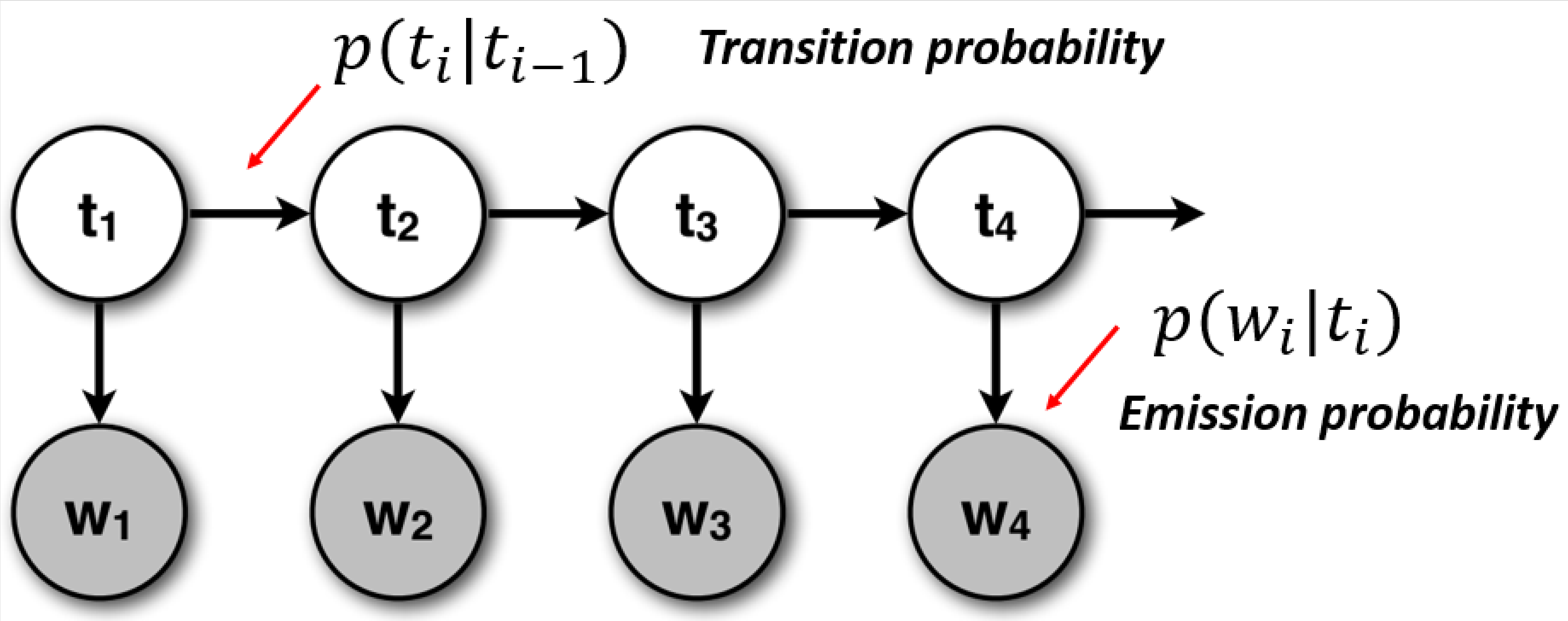
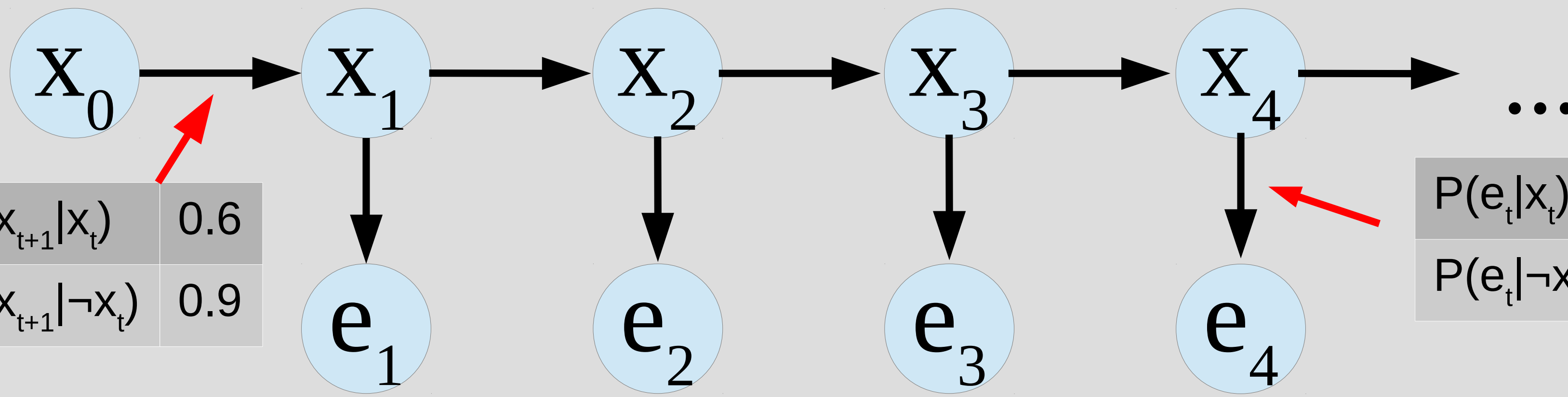
Hidden Markov Models (Ch. 15)



