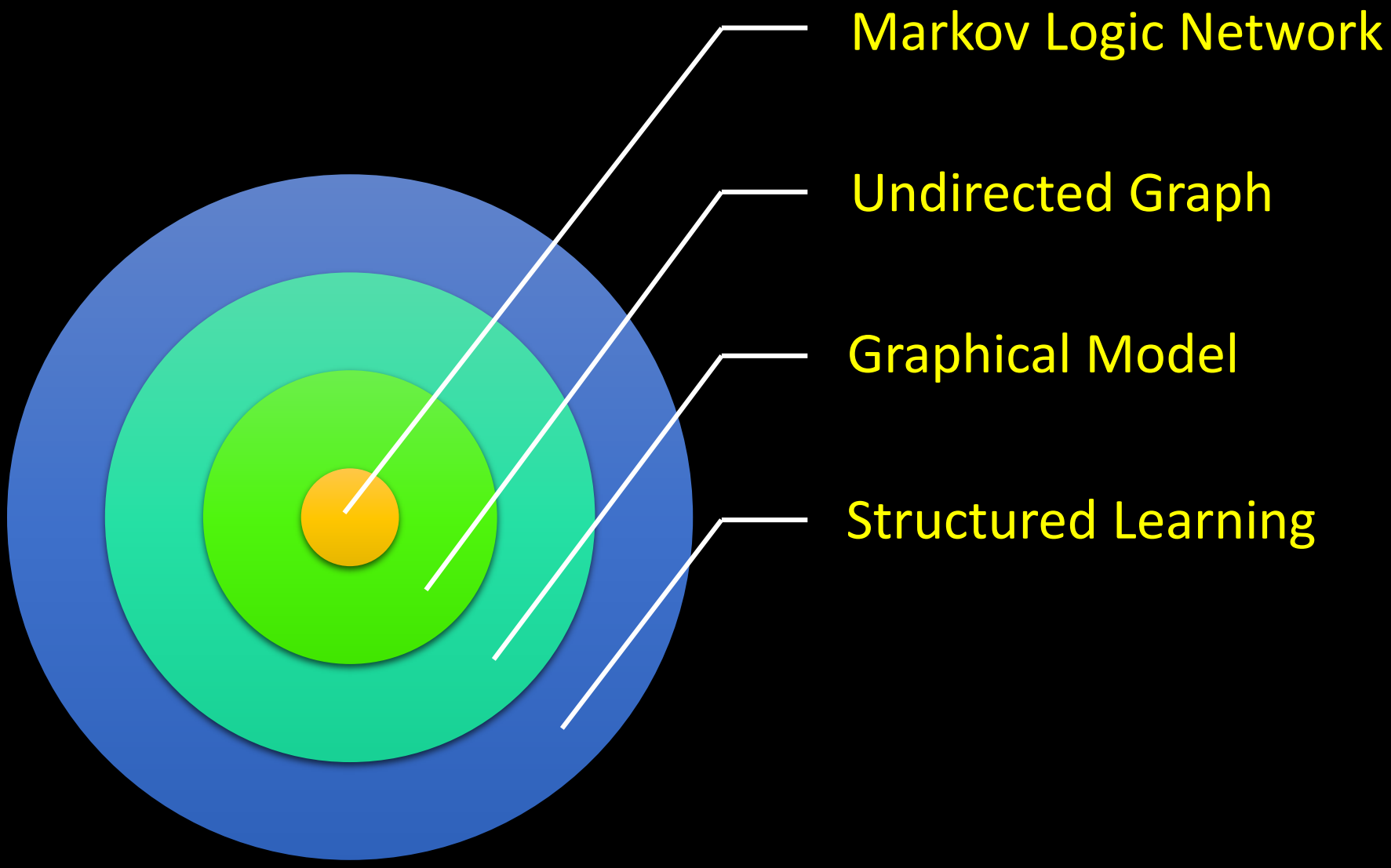
Markov Logic Network

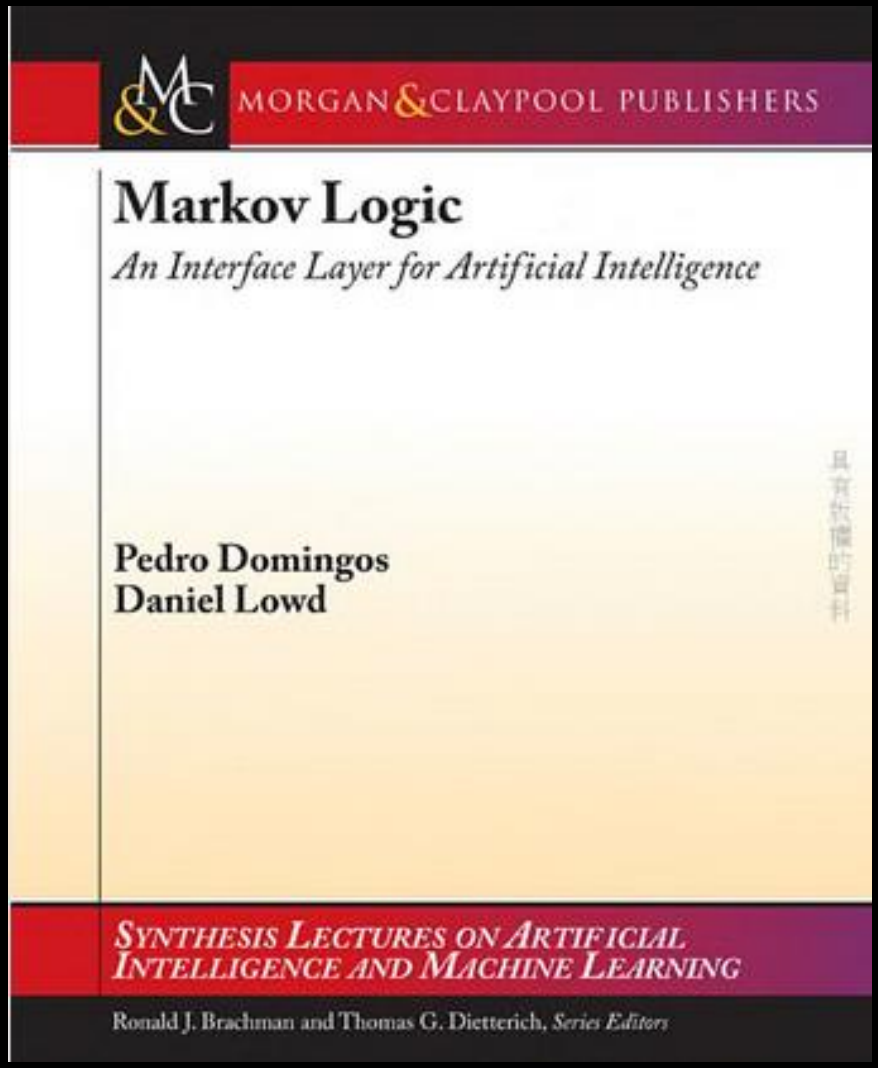
Hung-yi Lee

Markov Logic Network



Textbook

Domingos, Pedro, and Daniel Lowd. "Markov logic: An interface layer for artificial intelligence." *Synthesis Lectures on Artificial Intelligence and Machine Learning* 3.1 (2009): 1-155.



More Reference

- Course:
  - http://homes.cs.washington.edu/~pedrod/803/
- Toolkit:
  - Alchemy: http://alchemy.cs.washington.edu/
  - Tuffy: http://i.stanford.edu/hazy/hazy/tuffy/

Machines use logic as humans?

- This is also a structured learning problem.
- ***Evaluation***:
  - Evaluate how logical a possibility is based on the knowledge base
    - Using as graphical model (?)
- ***Inference***:
  - Evaluate all possibilities and find the most logical one
- ***Training***:
  - Learn the knowledge base

