Learning Analytics

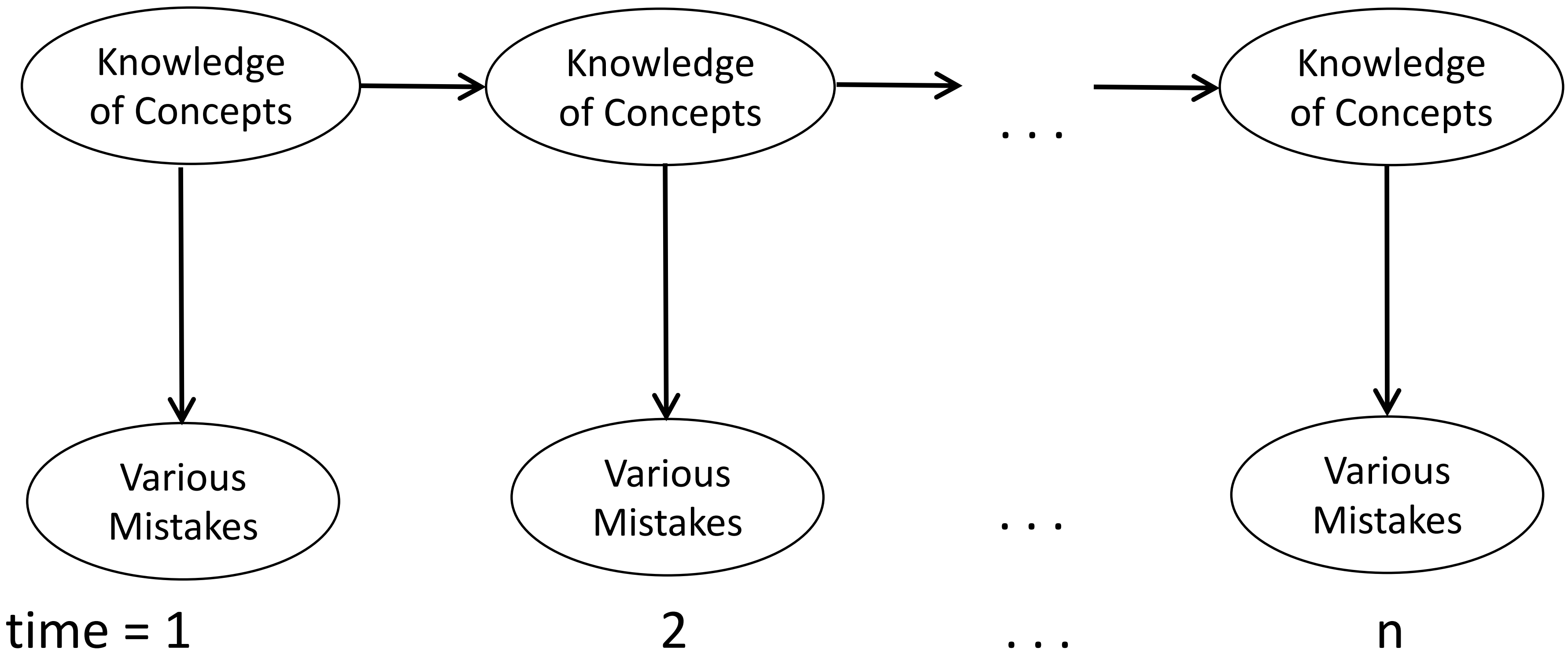
Dr. Bowen Hui

Computer Science

University of British Columbia Okanagan

Reasoning Over Time

- What if model dynamics change over time?

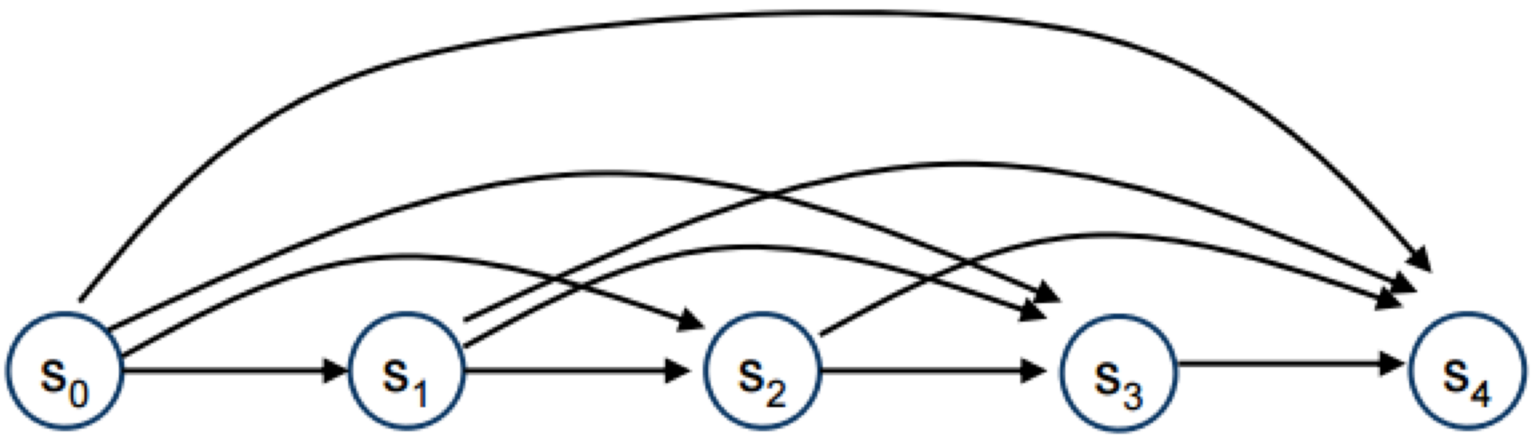


Dynamic Inference

- Model time-steps (snapshots of the world)
  - Granularity: 1 second, 5 minutes, 1 session, …
- Allow different probability distributions over states at each time step
- Define temporal causal influences
  - Encode how distributions change over time

The Big Picture

- Life as one big stochastic process:
  - Set of States: S
  - Stochastic dynamics: $Pr(s_t | s_{t-1}, \ldots, s_0)$



  - Can be viewed as a Bayes net with one random variable per time step

Challenges with a Stochastic Process

- Problems:
  - Infinitely many variables
  - Infinitely large conditional probability tables

Challenges with a Stochastic Process

- Problems:
  - Infinitely many variables
  - Infinitely large conditional probability tables

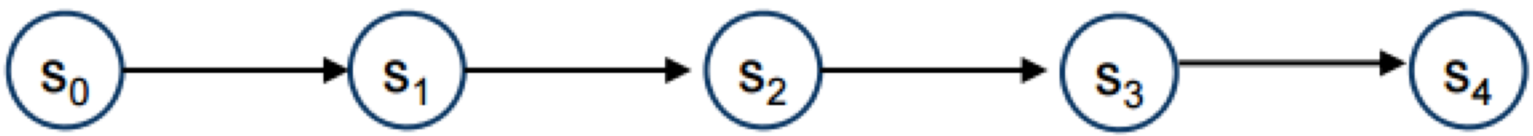
- Solutions:
  - Stationary process: dynamics do not change over time
  - Markov assumption: current state depends only on a finite history of past states

K-order Markov Process

- Assumption: last k states sufficient

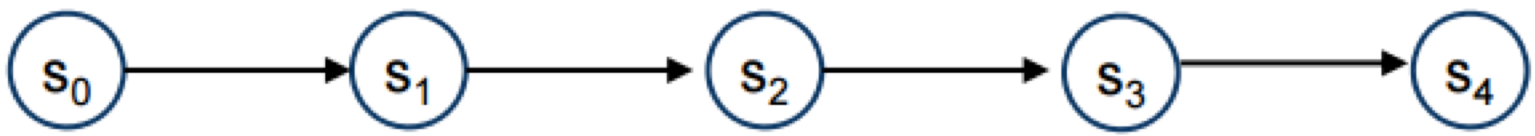
K-order Markov Process

- Assumption: last k states sufficient
- First-order Markov process
  - $Pr(s_t | s_{t-1}, \ldots, s_0) = Pr(s_t | s_{t-1})$



K-order Markov Process

- Assumption: last k states sufficient
- First-order Markov process
  - $\Pr(s_t | s_{t-1}, \ldots, s_0) = \Pr(s_t | s_{t-1})$



- Second-order Markov process
  - $\Pr(s_t | s_{t-1}, \ldots, s_0) = \Pr(s_t | s_{t-1}, s_{t-2})$

K-order Markov Process

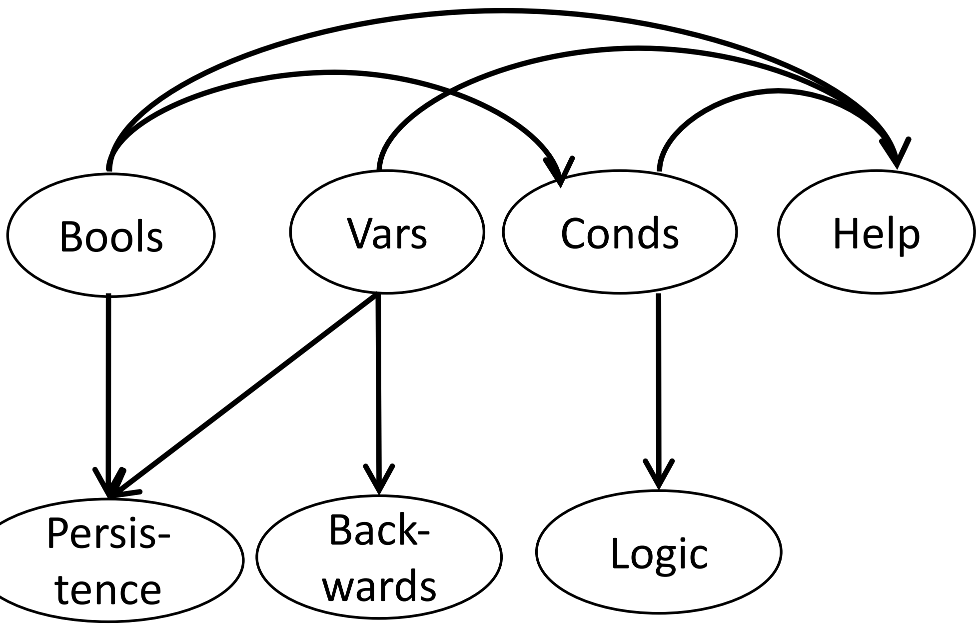
- Advantage:
  - Can specify entire process with *finitely* many time steps

- Two time steps sufficient for a first-order Markov process
  - Graph: 