Hidden Markov Models



$P(x_{t+1}\|x_t)$	0.6
$P(x_{t+1}\|\neg x_t)$	0.9

$P(e_t\|x_t)$	0.3
$P(e_t\|\neg x_t)$	0.8

So in this Bayesian network (bigger):

$$P(X_{0:t}, E_{1:t}) = P(X_0) \prod_{i=1}^{t} P(X_i|X_{i-1})P(E_i|X_i)$$

Most Likely Explanation

Typically, use above to compute four things:

Filtering	Prediction	Smoothing	MLE
$P(x_t\|e_{1:t})$	$P(x_{t+k}\|e_{1:t})$	$P(x_k\|e_{1:t}), k < t$	$P(x_{1:t}\|e_{1:t})$

Hidden Markov Model

- $P(x_t|e_{1:t})$ — Filtering
- $P(x_{t+k}|e_{1:t})$ — Prediction
- $P(x_k|e_{1:t}), k < t$ — Smoothing
- $P(x_{1:t}|e_{1:t})$ — Most likely-explanation

Quick recap...

Hidden Markov Model

Filering: $P(x_t|e_{1:t})$

$$f(0) = P(x_0)$$

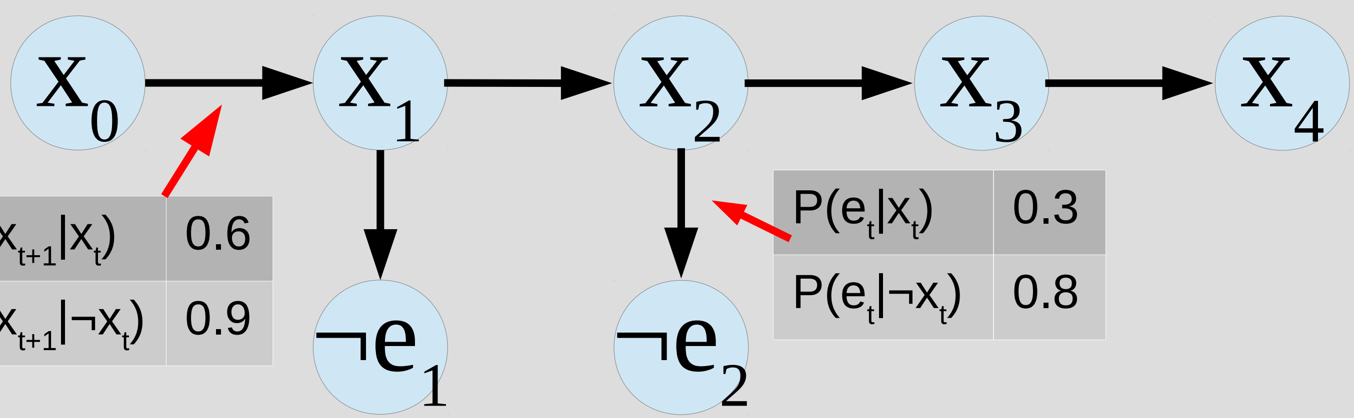
$$f(t) = \alpha P(e_t|x_t) \sum_{x_{t-1}} \Big(P(x_t|x_{t-1}) f(t-1) \Big)$$