Terminology

- ***Knowledge base***: a set of ***Formulas***
- ***Formula***: $\mathrm{P} \Rightarrow Q$, $\mathrm{P} \Leftrightarrow Q$ …
  - Conjunction of ***predicates*** by logic operation
  - Logic operation: ~, ∧, ∨, ⇒, ⇔
- ***Predicate***: P, Q
  - Predicate is a function
  - Input: one or several objects
  - Output: True or False

Predicate:

姓宇智波 (x)

有寫輪眼(x)

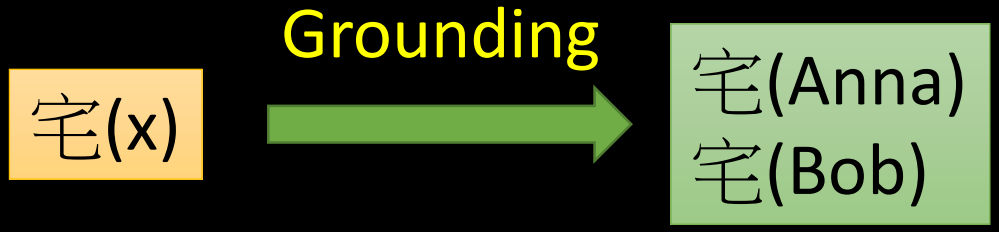
Formula:

姓宇智波(x) ⇒有寫輪眼(x)

Terminology

- ***Grounding***: Replace the variables in the predicates with all possible constant
- Example:
  - Predicate: 宅(x)
  - x: is a variable which can be any person

Person = {Anna, Bob} → constant



Terminology

World: 一個可能的情況

- ***World***: Grounding all the predicates, and assign a truth value to each grounded predicate
  - Predicate: 是朋友(x,y), 宅(x), 做研究(y)
  - Person: {A, B, C}