  - Dynamics: $Pr(s_t|s_{t-1})$
  - Prior: $Pr(s_0)$

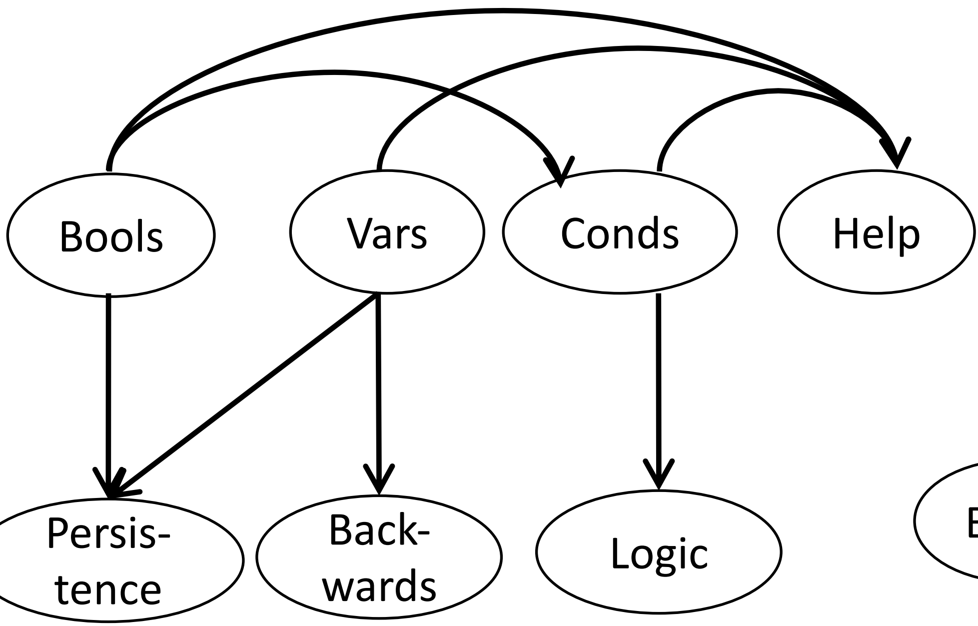
Building a 2-Stage Dynamic Bayes Net

Use same CPTs as BN



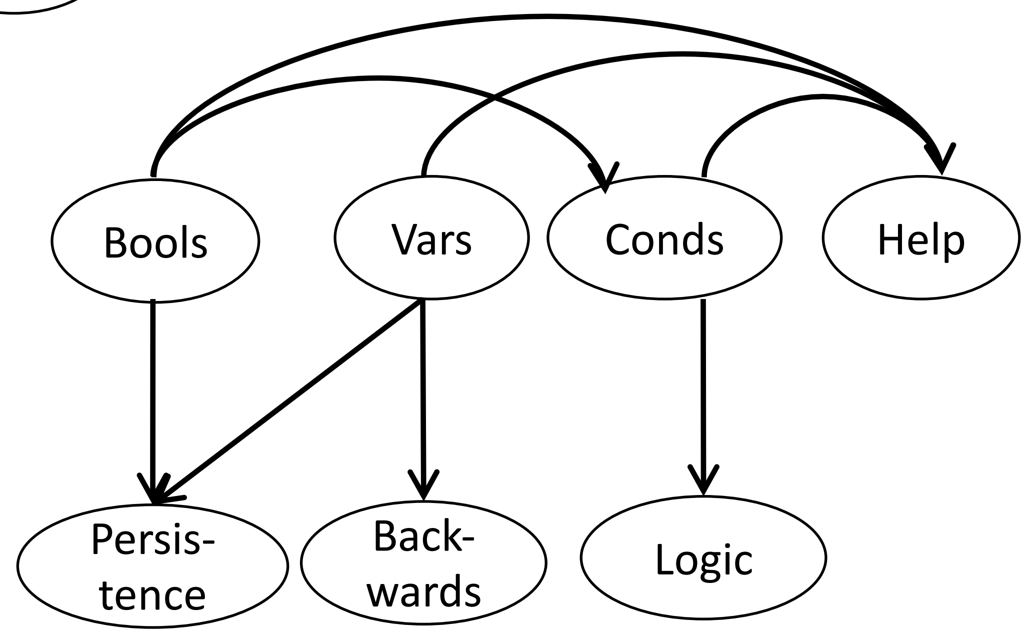
Building a 2-Stage Dynamic Bayes Net

Use same CPTs as BN



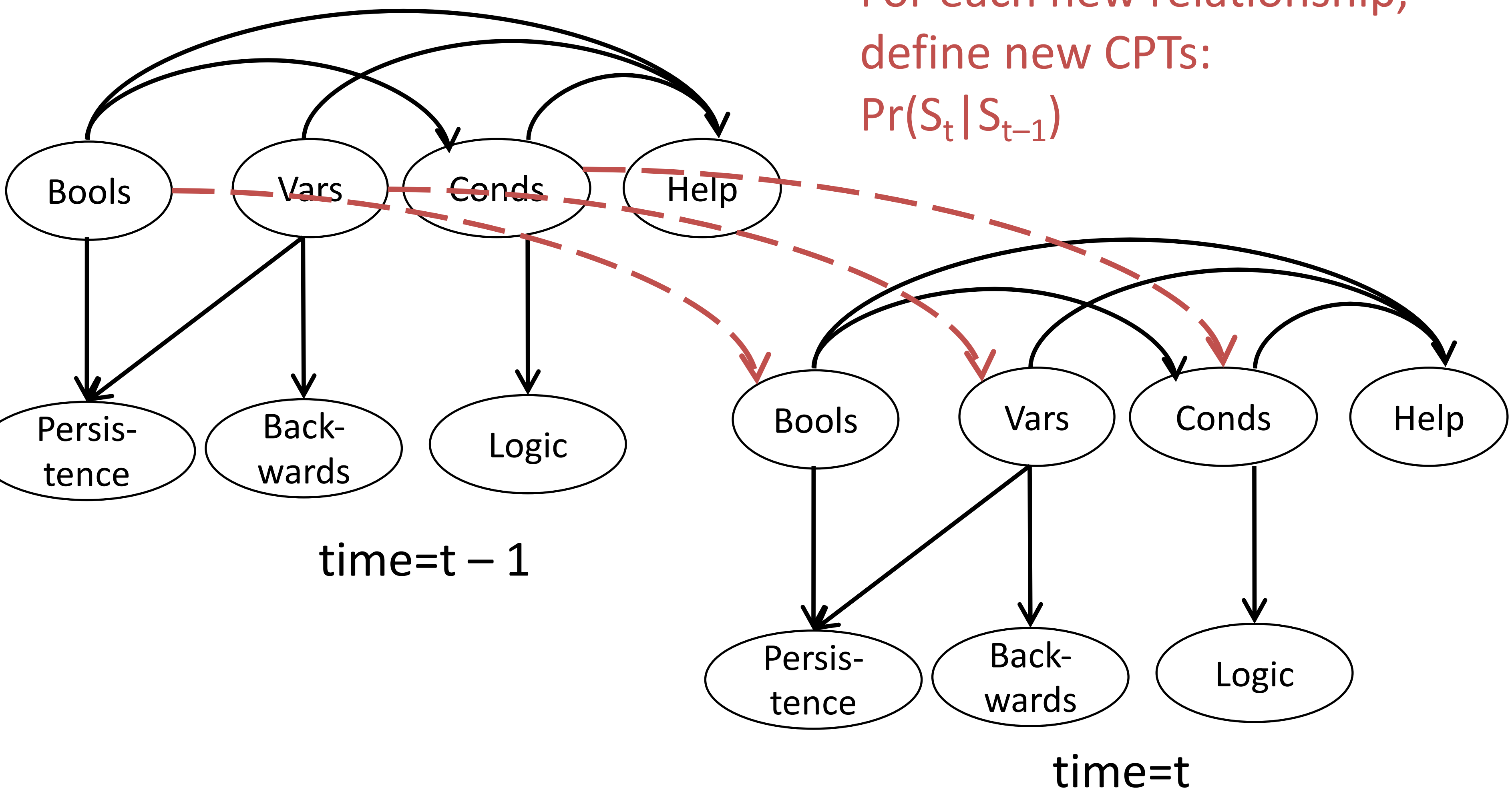
time=t – 1

Use same or different CPTs as other time step



time=t

Building a 2-Stage Dynamic Bayes Net



For each new relationship, define new CPTs: $\Pr(S_t | S_{t-1})$

Defining a DBN

- Recall that a BN over variables **X** = $\{X_1, X_2, \ldots, X_n\}$ consists of:
  - A directed acyclic graph whose nodes are variables
  - A set of CPTs $Pr(X_i | Parents(X_i))$ for each X_i

- Dynamic Bayes Net is an extension of BN
- Focus on 2-stage DBN

DBN Definition

- A DBN over variables **X** consists of:
  - A directed acyclic graph whose nodes are variables
  - **S** is a set of hidden variables, $\mathbf{S} \subset \mathbf{X}$
  - **O** is a set of observable variables, $\mathbf{O} \subset \mathbf{X}$
  - Two time steps: *t–1* and *t*
  - Model is first order Markov s.t. $\Pr(U_t | U_{1:t-1}) = \Pr(U_t | U_{t-1})$

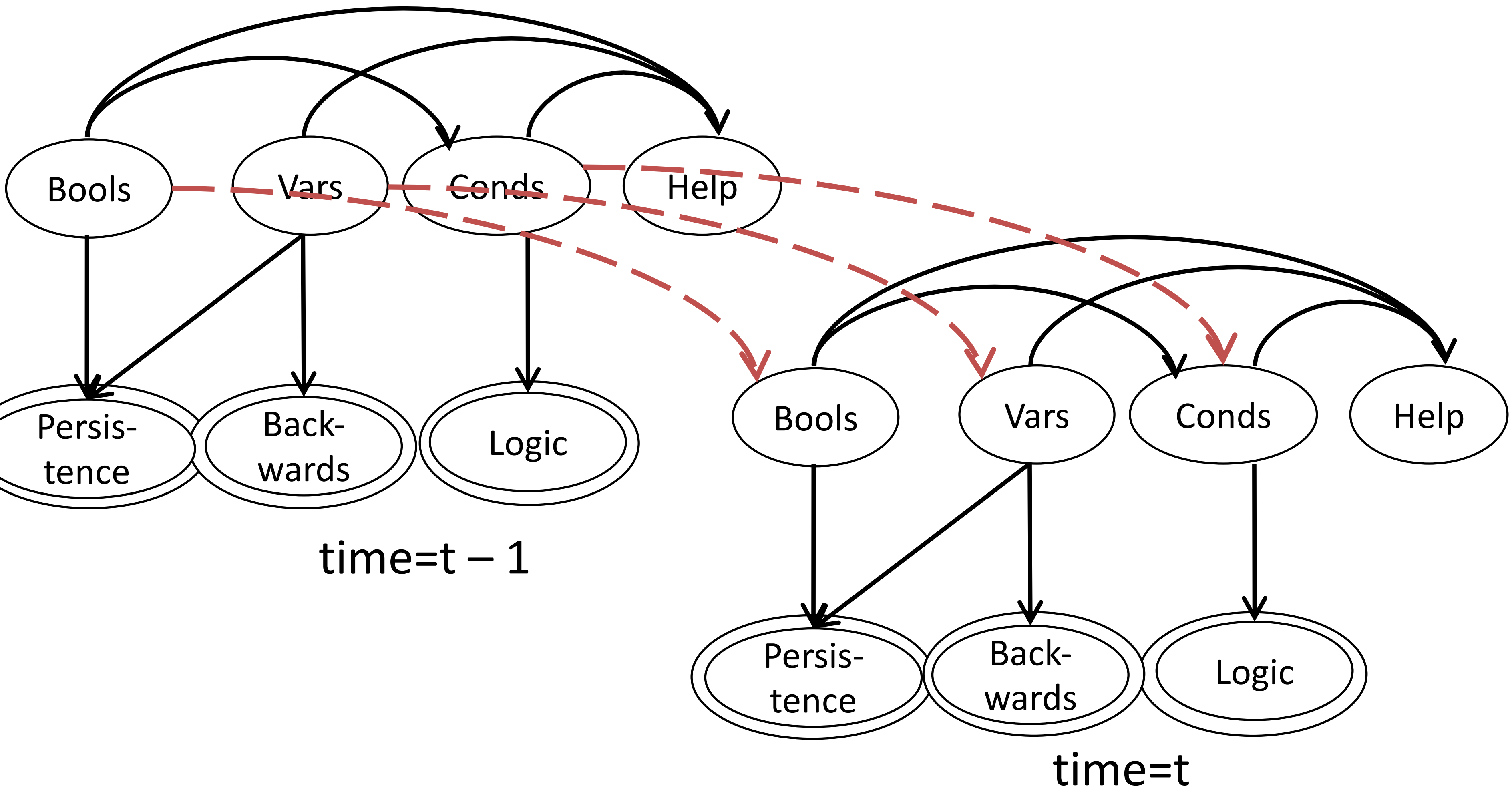
DBN Definition (cont.)