Compute P(¬x$_1$|¬e$_1$) and then normalize

$$P(x_1|\neg e_1) = \alpha(1-0.3)\cdot(0.6\cdot 0.5 + 0.9\cdot 0.5) \approx 0.913$$

$$P(x_2|\neg e_1, \neg e_2) = \alpha(1-0.3)(0.6\cdot 0.913 + 0.9\cdot(1-0.913)) \approx 0.854$$

$P(x_0)$	0.5



$P(x_{t+1}|x_t)$	0.6
$P(x_{t+1}|\neg x_t)$	0.9

$P(e_t|x_t)$	0.3
$P(e_t|\neg x_t)$	0.8

Hidden Markov Model

Filtering: $P(x_t|e_{1:t})$ $\qquad f(0) = P(x_0)$

$$f(t) = \alpha P(e_t|x_t) \sum_{x_{t-1}} \Big(P(x_t|x_{t-1}) f(t-1) \Big)$$

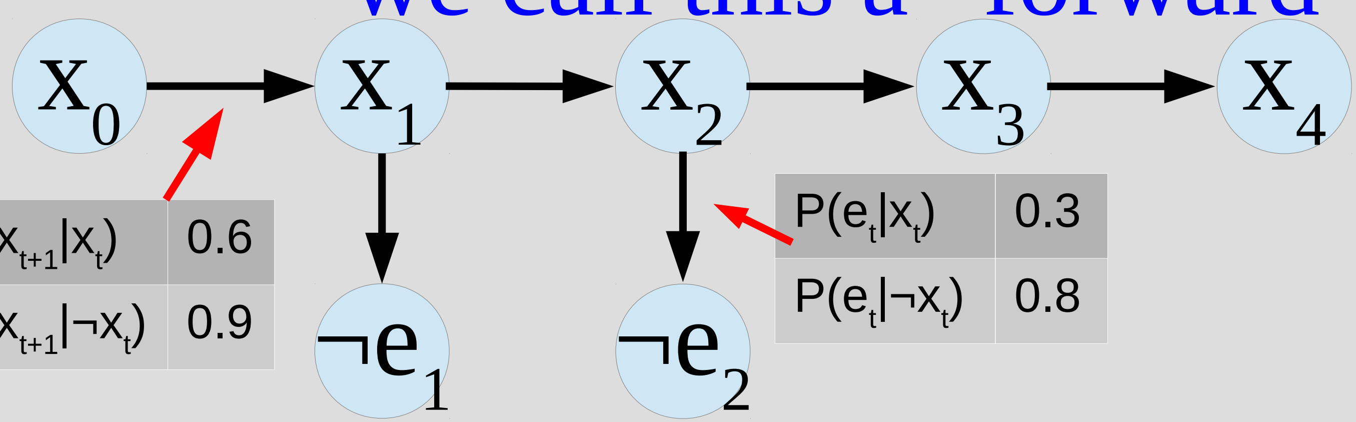
Compute P(¬x$_1$|¬e$_1$) and then normalize

$$P(x_1|\neg e_1) = \alpha(1-0.3)\cdot(0.6\cdot 0.5 + 0.9 \cdot 0.5) \approx \boxed{0.913}$$

$$P(x_2|\neg e_1, \neg e_2) = \alpha(1-0.3)(0.6 \cdot \boxed{0.913} + 0.9 \cdot (1 - \boxed{0.913})) \approx 0.854$$