是朋友(A,B)	=T
是朋友(B,C)	=T
是朋友(A,C)	=T
宅(A)	=T
宅(B)	=T
宅(C)	=T
做研究(A)	=T
做研究(B)	=T
做研究(C)	=T

U

是朋友(A,B)	=F
是朋友(B,C)	=F
是朋友(A,C)	=F
宅(A)	=F
宅(B)	=F
宅(C)	=F
做研究(A)	=F
做研究(B)	=F
做研究(C)	=F

U’

Evaluation – Logic

- Evaluating a world is logic or not given a knowledge base

Knowledge Base L:

宅(x) ⇒做研究(x)
是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))

Person:
{A, B}

World U:

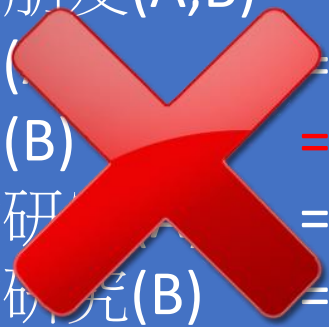
L(U) = T

是朋友(A,B) =T
宅(A) =T
宅(B) =T
做研究(A) =T
做研究(B) =T

World U':

L(U') = F

是朋友(A,B) =T
宅(=T
宅(B) =F
做研 =T
做研究(B) =T



Inference - Logic

Knowledge Base L:

宅(x) ⇒做研究(x)
是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))

Evidence:

是朋友(A,B) =T
宅(A)=T

World U:

是朋友(A,B)	=T
宅(A)	=T
宅(B)	=?
做研究(A)	=?
做研究(B)	=?

X (evidence)

Y (to be inferred)

Check the 8 possible worlds

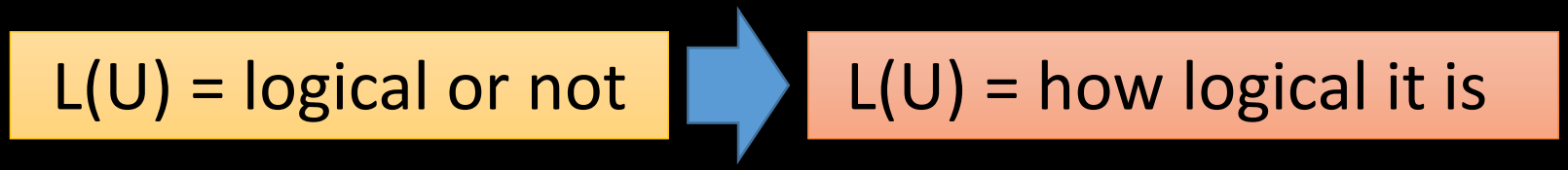
By L(U)

是朋友(A,B)	=T
宅(A)	=T
宅(B)	=T
做研究(A)	=T
做研究(B)	=T

There is only one logical world.

Evaluation – Soft Logic

- The real world is complex.



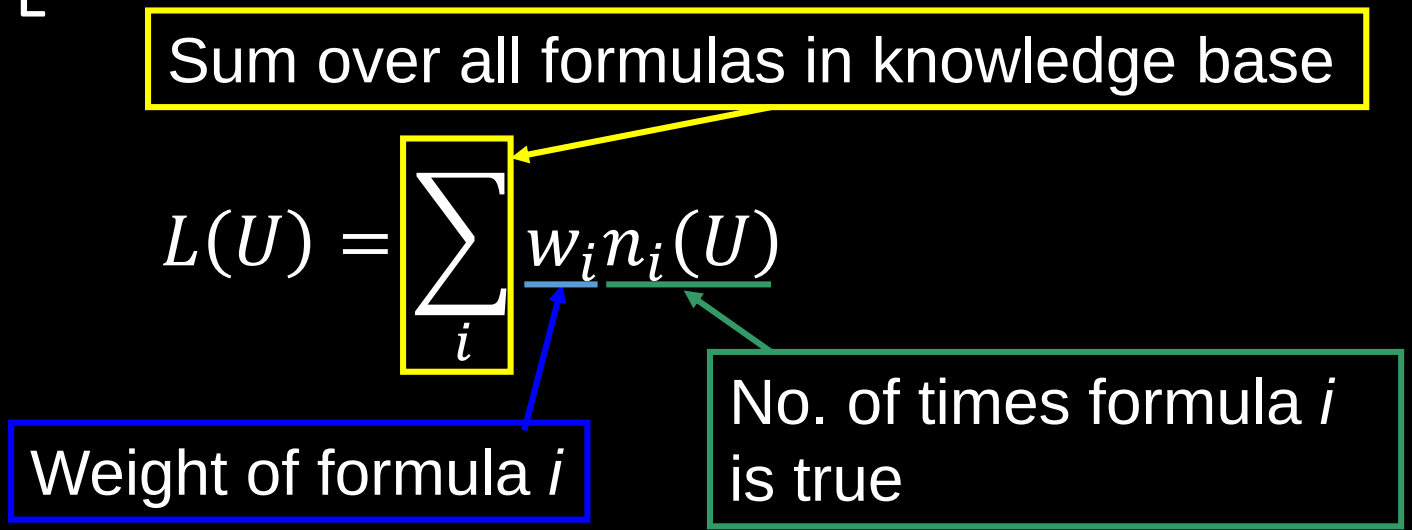
- Each formula is assigned a weight representing confidence score.