For BNs: CPTs *$Pr(X_i|Parents(X_i))$* for each *X_i*

- Numerical component:
  - Prior distribution, $\Pr(\mathbf{X}_0)$
  - State transition function, $\Pr(S_t|S_{t-1})$
  - Observation function, $\Pr(O_t|S_t)$

- See previous example

Example


Bools
Vars
Conds
Help
Persis-
tence
Back-
wards
Logic
time=t – 1
Bools
Vars
Conds
Help
Persis-
tence
Back-
wards
Logic
time=t

Exercise #1

- Changes in ***skill*** from t–1 to t?
- Context:
  My son is potty training. When he has to go, he either tells us in advance (yay!), he starts but continues in the potty, or an accident happens. With feedback and practice, his ability to go potty improves.

- Does a temporal relationship exist? What might the CPT look like?

Exercise #2

- Changes in ***hint quality*** from t–1 to t?
- Context:
  You’re tutoring a friend in Math. Rather than solving the exercises for him, you offer to give hints. Sometimes the hints aren’t so great. Depending on the quality, you may or may not give hints each time he is stuck.
- Does a temporal relationship exist? What might the CPT look like?

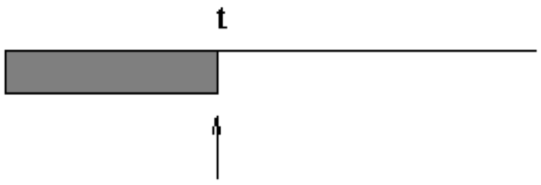
Exercise #3

- Changes in ***blood type*** from t–1 to t?
- Context:
  People with blood type A tend to be more cooperative. We have the opportunity to observe a student's teamwork in class twice each week. We want to infer the student's blood type.

- Does a temporal relationship exist? What might the CPT look like?

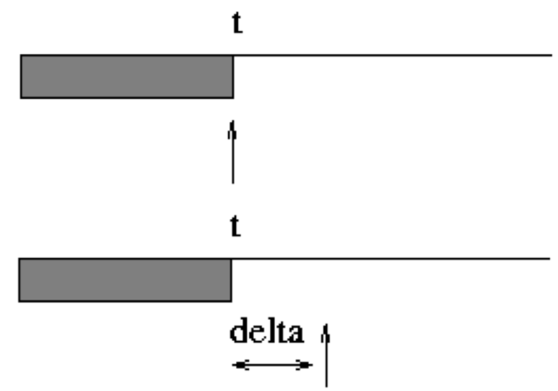
Temporal Inference Tasks

- Common tasks:
  - Belief update: $Pr(s_t | o_t, \ldots, o_1)$



Temporal Inference Tasks

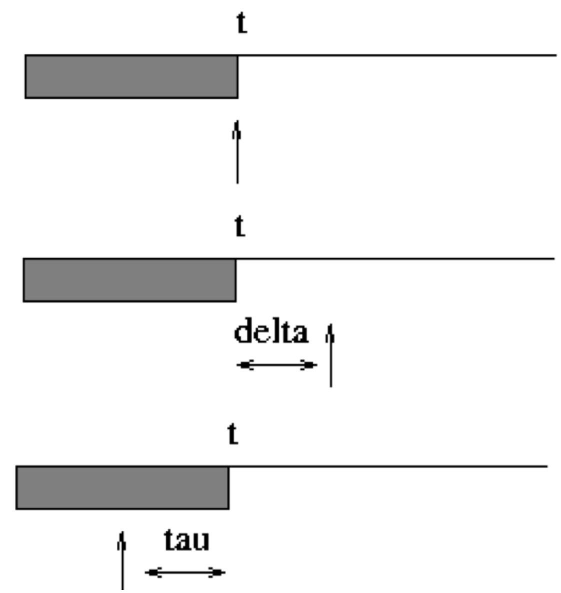
- Common tasks:
  - Belief update:
    $Pr(s_t | o_t, \ldots, o_1)$
  - Prediction:
    $Pr(s_{t+\delta} | o_t, \ldots, o_1)$



$Pr(st | o_T, \ldots, o_1)$

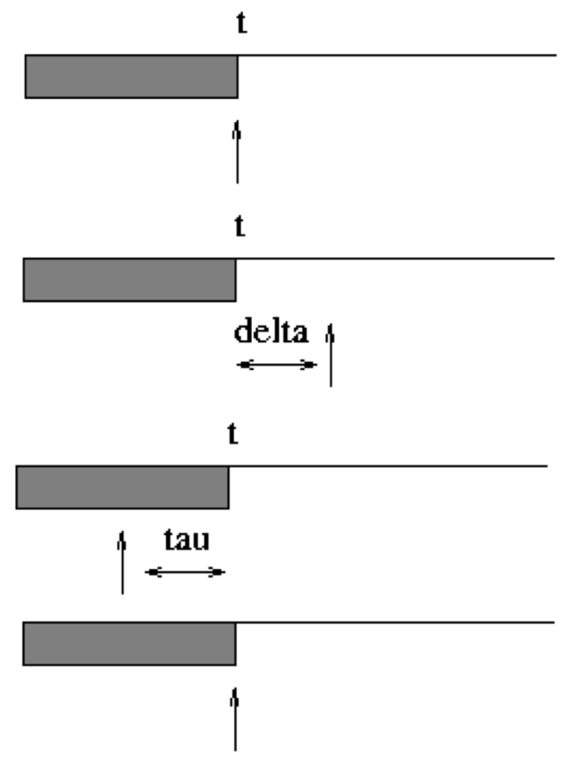
Temporal Inference Tasks

- Common tasks:
  - Belief update:
    $Pr(s_t | o_t, \ldots, o_1)$
  - Prediction:
    $Pr(s_{t+\delta} | o_t, \ldots, o_1)$
  - Hindsight:
    $Pr(s_{t-\tau} | o_t, \ldots, o_1)$ where k < t



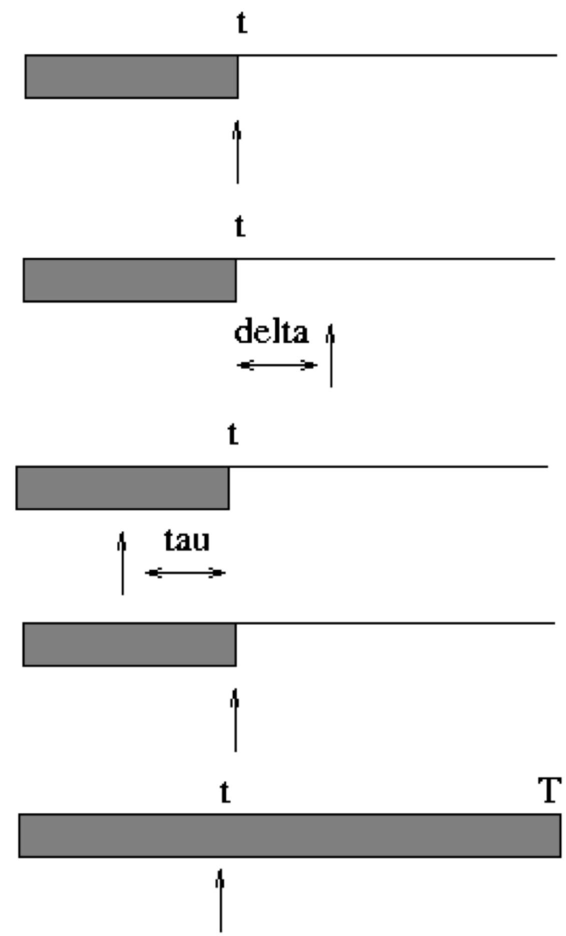
Temporal Inference Tasks

- Common tasks:
  - Belief update:
    $Pr(s_t | o_t, \ldots, o_1)$
  - Prediction:
    $Pr(s_{t+\delta} | o_t, \ldots, o_1)$
  - Hindsight:
    $Pr(s_{t-\tau} | o_t, \ldots, o_1)$ where $k < t$
  - Most likely explanation:
    $\underset{s_t, \ldots, s_1}{\operatorname{argmax}} Pr(s_t, \ldots, s_1 | o_t, \ldots o_1)$



Temporal Inference Tasks

- Common tasks:
  - Belief update:
    $\Pr(s_t \mid o_t, \ldots, o_1)$
  - Prediction:
    $\Pr(s_{t+\delta} \mid o_t, \ldots, o_1)$
  - Hindsight:
    $\Pr(s_{t-\tau} \mid o_t, \ldots, o_1)$ where $k < t$
  - Most likely explanation:
    $\operatorname{argmax}_{s_t, \ldots, s_1} \Pr(s_t, \ldots, s_1 \mid o_t, \ldots o_1)$
  - Fixed interval smoothing:
    $\Pr(s_t \mid o_T, \ldots, o_1)$



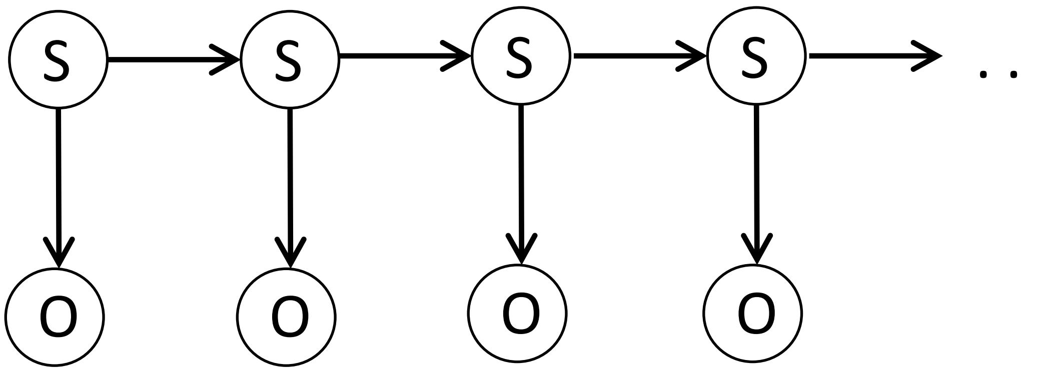
Inference Over Time

- Same algorithm: clique inference

- Added setup:
  - – Enter evidence in stage *t*
  - – Take marginal in stage *t*, keep it aside
  - – Create new copy
  - – Enter marginal into stage *t–1*
  - – Repeat

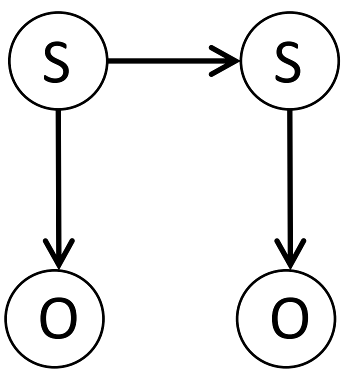
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

