Knowledge representation : (Markov logic networks)

A set of first-order logic rules can be represented as a Markov logic networks

f_1: $A \wedge B \Rightarrow Y \iff \neg A \vee \neg B \vee Y$

Predicate set $X = (X_1, X_2, \ldots X_n) \in \mathcal{X}$

Energy-based model : (Markov logic Networks)

$$P(X = x) = \frac{1}{Z} \exp\left\{ \sum_{j \in |F|} w_j f_j(x) \right\}$$

where $f_j(x)$ is logic-informed feature.

(KDD) - draft

$K = (C, X, O)$ | ICML 2024)

(knowledge base) entity relation facts

$C = \{C_1, C_2, \cdots C_M\}$. entity set

(in our case. we can only consider one entity. i.e. focus on individual patient)

Consider a set of N predicates (properties or relations)

(e.g. measurements, Treatments, ...)

$X = \{X_1 \cdots X_N\}$

$X(\cdot) : C \to \{0, 1\}$ logic variable

each grounded predicate $\equiv$ a binary random variable

by data

$O = \{o_1, \cdots o_L\}$, where each observation is a true value $\{0, 1\}$. assigned to a grounded predicate.

Knowledge base $\mathcal{F} = \{f_1, \cdots f_K\}$.

$f(\ ): \quad \in \{0,1\}$. logic formula.

$f(\ ): \quad A \wedge B \Rightarrow Y$ (Horn rule)

$A(c) \wedge B(c) \Rightarrow Y(c)$ if-then

Given De Morgan's law.

a disjunctive form

$f(\ ): \neg A(c) \vee \neg B(c) \vee Y(c)$

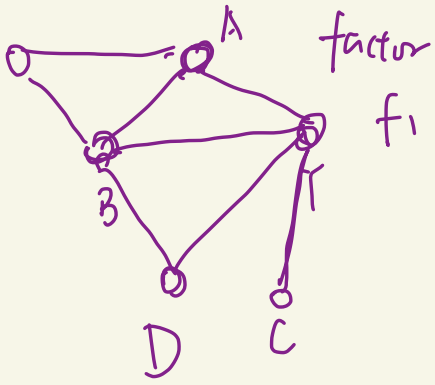
General form:

$$f \overset{\Delta}{=} \left(\bigvee_{x \in X^+} x \right) \vee \left(\bigvee_{x \in X^-} \neg x \right) \quad \bigstar$$

Given Markov logic networks

Any f_j can be evaluated by $\bigstar$.

$$P(X = x) = \frac{1}{Z} \exp\left(\sum_{j \in |F|} w_j f_j(x_{\{j\}}) \right)$$



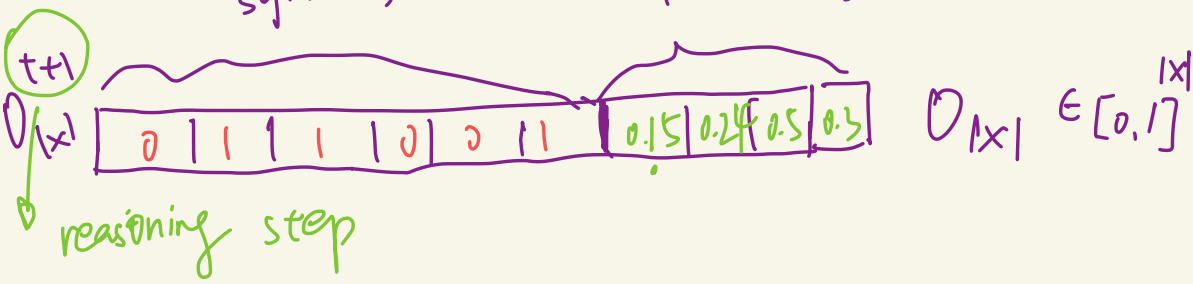
f_1: A ∨ B ∨ Y

f_2: C ∨ Y

f_3: D ∨ B ∨ Y

Consider a setting where we have a very huge knowledge graph "$\bar{N}$ is very big.

We first consider "exact reasoning" procedure.