1.5	宅(x) ⇒做研究(x)
1.1	是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))

- When a world violates a formula, It becomes less probable, but not impossible

Evaluation – Soft Logic

- $L(U)$: How logical a world U is given knowledge base L



- Probability point of view:

  (The probability that the world U appears.)

$$P(U) = \frac{e^{L(U)}}{\sum_{U'} e^{L(U')}}$$

Inference – Soft Logic

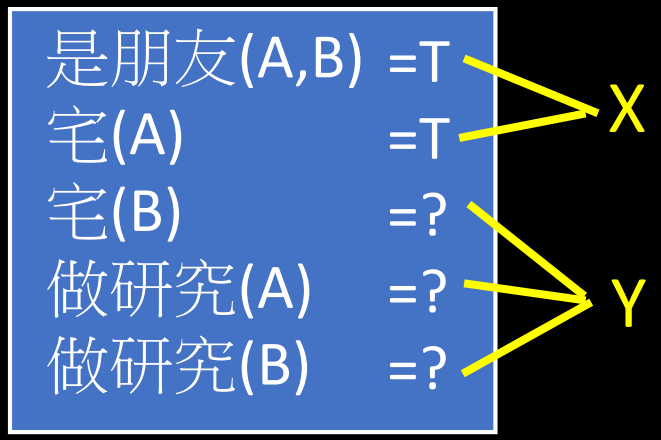
- Given a world U = {X, Y}, X is known, find the most possible Y

$$\tilde{Y} = arg\max_{Y} L(X,Y)$$

$$= arg\max_{Y} \sum_{i} w_i n_i(X,Y)$$

$$\tilde{Y} = arg\max_{Y} P(Y|X)$$

$$= arg\max_{Y} \frac{e^{L(X,Y)}}{\sum_{Y'} e^{L(X,Y')}}$$



You can use Gibbs sampling if you do not know how to solve the problem.

Example

w_1 1.5 宅(x) ⇒做研究(x) f_1

Person = {Anna}

If we observe that 宅(A) = T

P(做研究(A) = T) = 0.82

P(做研究(A) = F) = 0.18

宅(A)	T	T	F	F
做研究(A)	T	F	T	F
$n_1(U)$	1	0	1	1
$L(U)$	1.5	0	1.5	1.5
$P(U)$	0.31	0.07	0.31	0.31

w_1	1.5	宅(x) ⇒做研究(x)	f_1
w_2	1.1	是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))	f_2

Evidence:

是朋友(A,B) =T
宅(A)=T

Person = {Anna, Bob}

做研究(A)	宅(B)	做研究(B)	$n_1(U)$	$n_2(U)$	score	Prob
T	T	T	1+1=2	1	4.1	0.43
T	T	F	1+0=1	1	2.6	0.10
T	F	T	1+1=2	0	3.0	0.14
T	F	F	1+1=2	0	3.0	0.14
F	T	T	0+1=1	1	2.6	0.10
F	T	F	0+0=0	1	1.1	0.02
F	F	T	0+1=1	0	1.5	0.03
F	F	F	0+1=1	0	1.5	0.03

Q: Bob 是否喜歡做研究?