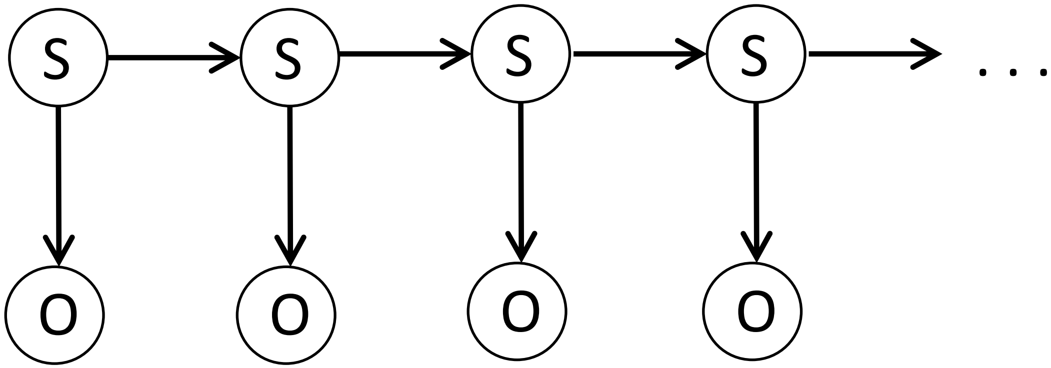
Setup:



time = t–1 t

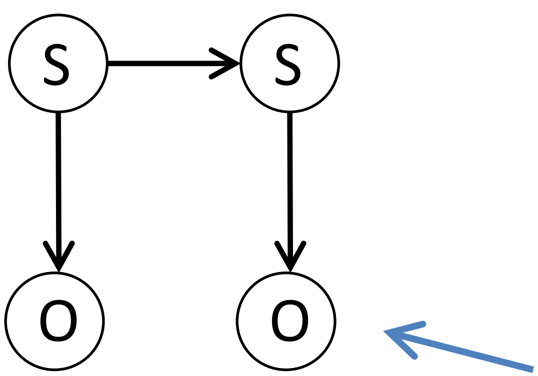
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

Start, time=1:



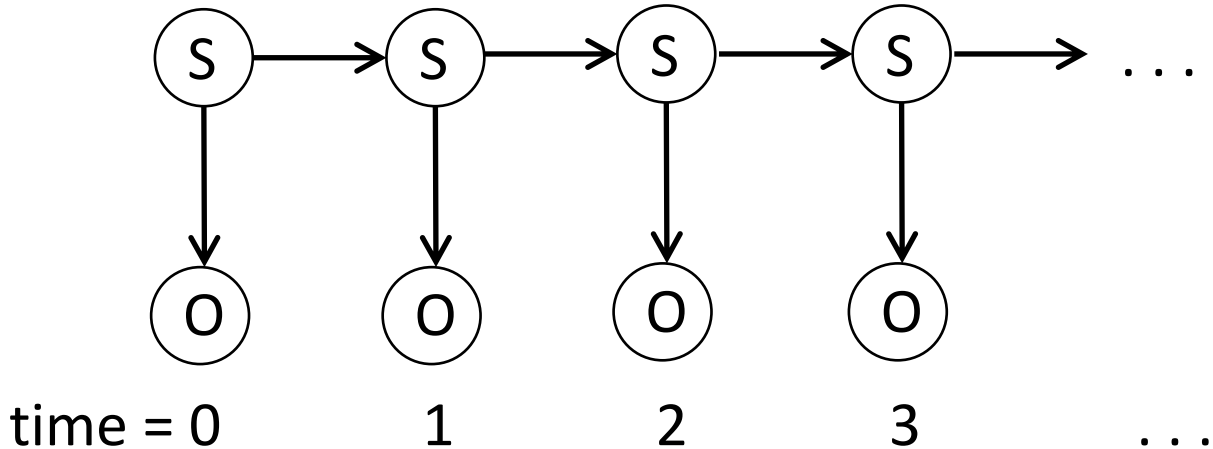
time = 0 1

Observe user behaviour

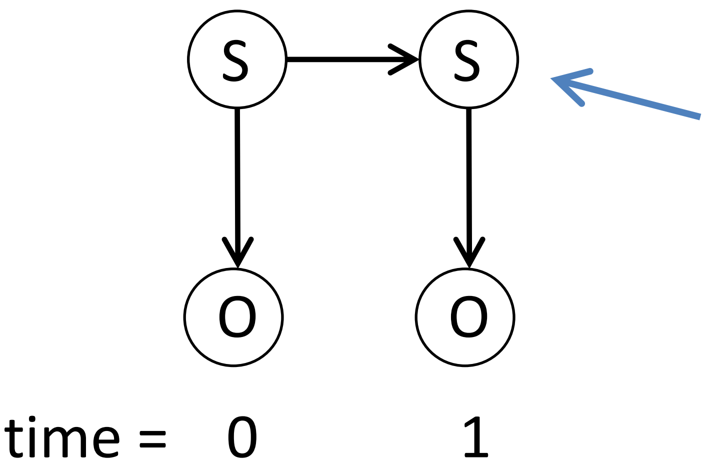
Enter evidence

Inference Over Time

Mimic:



Start, time=1:

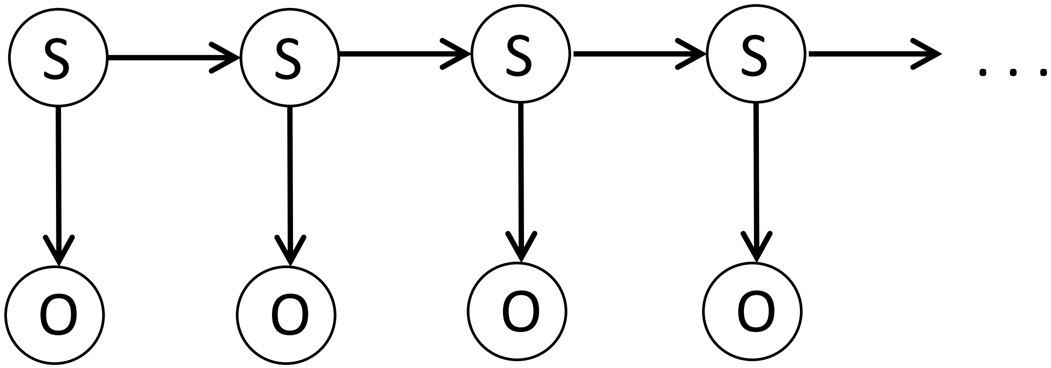


Compute marginal of interest

Get: $Pr(S_1)$

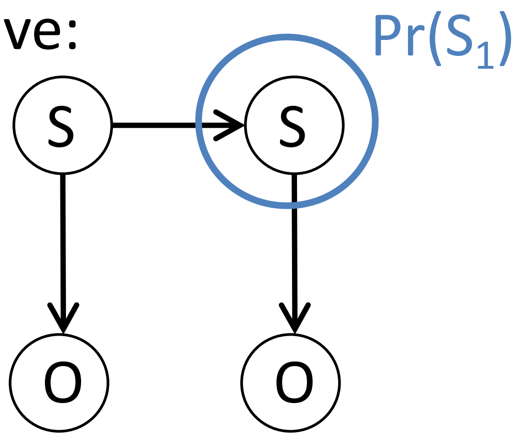
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

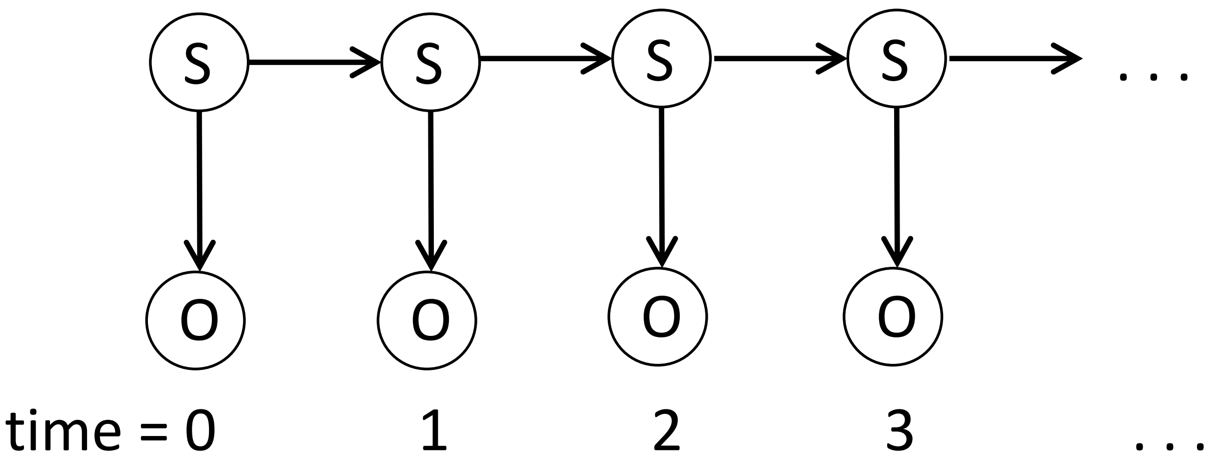
We have:



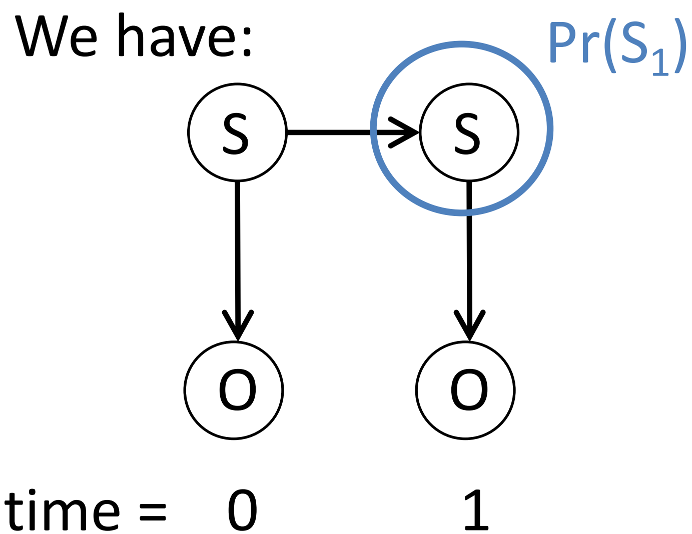
time = 0 1

Inference Over Time

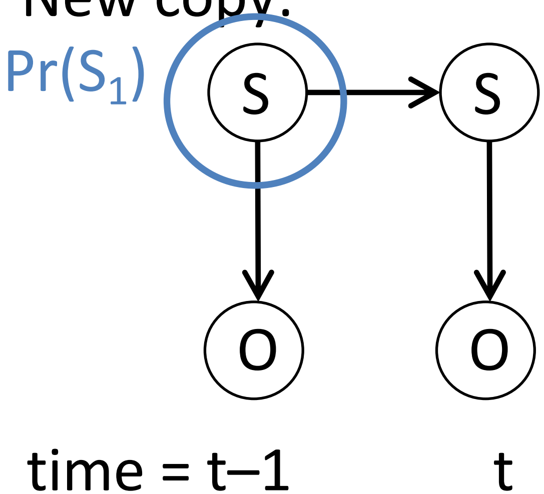
Mimic:



We have:

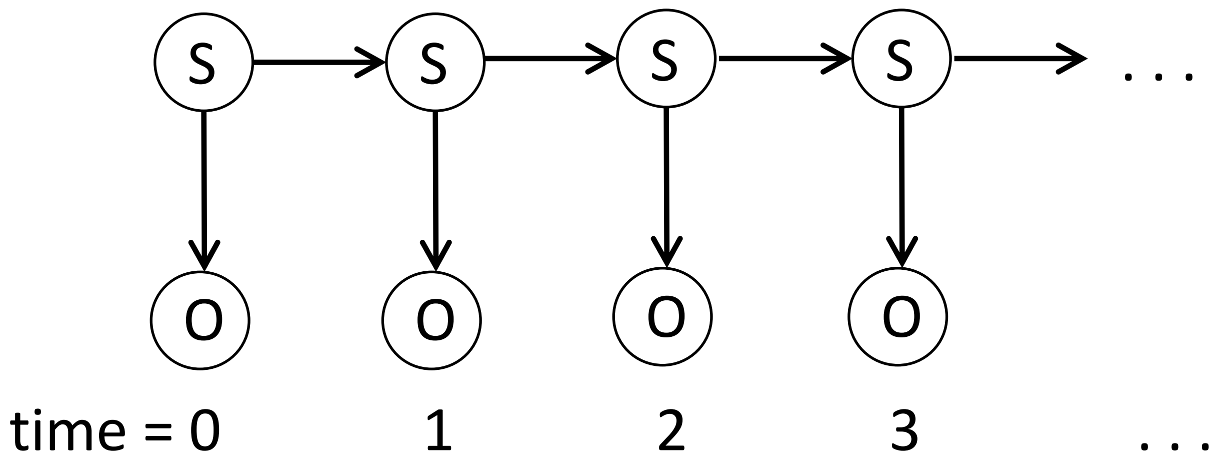


New copy:

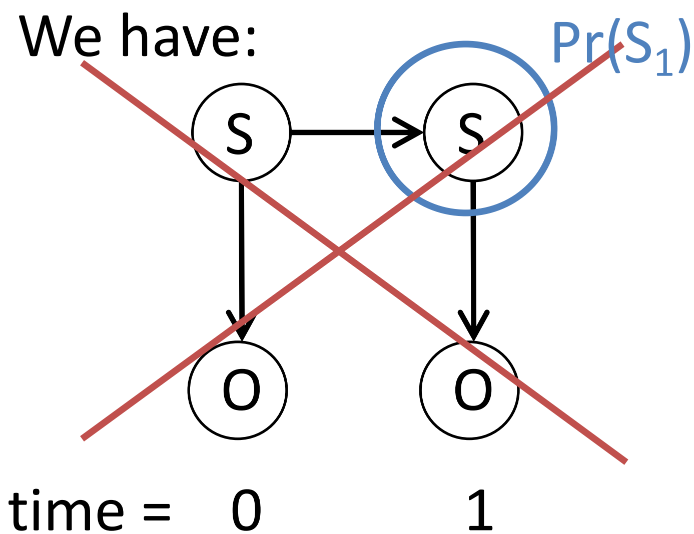


Inference Over Time

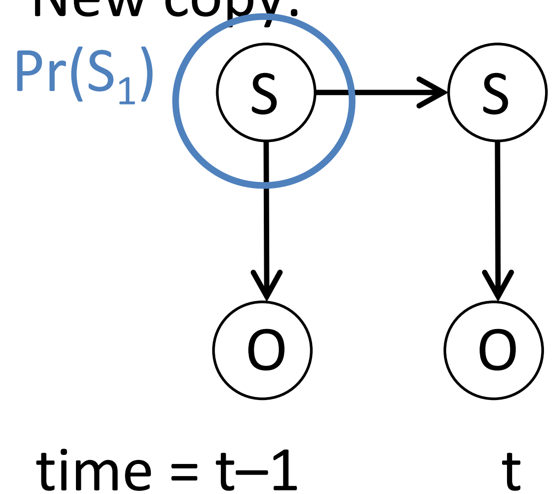
Mimic:



We have:

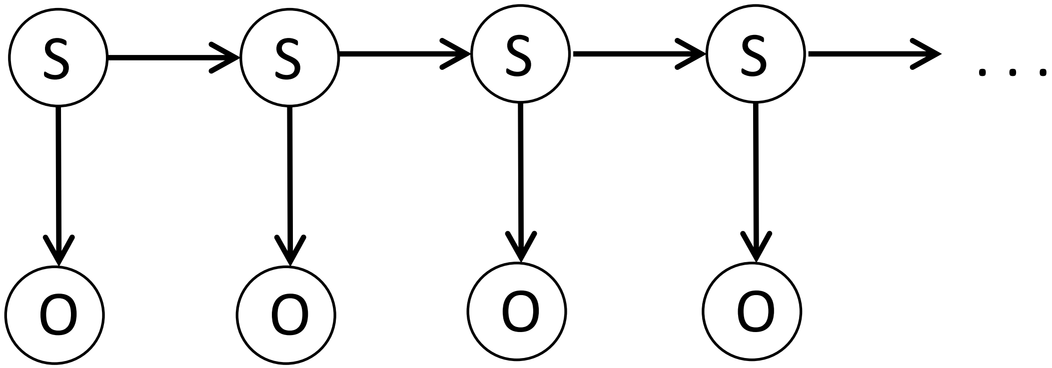


New copy:



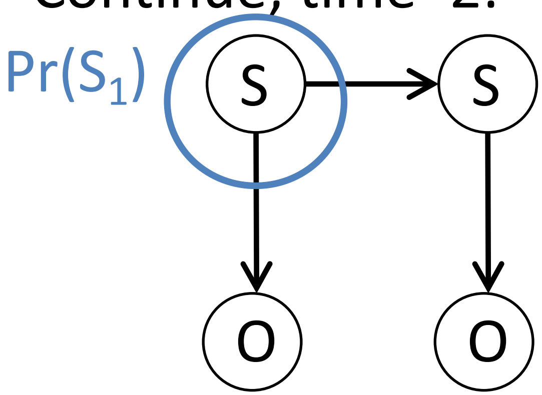
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

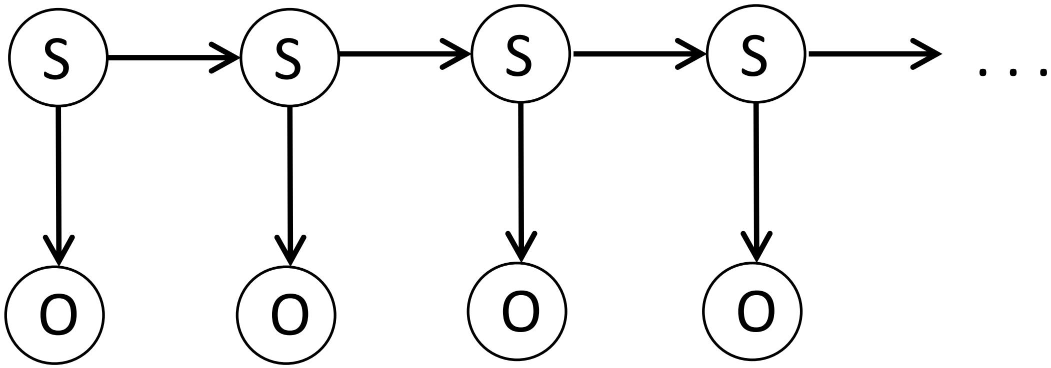
Continue, time=2:



time = 1 2

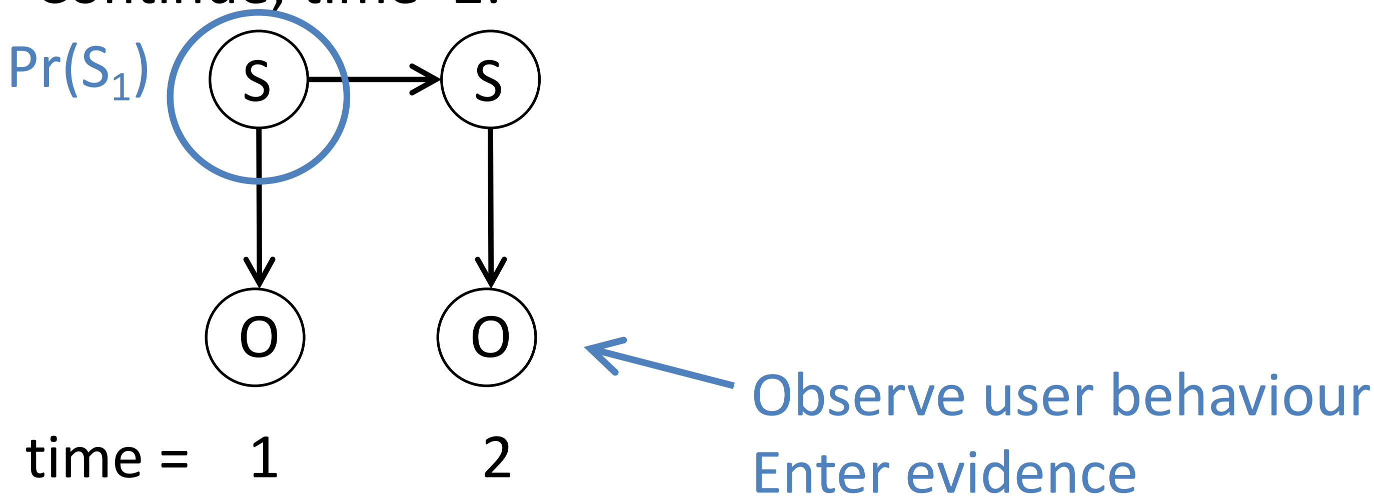
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

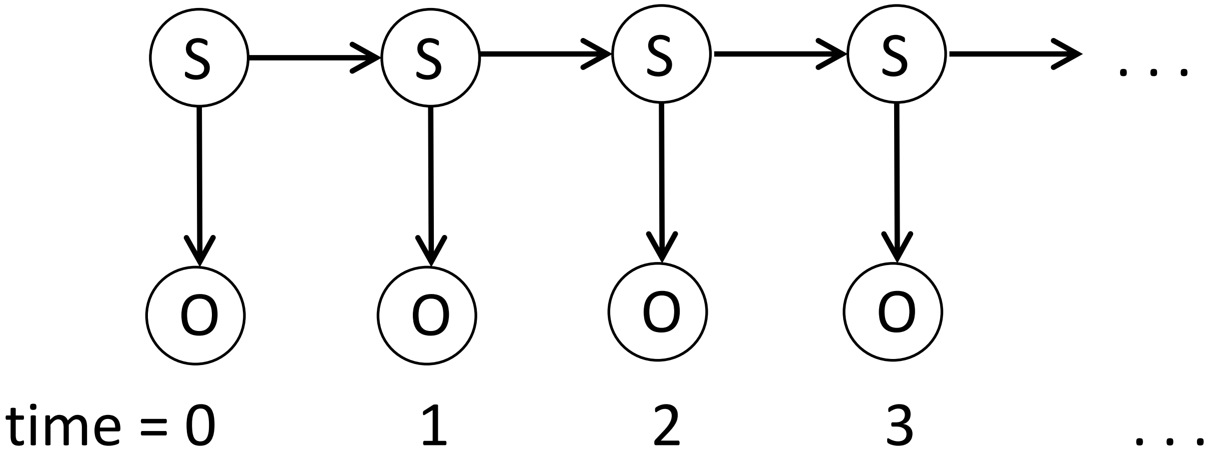
Continue, time=2:



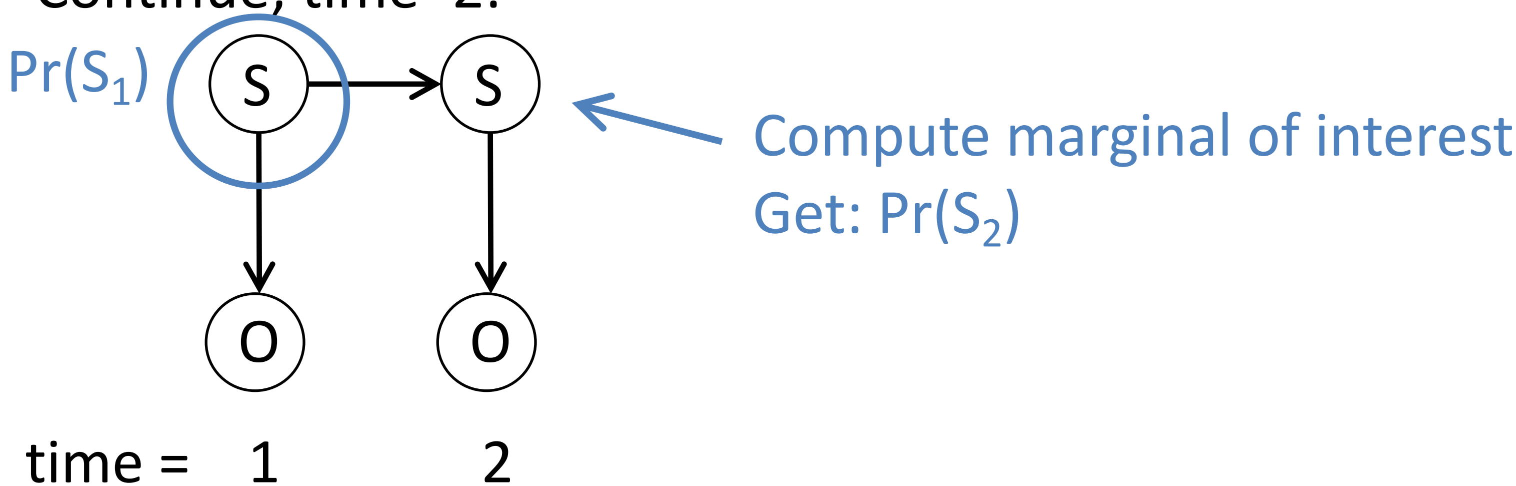
time = 1 2

Inference Over Time

Mimic:

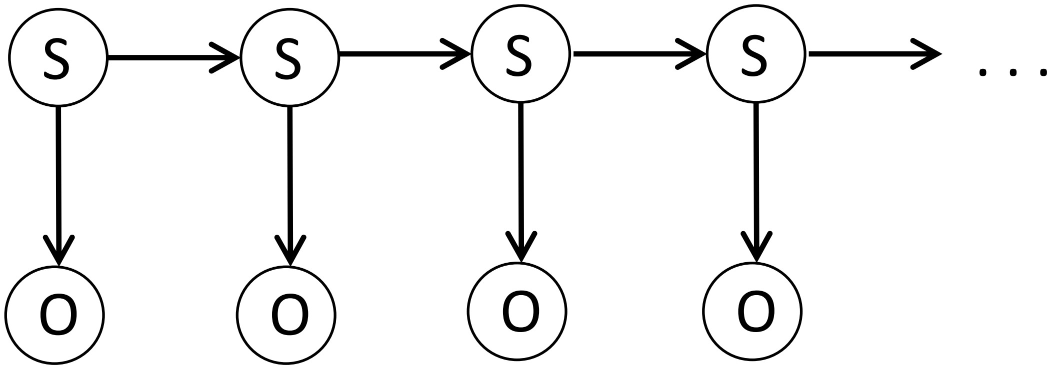


Continue, time=2:



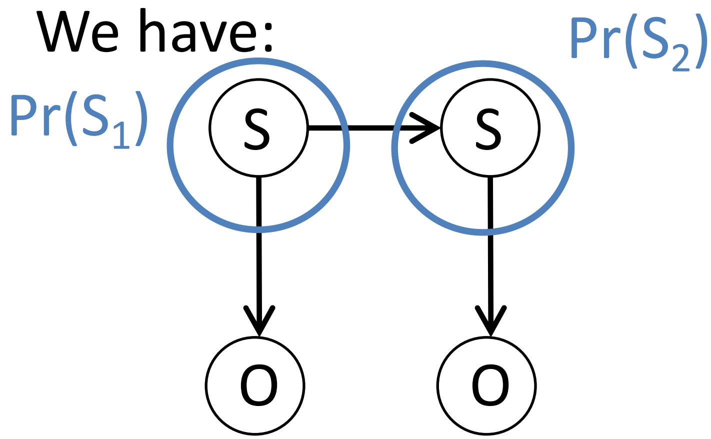
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

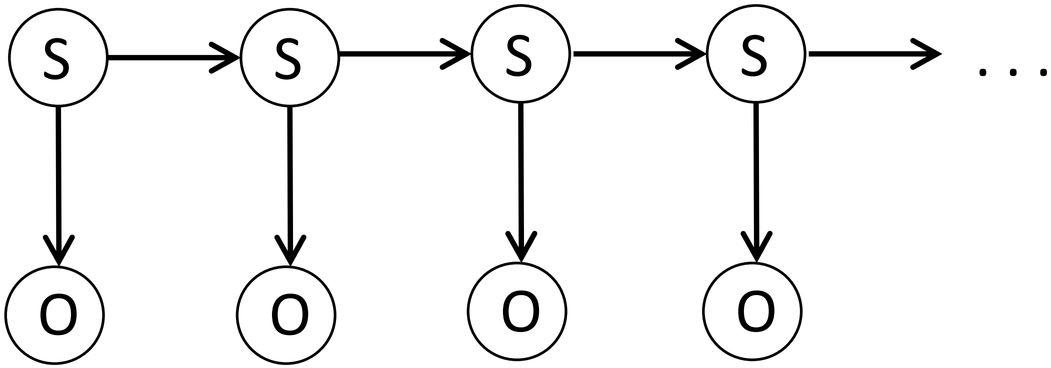
We have:



time = 1 2

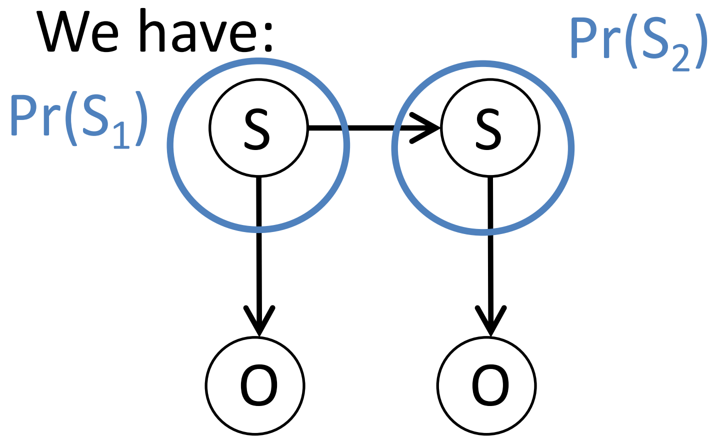
Inference Over Time

Mimic:



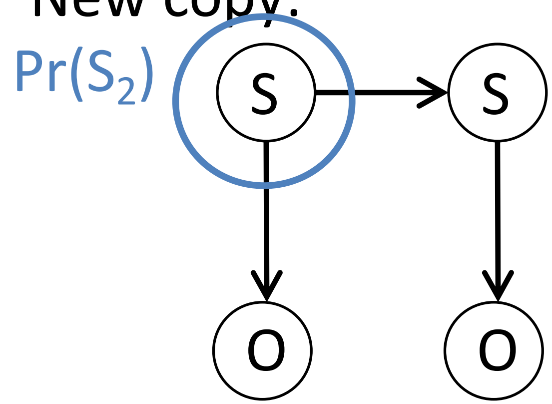
time = 0 1 2 3 . . .

We have:



time = 1 2

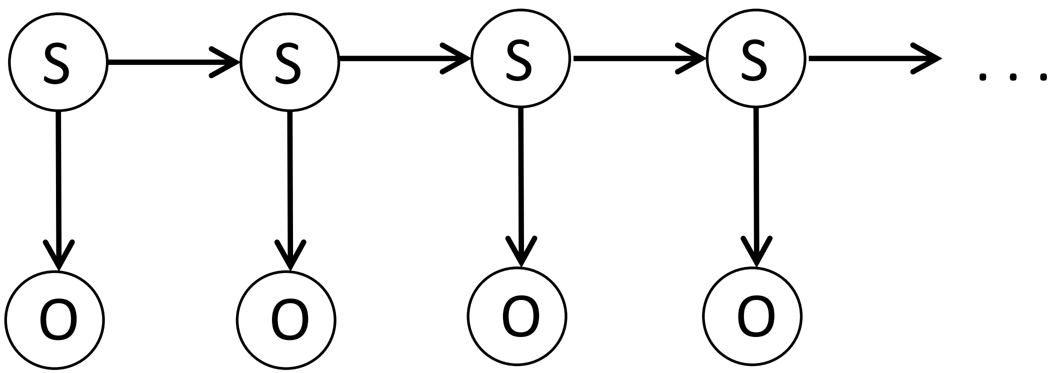
New copy:



time = t–1 t

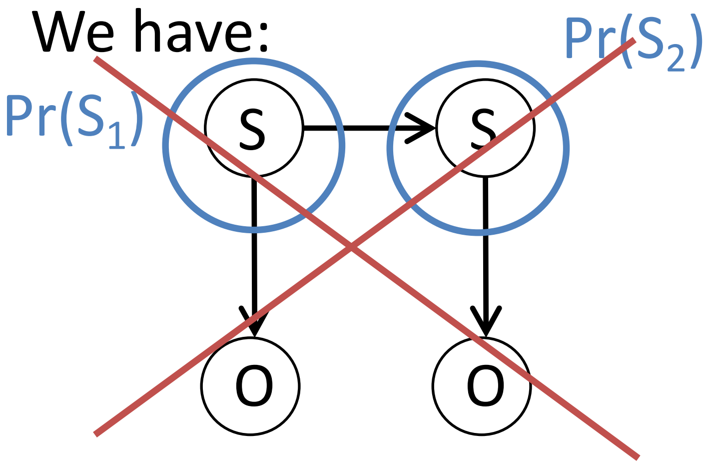
Inference Over Time

Mimic:



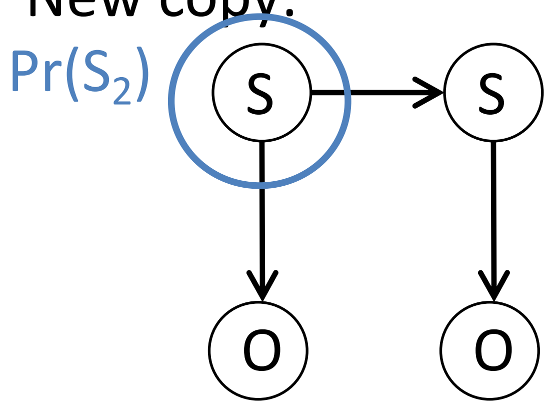
time = 0 1 2 3 . . .