This storing for the next iteration, we call this a “forward” message

$P(x_0)$	0.5



$P(x_{t+1}|x_t)$	0.6
$P(x_{t+1}|\neg x_t)$	0.9

$P(e_t|x_t)$	0.3
$P(e_t|\neg x_t)$	0.8

Hidden Markov Model

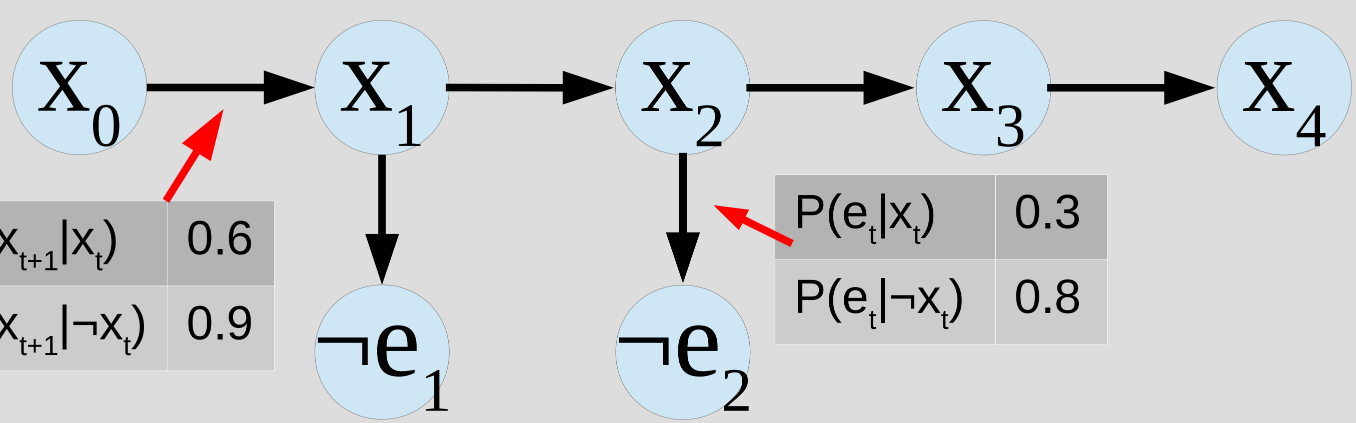
Prediction: $P(x_{t+k}|e_{1:t})$ $g(t) = f(t)$ from filtering

$$g(t) = \alpha \sum_{x_{t-1}} \Big(P(x_t|x_{t-1}) g(t-1) \Big)$$

$P(x_2|\neg e_1, \neg e_2) = f(2) \approx 0.854$

$$P(x_3|\neg e_1, \neg e_2) = \alpha(0.6 \cdot 0.854 + 0.9 \cdot (1 - 0.854)) \approx 0.644$$

$$P(x_4|\neg e_1, \neg e_2) = \alpha(0.6 \cdot 0.644 + 0.9 \cdot (1 - 0.644)) \approx 0.707$$



$P(x_0)$	0.5

$P(x_{t+1}|x_t)$	0.6
$P(x_{t+1}|\neg x_t)$	0.9

$P(e_t|x_t)$	0.3
$P(e_t|\neg x_t)$	0.8

Hidden Markov Model

$$
\begin{aligned}
P(x_2|\neg e_1, \neg e_2) &= \alpha P(x_2, \neg e_1, \neg e_2) \\
&= \alpha \sum_{x_0} \sum_{x_1} P(x_0, x_1, x_2 \neg e_1, \neg e_2) \\
&= \alpha \sum_{x_0} \sum_{x_1} P(x_0) P(x_1|x_0) P(\neg e_1|x_1) P(x_2|x_1) P(\neg e_2|x_2) \\
&= \alpha \sum_{x_1} \sum_{x_0} P(\neg e_2|x_2) P(x_2|x_1) P(\neg e_1|x_1) P(x_0) P(x_1|x_0)
\end{aligned}
$$

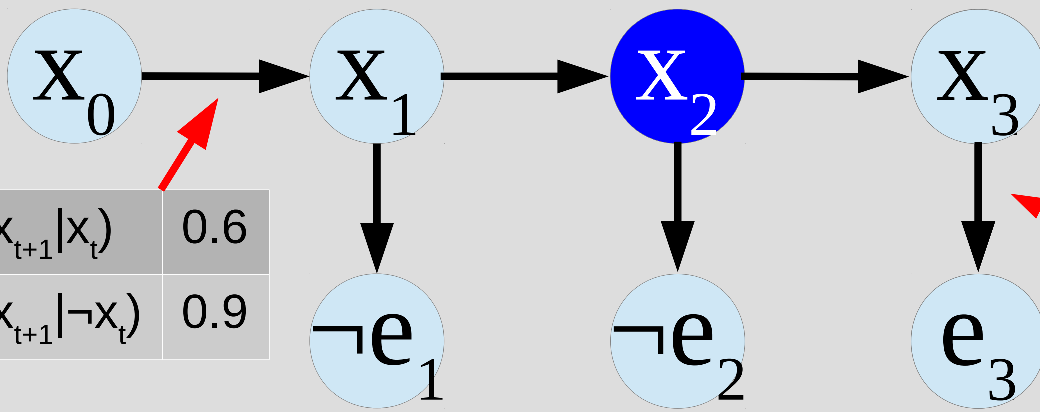
$P(x_1|e_1)$, last message

$$
\begin{aligned}
&= \alpha P(\neg e_2|x_2) \sum_{x_1} P(x_2|x_1) \boxed{P(\neg e_1|x_1) \sum_{x_0} P(x_0) P(x_1|x_0)} \\
&= \alpha P(\neg e_2|x_2) \big(P(x_2|x_1) \cdot 0.913 + P(x_2|\neg x_1) \cdot (1 - 0.913) \big) \\
&= \alpha (1 - 0.3)(0.6 \cdot 0.913 + 0.9 \cdot (1 - 0.913)) \\
&\approx 0.438\alpha
\end{aligned}
$$