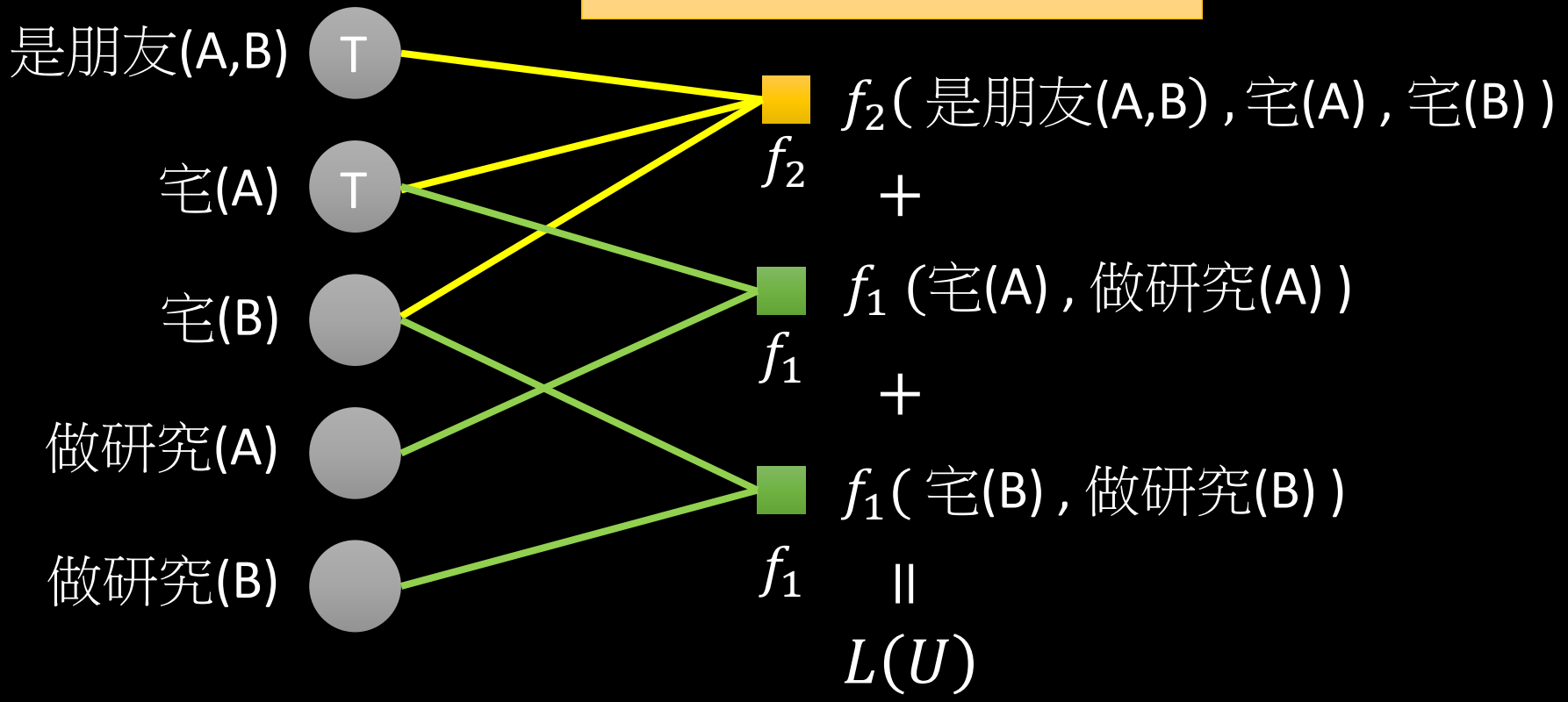
A: 70% 的機率 Bob 喜歡做研究

Graphical Model

1.5	宅(x) ⇒做研究(x)	f_1
1.1	是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))	f_2

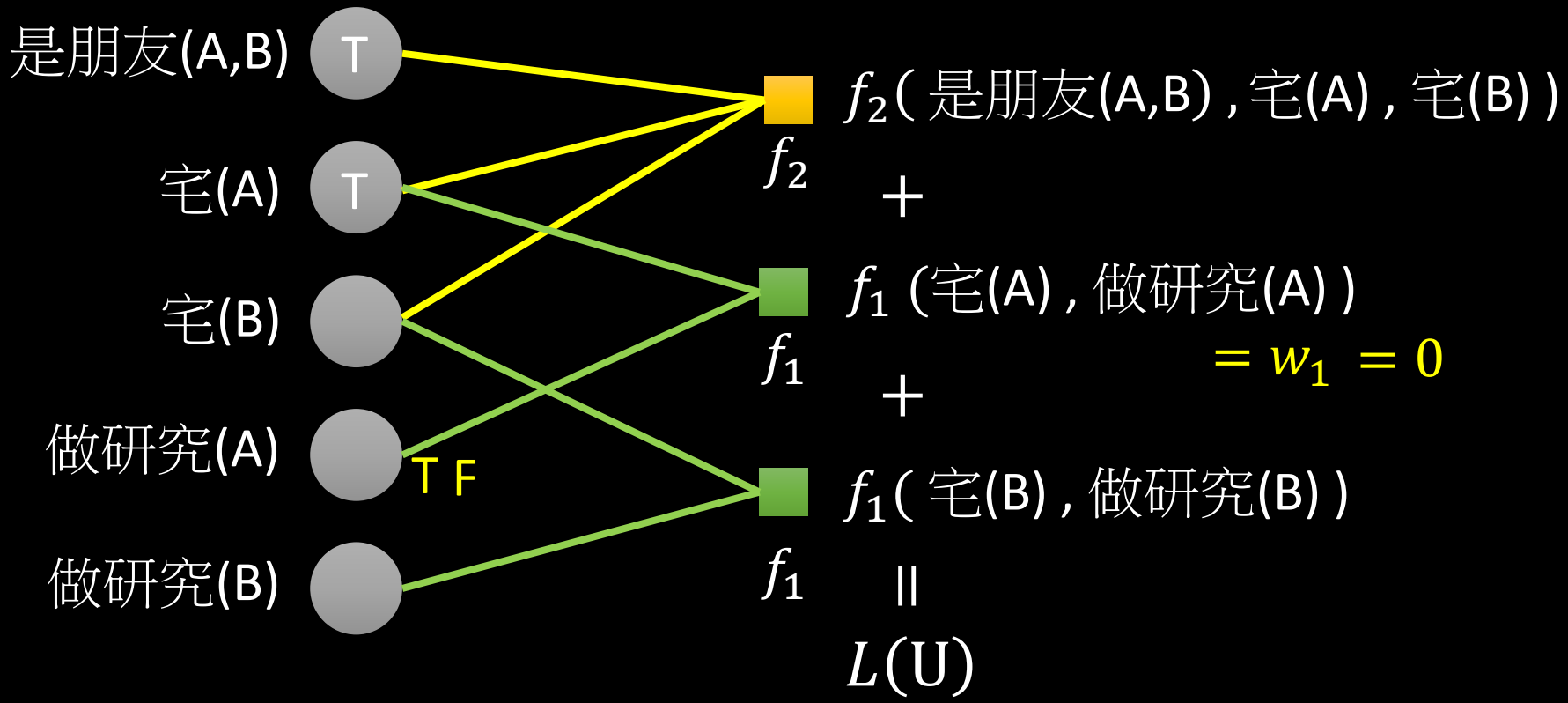
Each ground predicate is a node.