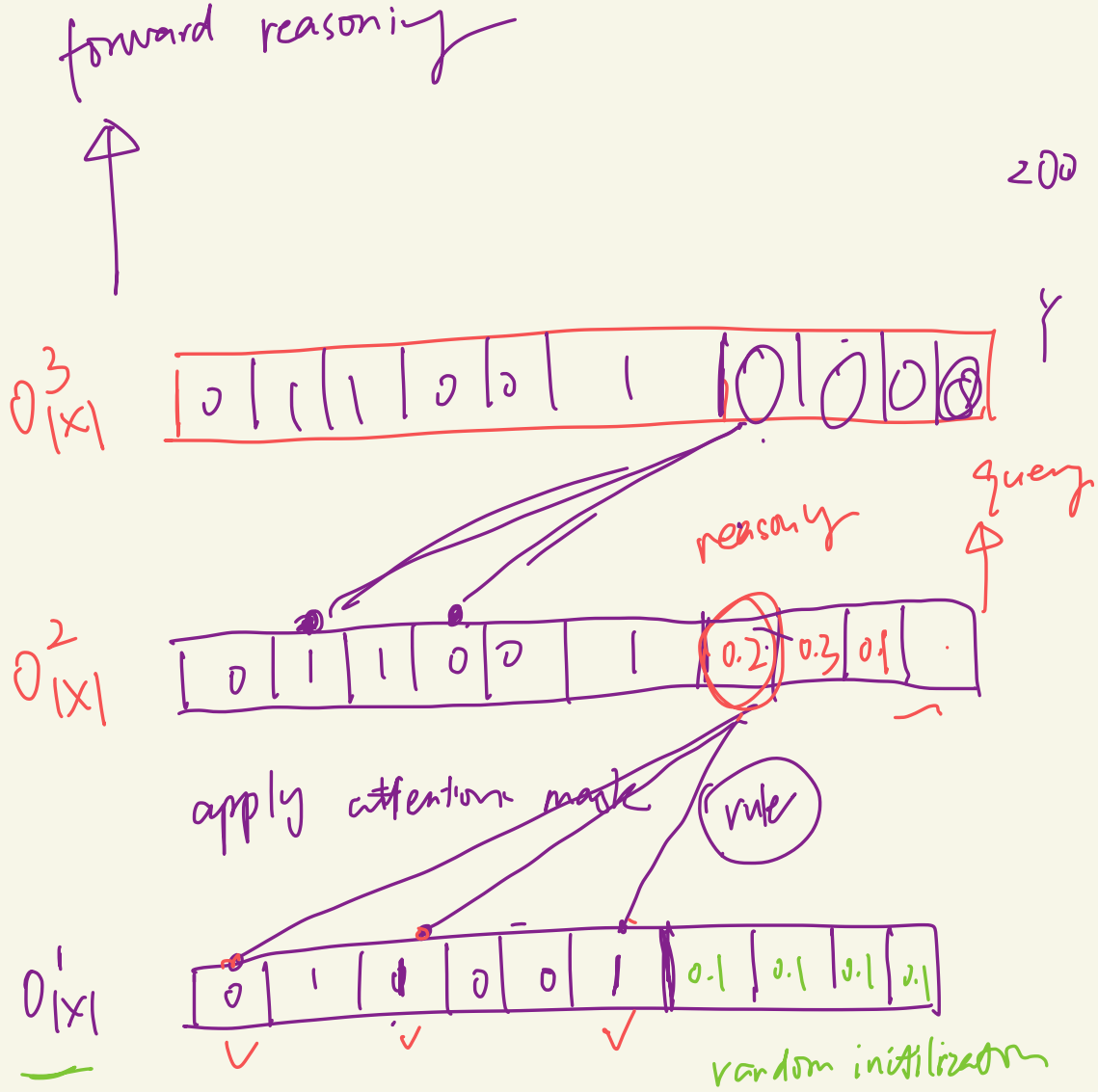
For an AI algorithm (no information processing limit)

↓ if $O_{|x|}$ (symptoms part) is fully observed. Reasoning only requires one-step.

For a real human, however, her information processing ability is limited.

↓

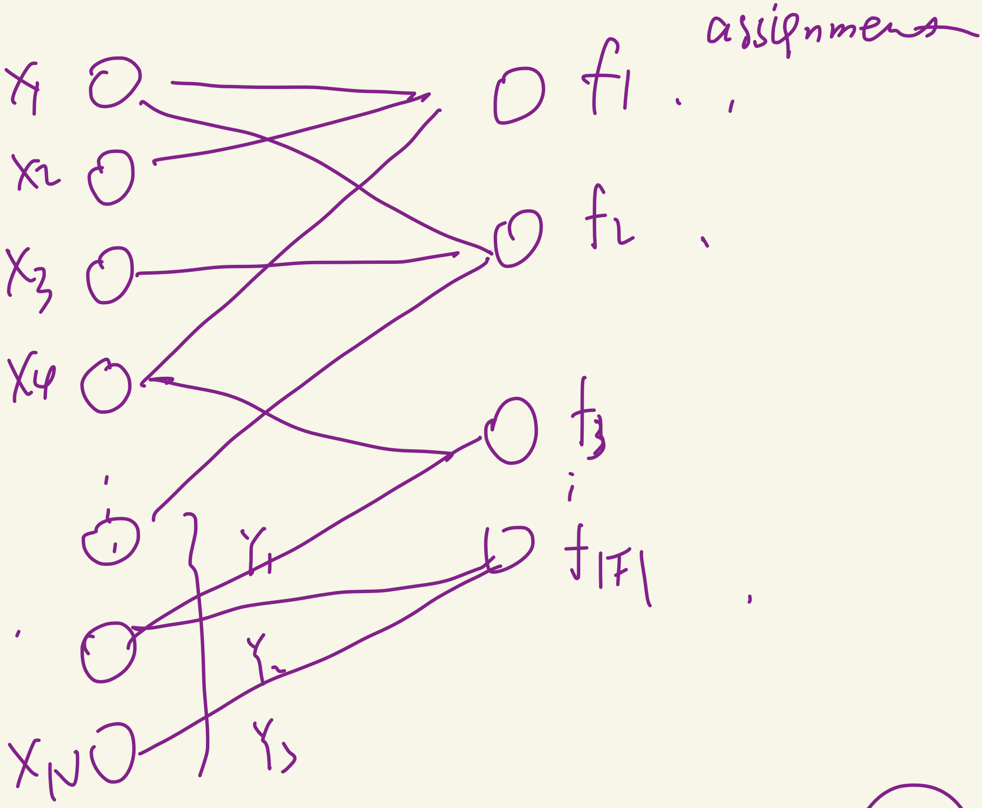
A snapshot can only focus on some segments of information.



attention is used to select information from all the observations

Modelling: ① how to compute the current reasoning results given information in the memory. (increasing)

(ICML 2024 paper results)



500

Forward reasoning: (ICLR: KB)

chaining (always from evidence to goals) ✓

Given $\{X_1, X_3, X_5\}$ in the memory.

trigger $\{f_1, f_2, f_4, f_6, \ldots\}$

(determine neighbor hood)