... after normalizing you should get: ≈0.854

Smoothing in HMMs

$P(x_0)$	0.5



$P(x_{t+1} \mid x_t)$	0.6
$P(x_{t+1} \mid \neg x_t)$	0.9

$P(e_t \mid x_t)$	0.3
$P(e_t \mid \neg x_t)$	0.8

Note: $P(x_2 \mid \neg e_1, \neg e_2) = 0.854$

Find: $P(x_2 \mid \neg e_1, \neg e_2, e_3)$

Smoothing in HMMs

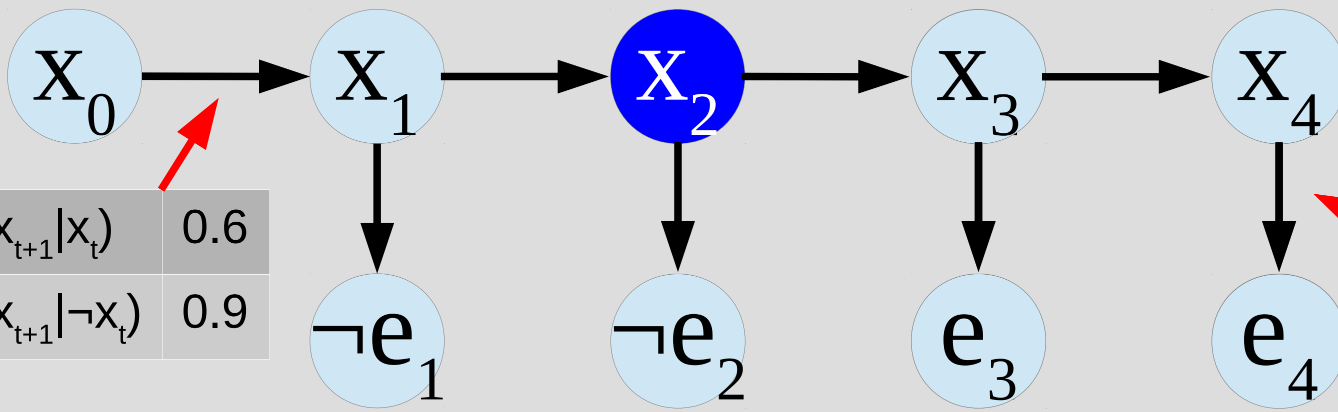
$$
\begin{aligned}
P(x_2|\neg e_1,\neg e_2,e_3) &= \alpha P(x_2,\neg e_1,\neg e_2,e_3)\\
&= \alpha \sum_{x_0}\sum_{x_1}\sum_{x_3} P(x_0,x_1,x_2,x_3,\neg e_1,\neg e_2,e_3)\\
&= \alpha \sum_{x_0}\sum_{x_1}\sum_{x_3} P(x_0)P(x_1|x_0)P(\neg e_1|x_1)P(x_2|x_1)P(\neg e_2|x_2)P(x_3|x_2)P(e_3|x_2)\\
&= \alpha \sum_{x_3}\sum_{x_1}\sum_{x_0} P(e_3|x_3)P(x_3|x_2)P(\neg e_2|x_2)P(x_2|x_1)P(\neg e_1|x_1)P(x_0)P(x_1|x_0)\\
&= \alpha \sum_{x_3} P(e_3|x_3)P(x_3|x_2)P(\neg e_2|x_2)\sum_{x_1} P(x_2|x_1)P(\neg e_1|x_1)\sum_{x_0} P(x_0)P(x_1|x_0)\\
&= \alpha \sum_{x_3} P(e_3|x_3)P(x_3|x_2)P(x_2|\neg e_1,\neg e_2)\\
&\approx \alpha\,(\underbrace{0.3\cdot 0.6\cdot 0.854}_{x_3,x_3=true}+\underbrace{0.8\cdot(1-0.6)\cdot 0.854)}_{\neg x_3,x_3=false}\\
&\approx 0.427\alpha
\end{aligned}
$$

$$
\boxed{\begin{aligned}
&P(\neg x_2|\neg e_1,\neg e_2,e_3)\\
&\approx \alpha 0.3\cdot 0.9\cdot(1-0.854)+0.8\cdot(1-0.9)\cdot(1-0.854)\\
&\approx 0.0511\alpha
\end{aligned}}
$$