We have:



time = 1 2

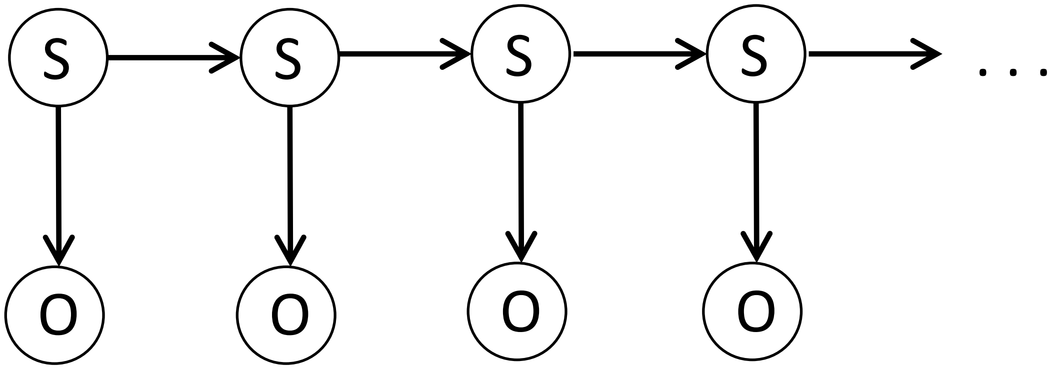
New copy:



time = t–1 t

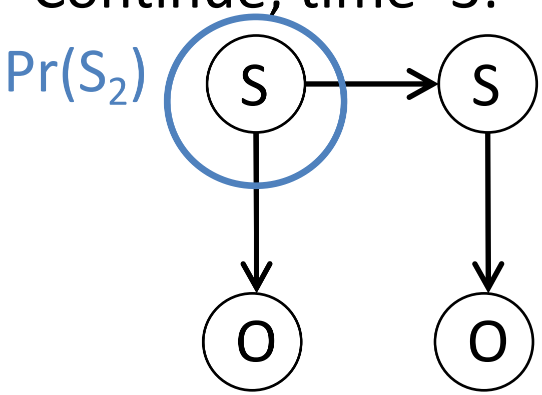
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

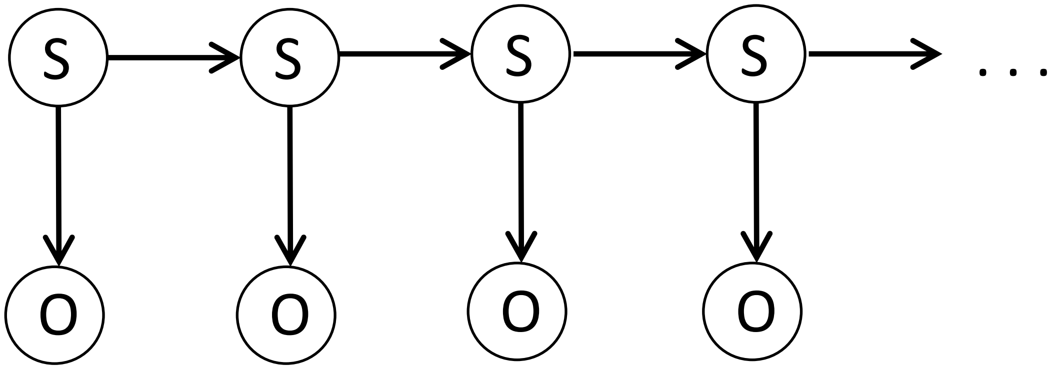
Continue, time=3:



time = 2 3

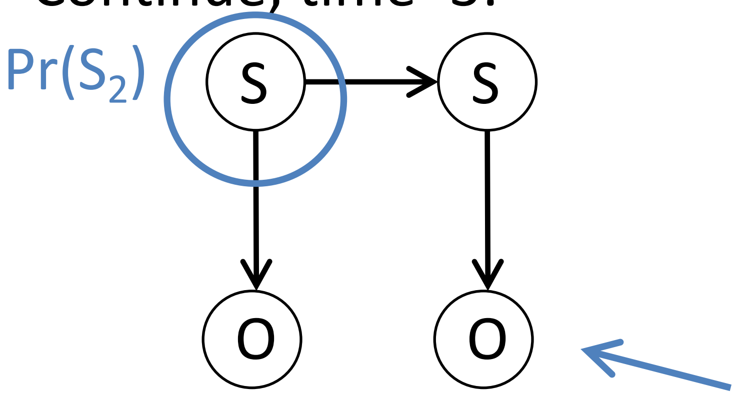
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

Continue, time=3:



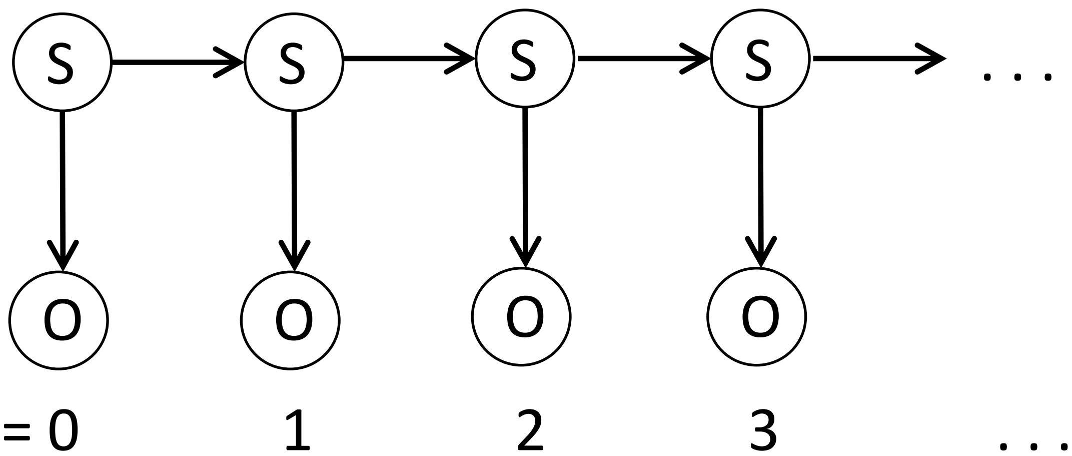
time = 2 3

Observe user behaviour

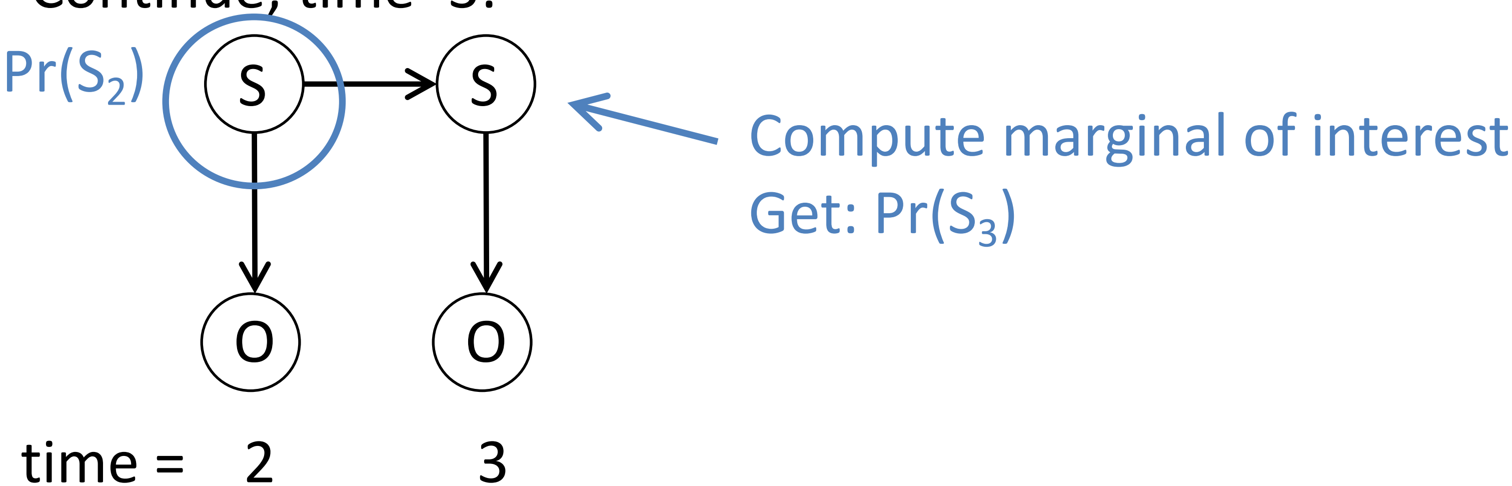
Enter evidence

Inference Over Time

Mimic:

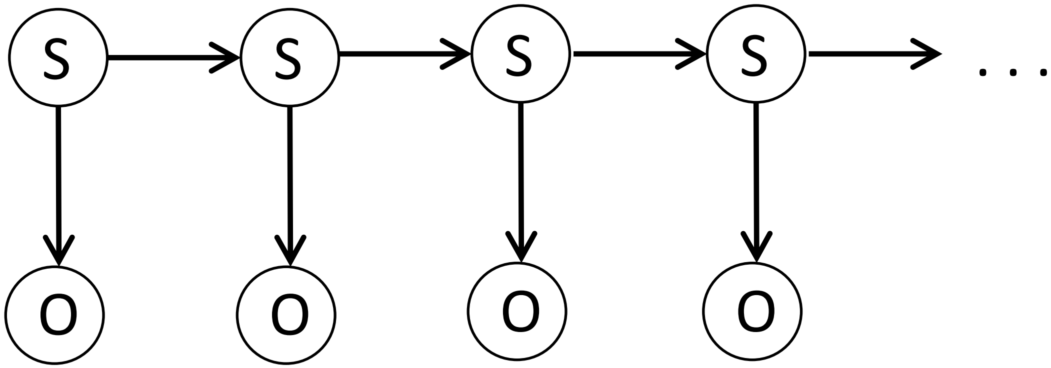


Continue, time=3:



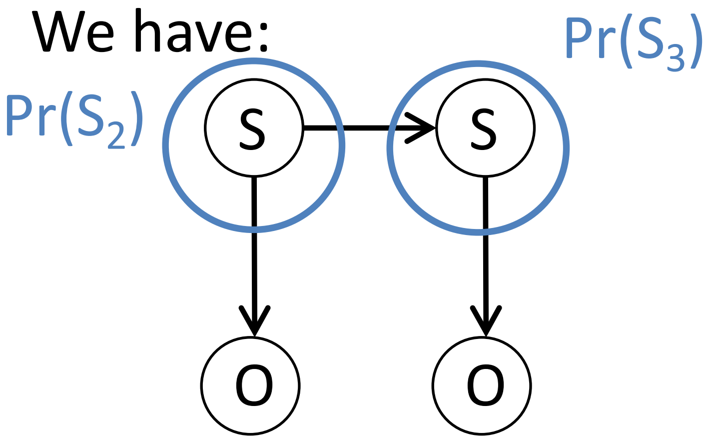
Inference Over Time

Mimic:



time = 0 1 2 3 . . .