The formulas are Factors.



w_1	1.5	宅(x) ⇒做研究(x)	f_1
w_2	1.1	是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))	f_2

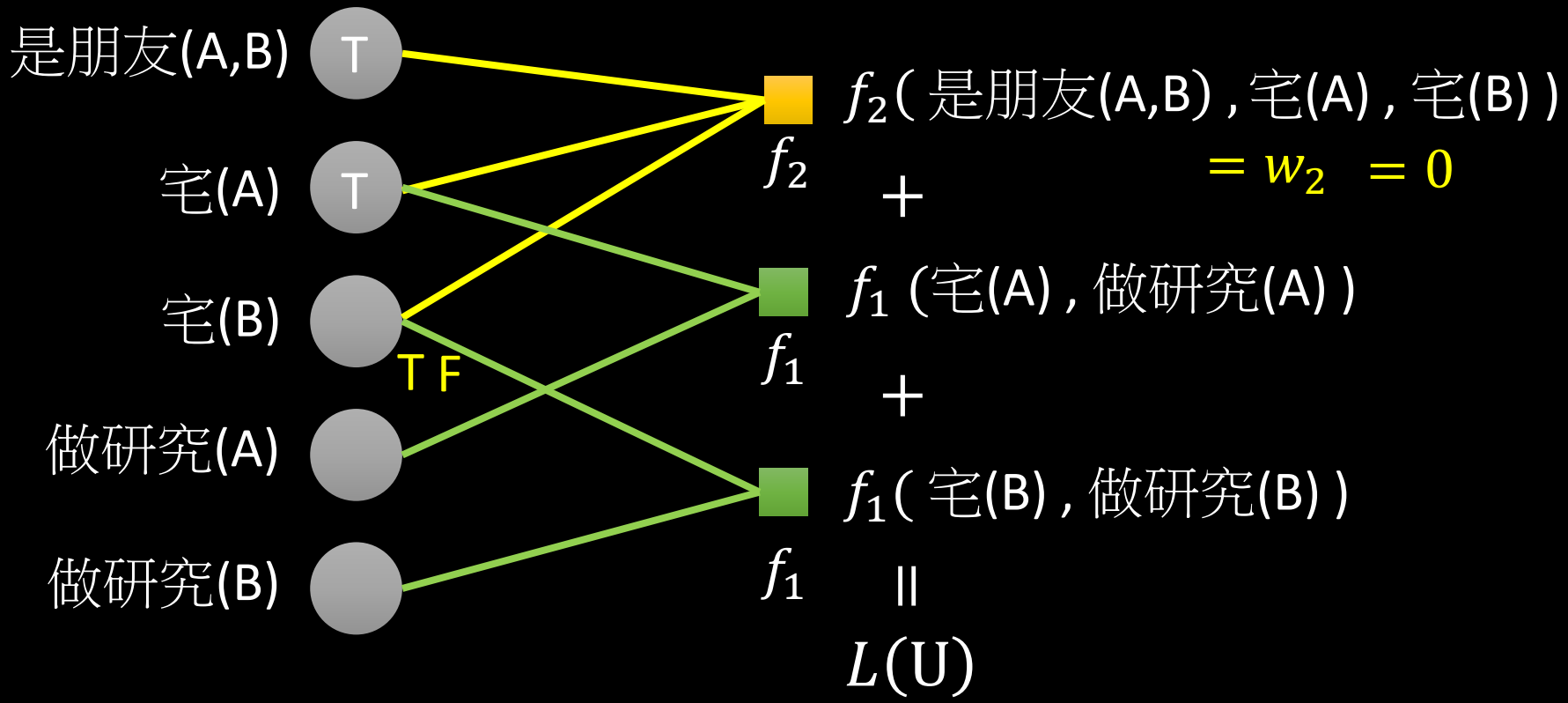
$$f_1\,(\text{宅(x)}\,,\,\text{做研究(x)}\,) = \begin{cases} w_1 & (\text{宅(x)} \Rightarrow \text{做研究(x)}) \text{ is true} \\ 0 & otherwise \end{cases}$$



w_1	1.5	宅(x) ⇒做研究(x)	f_1
w_2	1.1	是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))	f_2