... after normalizing, $P(x_2|\neg e_1,\neg e_2,e_3)$ ≈0.89311

Smoothing in HMMs

$P(x_0)$	0.5



$P(x_{t+1}\|x_t)$	0.6
$P(x_{t+1}\|\neg x_t)$	0.9

$P(e_t\|x_t)$	0.3
$P(e_t\|\neg x_t)$	0.8

Note: $P(x_2|\neg e_1,\neg e_2) = 0.854$

Note: $P(x_2|\neg e_1,\neg e_2,e_3) = 0.893$

Find: $P(x_2|\neg e_1,\neg e_2,e_3,e_4)$

Smoothing in HMMs

$$P(x_2|\neg e_1,\neg e_2,e_3,e_4) = \alpha P(x_2,\neg e_1,\neg e_2,e_3,e_4)$$

$$= \alpha \sum_{x_0}\sum_{x_1}\sum_{x_3}\sum_{x_4} P(x_0,x_1,x_2,x_3,x_4,\neg e_1,\neg e_2,e_3,e_4)$$

$$= \alpha \sum_{x_0}\sum_{x_1}\sum_{x_3}\sum_{x_4} P(x_0)P(x_1|x_0)P(\neg e_1|x_1)P(x_2|x_1)P(\neg e_2|x_2)P(x_3|x_2)P(e_3|x_2)P(x_4|x_3)P(e_4|x_4)$$

$$= \alpha \sum_{x_4}\sum_{x_3}\sum_{x_1}\sum_{x_0} P(e_4|x_4)P(x_4|x_3)P(e_3|x_3)P(x_3|x_2)P(\neg e_2|x_2)P(x_2|x_1)P(\neg e_1|x_1)P(x_0)P(x_1|x_0)$$

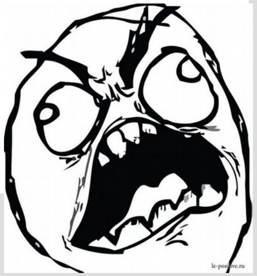
$$= \alpha \sum_{x_4} P(e_4|x_4)\sum_{x_3} P(x_4|x_3)P(e_3|x_3)P(x_3|x_2)P(\neg e_2|x_2)\sum_{x_1} P(x_2|x_1)P(\neg e_1|x_1)\sum_{x_0} P(x_0)P(x_1|x_0)$$

$$= \alpha \sum_{x_4} P(e_4|x_4)\sum_{x_3} P(x_4|x_3)P(e_3|x_3)P(x_3|x_2)P(x_2|\neg e_1,\neg e_2)$$

Smoothing in HMMs

$$
\begin{aligned}
P(x_2 \mid \neg e_1, \neg e_2, e_3, e_4) &= \alpha P(x_2, \neg e_1, \neg e_2, e_3, e_4) \\
&= \alpha \sum_{x_0} \sum_{x_1} \sum_{x_3} \sum_{x_4} P(x_0, x_1, x_2, x_3, x_4, \neg e_1, \neg e_2, e_3, e_4) \\
&= \alpha \sum_{x_0} \sum_{x_1} \sum_{x_3} \sum_{x_4} P(x_0) P(x_1 \mid x_0) P(\neg e_1 \mid x_1) P(x_2 \mid x_1) P(\neg e_2 \mid x_2) P(x_3 \mid x_2) P(e_3 \mid x_2) P(x_4 \mid x_3) P(e_4 \mid x_4) \\
&= \alpha \sum_{x_4} \sum_{x_3} \sum_{x_1} \sum_{x_0} P(e_4 \mid x_4) P(x_4 \mid x_3) P(e_3 \mid x_3) P(x_3 \mid x_2) P(\neg e_2 \mid x_2) P(x_2 \mid x_1) P(\neg e_1 \mid x_1) P(x_0) P(x_1 \mid x_0) \\
&= \alpha \sum_{x_4} P(e_4 \mid x_4) \sum_{x_3} P(x_4 \mid x_3) P(e_3 \mid x_3) P(x_3 \mid x_2) P(\neg e_2 \mid x_2) \sum_{x_1} P(x_2 \mid x_1) P(\neg e_1 \mid x_1) \sum_{x_0} P(x_0) P(x_1 \mid x_0) \\
&= \alpha \sum_{x_4} P(e_4 \mid x_4) \sum_{x_3} \boxed{P(x_4 \mid x_3)} P(e_3 \mid x_3) P(x_3 \mid x_2) P(x_2 \mid \neg e_1, \neg e_2)
\end{aligned}
$$