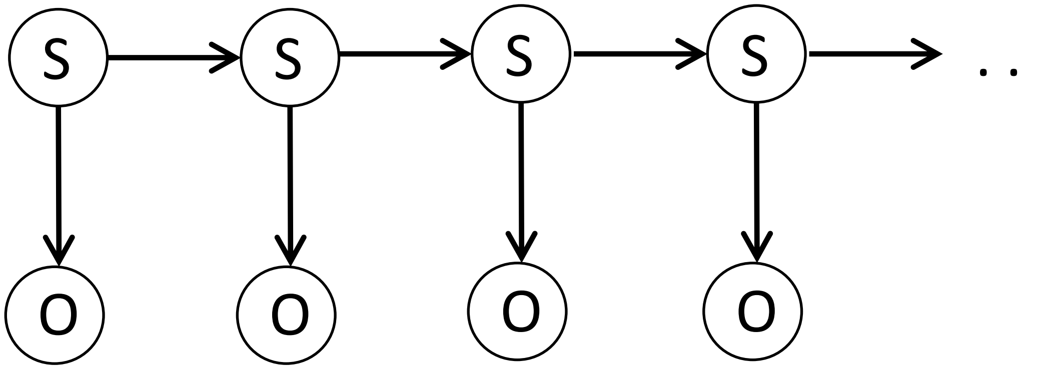
We have:



time = 2 3

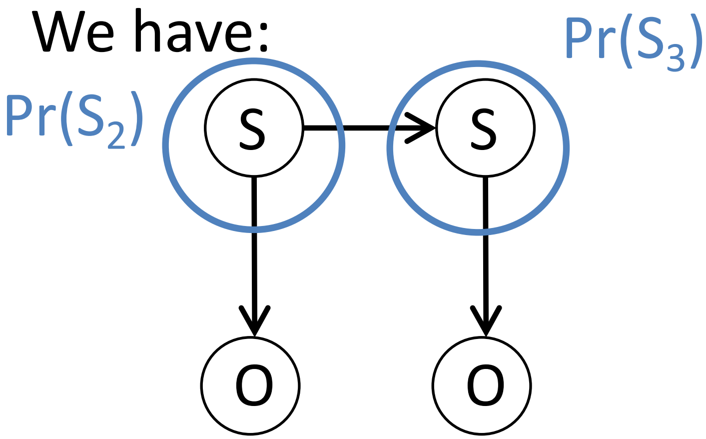
Inference Over Time

Mimic:



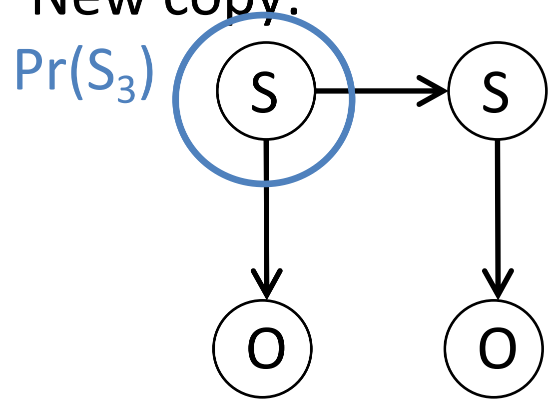
time = 0 1 2 3 . . .

We have:



time = 2 3

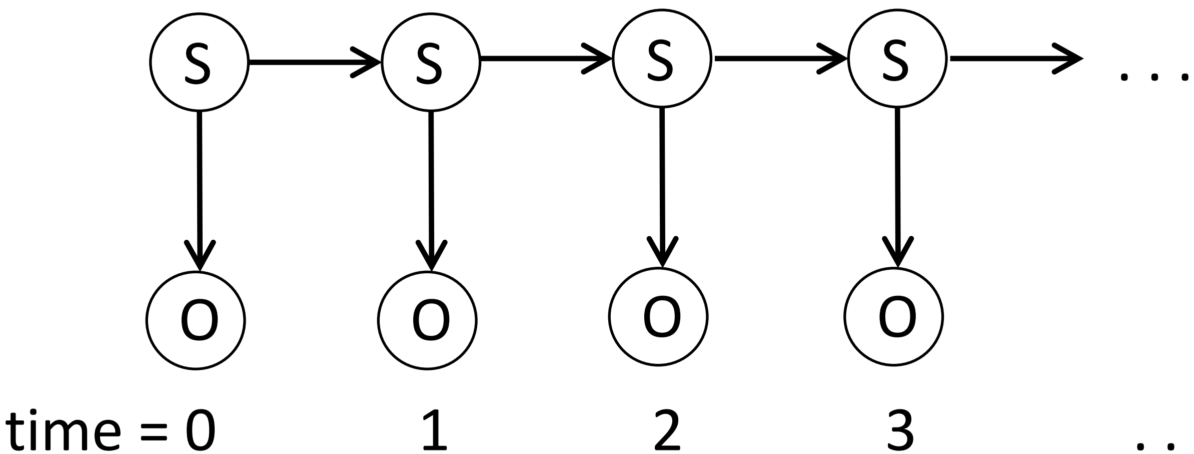
New copy:



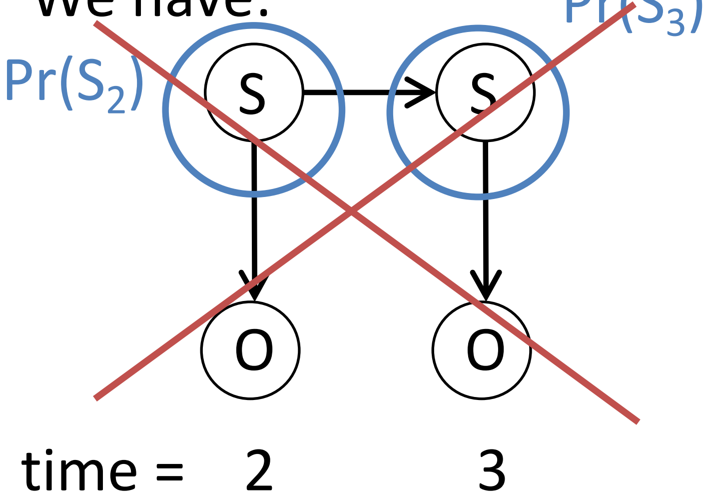
time = t–1 t

Inference Over Time

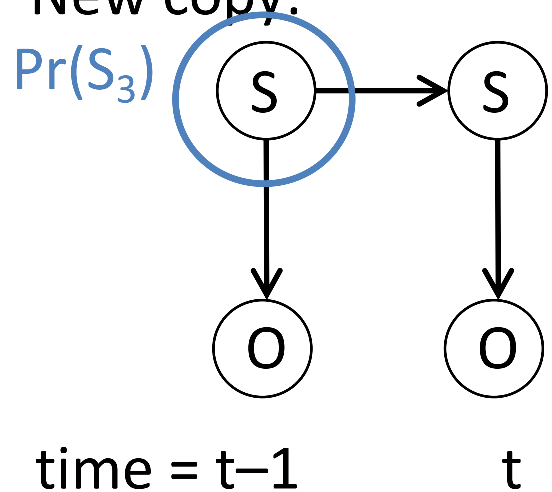
Mimic:



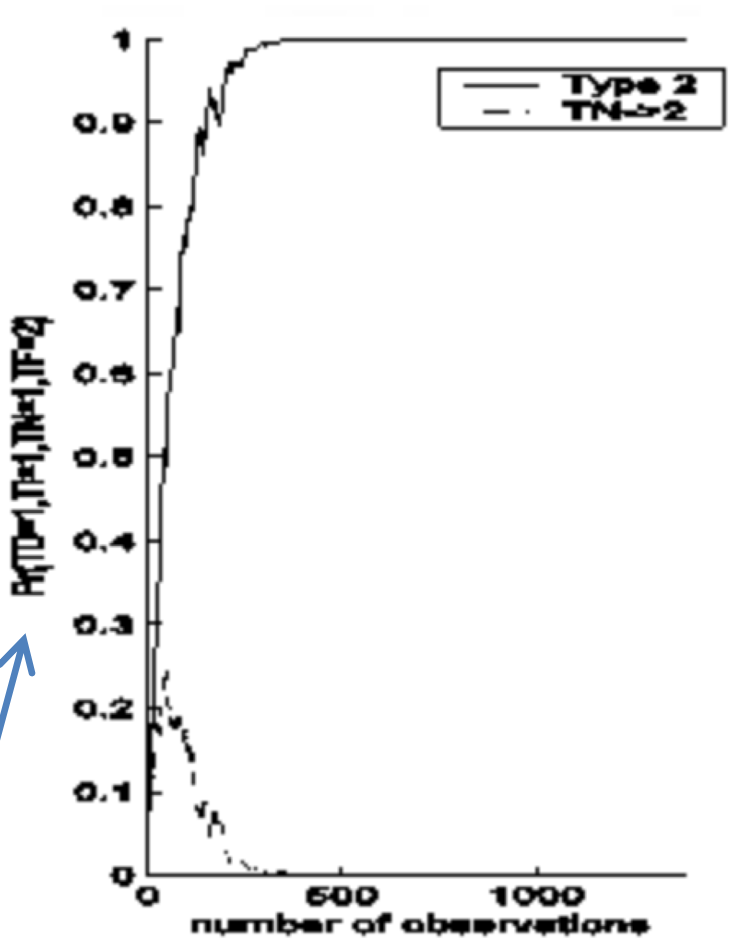
We have:

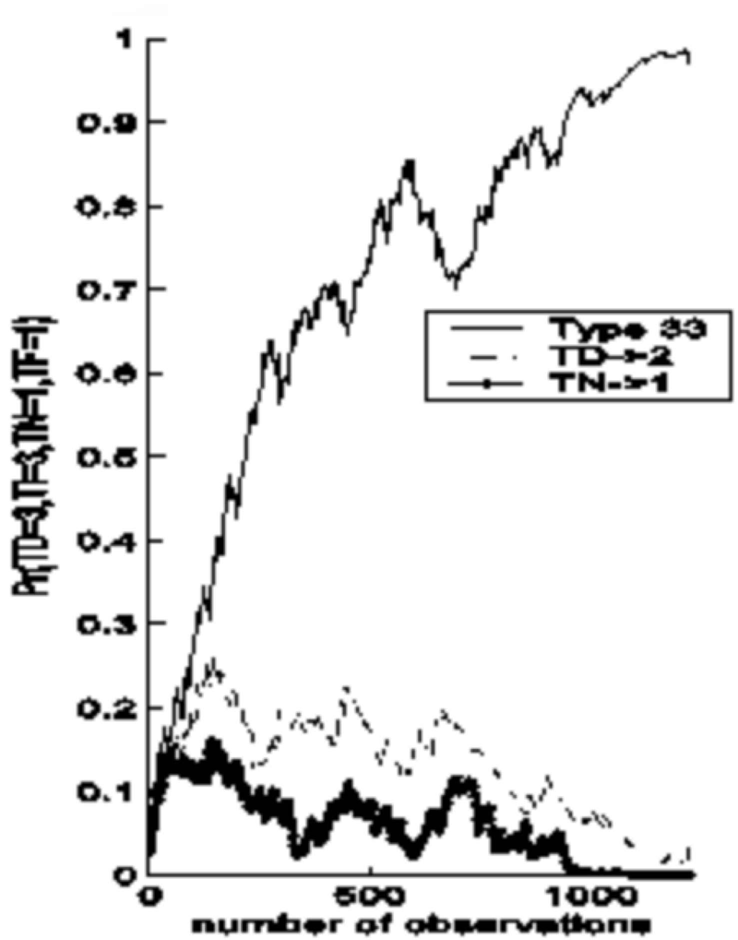


New copy:



Example: Belief Update Over Time





Marginal of interest

Key Ideas

- Main concepts
  - Reasoning over time considers evolving dynamics in the model
- Representation:
  - Stationarity assumption: given model, dynamics don’t change over time
  - Markov assumption: current state depends only on a finite history
  - DBN is an extension of BN with an additional set of temporal relationships (transition functions)