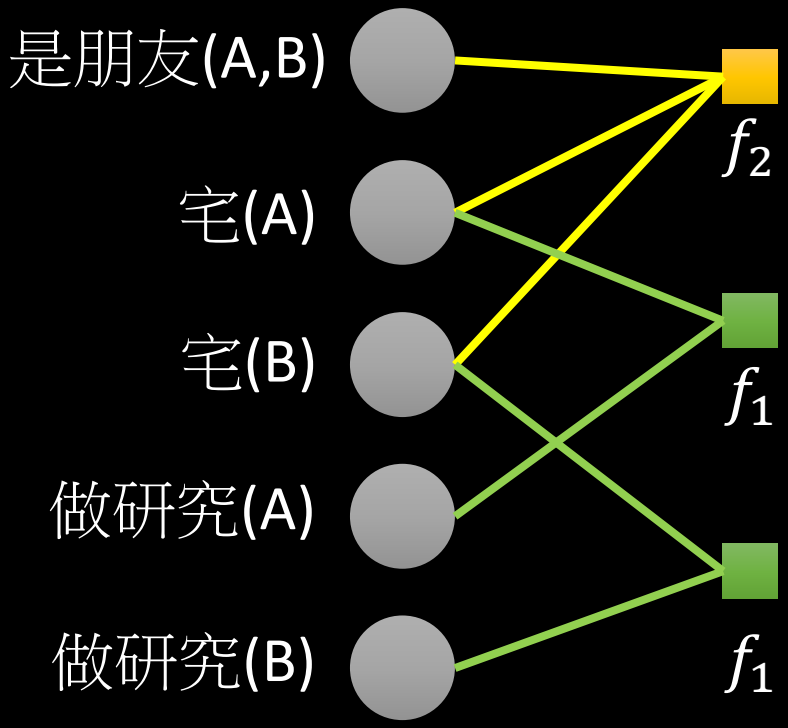
(是朋友(x,y) ⇒ (宅(x) ⇔ 宅(y))) is true

f_2(是朋友(A,B) , 宅(A) , 宅(B))= $\begin{cases} w_2 \\ 0 \end{cases}$ $otherwise$

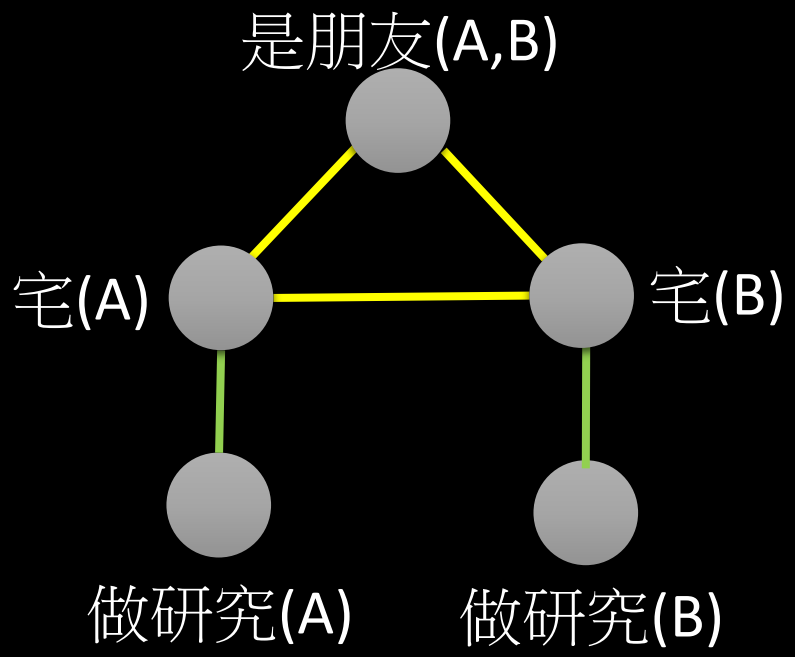


Graphical Model

Factor Graph



Markov Random Field



This is why the model is named ***Markov Logic Network***

Learning

- Given a set of formulas $\{F_1 \cdots F_i \cdots F_N\}$ and world U, assign weights $\{w_1 \cdots w_i \cdots w_N\}$ for each formulas
- Maximizing the likelihood of world $P(U)$

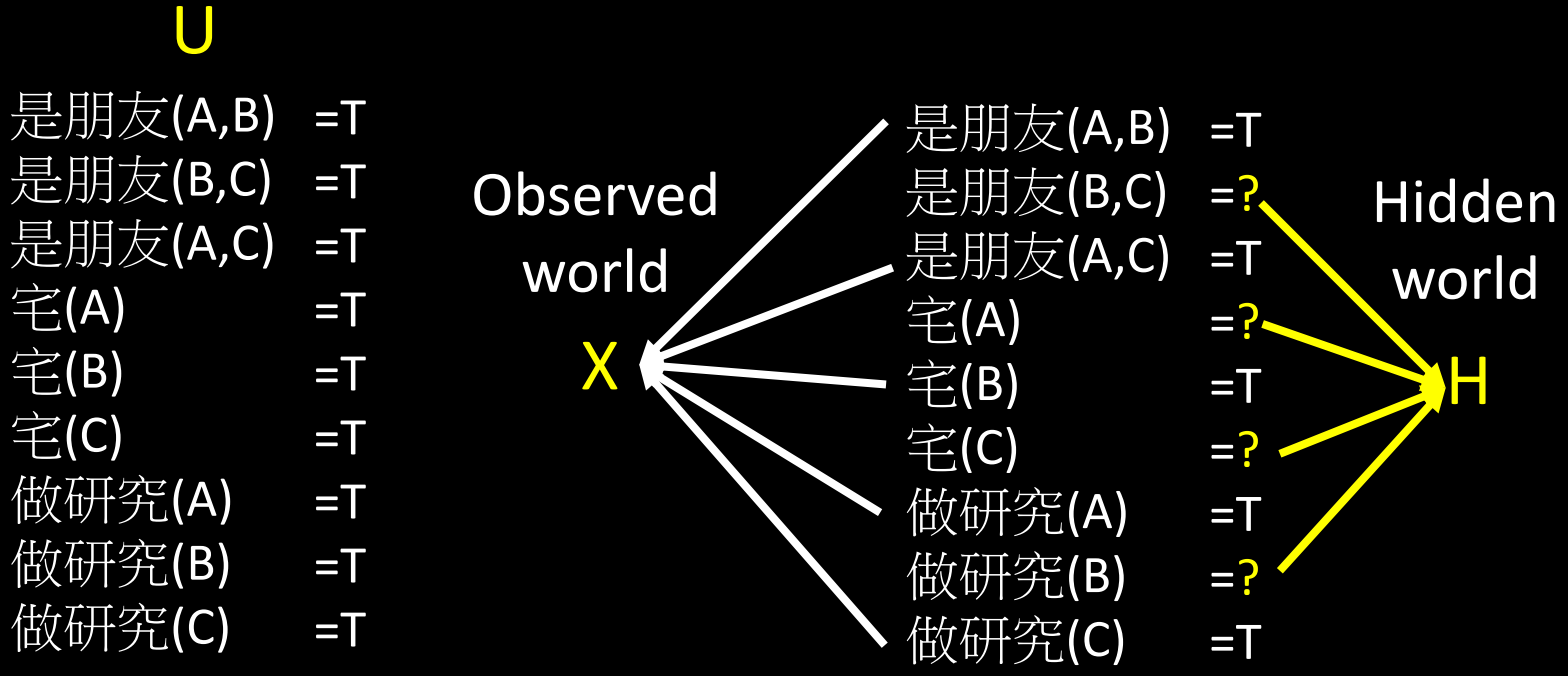
$$P(U) = \frac{e^{L(U)}}{\sum_{U'} e^{L(U')}} \quad L(U) = \sum_i w_i n_i(U)$$

$$logP(U) = L(U) - log \sum_{U'} e^{L(U')}$$

Gradient ascent: $w_i \leftarrow w_i + \eta \frac{\partial logP(U)}{\partial w_i}$

Learning

- If there is some missing data in the world



Learning

- If there is some missing data in the world
  - X: observed part of the world, H: missing part

$$P(X) = \sum_{H'} P(X, H')$$

Everything you don't observe can happen

$$\frac{\partial log P(X)}{\partial w_i} = \sum_{H'} P(X, H') n_i(X, H') - \sum_{X', H'} P(X', H') n_i(X', H')$$

Learning the Correctness - UW-CSE database

- Available: http://www.cs.washington.edu/ai/mln/database.html
- 1158 constants: person, course, paper title ……
- 22 predicates: Professor(x), AdvicedBy(x,y)
- 4M grounding predicates, 3k are true
- 94 hand-crafted formulas are given
- Learn the weight
  - The formulas with the highest weights
    - (course c is taught by x) => (x is a professor) 3.5
    - (x is advised by y) => (y is the co-author of the paper x published) 3.1

Outlook

- Markov Logic Network can learn more than just weights
- It can ……
    - Discover the knowledge
    - Discover the predicates
    - Transfer Learning
    - Unsupervised Learning
- Markov Logic Network can be used in general supervised learning like classification
    - Not especially powerful, but interpretable