This term was not in our last calculation



Smoothing in HMMs

Instead, lets put x_4 sum inside x_3 sum:

$$P(x_2|\neg e_1, \neg e_2, e_3, e_4) = \alpha \sum_{x_4} P(e_4|x_4) \sum_{x_3} P(x_4|x_3)P(e_3|x_3)P(x_3|x_2)P(x_2|\neg e_1, \neg e_2)$$

$$= \alpha P(x_2|\neg e_1, \neg e_2) \sum_{x_3} P(e_3|x_3)P(x_3|x_2) \sum_{x_4} P(e_4|x_4)P(x_4|x_3)$$

So... similarly if we knew e_5:

$$P(x_2|\neg e_1, \neg e_2, e_3, e_4, e_5) = \alpha P(x_2|\neg e_1, \neg e_2) \sum_{x_3} P(e_3|x_3)P(x_3|x_2) \sum_{x_4} P(e_4|x_4)P(x_4|x_3) \sum_{x_5} P(e_5|x_5)P(x_5|x_4)$$

calc first

This time the "inner most" is for large t's, so rather than a "forward" message, it's a "backwards" message (starting with large t)

Smoothing in HMMs

P(a,b) = P(a|b)P(b)
(with a conditional everywhere)

$$
\begin{aligned}
P(x_k|e_{1:t}) &= P(x_k|e_{1:k}, e_{k+1:t}) \\
&= \alpha P(x_k, e_{k+1:t}|e_{1:k}) \\
&= \alpha P(e_{k+1:t}|x_k, e_{1:k}) P(x_k|e_{1:k}) \\
&= \alpha P(e_{k+1:t}|x_k) P(x_k|e_{1:k})
\end{aligned}
$$

... where:

$$
\begin{aligned}
P(e_{k+1:t}|x_k) &= \sum_{x_{k+1}} P(e_{k+1:t}, x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1:t}|x_{k+1}, x_k) P(x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1}, e_{k+2:t}|x_{k+1}) P(x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1}|x_{k+1}) P(e_{k+2:t}|x_{k+1}, e_{k+1}) P(x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1}|x_{k+1}) P(e_{k+2:t}|x_{k+1}) P(x_{k+1}|x_k)
\end{aligned}
$$

Smoothing in HMMs

P(a,b) = P(a|b)P(b)
(with a conditional everywhere)

$$
\begin{aligned}
P(x_k|e_{1:t}) &= P(x_k|e_{1:k}, e_{k+1:t}) \\
&= \alpha P(x_k, e_{k+1:t}|e_{1:k}) \\
&= \alpha P(e_{k+1:t}|x_k, e_{1:k}) P(x_k|e_{1:k}) \\
&= \alpha P(e_{k+1:t}|x_k) P(x_k|e_{1:k})
\end{aligned}
$$

... where:

$$
\begin{aligned}
\boxed{P(e_{k+1:t}|x_k)} &= \sum_{x_{k+1}} P(e_{k+1:t}, x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1:t}|x_{k+1}, x_k) P(x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1}, e_{k+2:t}|x_{k+1}) P(x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1}|x_{k+1}) P(e_{k+2:t}|x_{k+1}, e_{k+1}) P(x_{k+1}|x_k) \\
&= \sum_{x_{k+1}} P(e_{k+1}|x_{k+1}) \boxed{P(e_{k+2:t}|x_{k+1})} P(x_{k+1}|x_k)
\end{aligned}
$$

recursive

Smoothing in HMMs

Thus for smoothing we have a recursive func:

$$h(k) = \sum_{x_{k+1}} P(e_{k+1}|x_{k+1})h(k+1)P(x_{k+1}|x_k)$$

... where: $h(t) =< 1, 1 >$

Then the final smoothing is:

$P(x_k|e_{1:t}) = \alpha f(t) * h(t)$

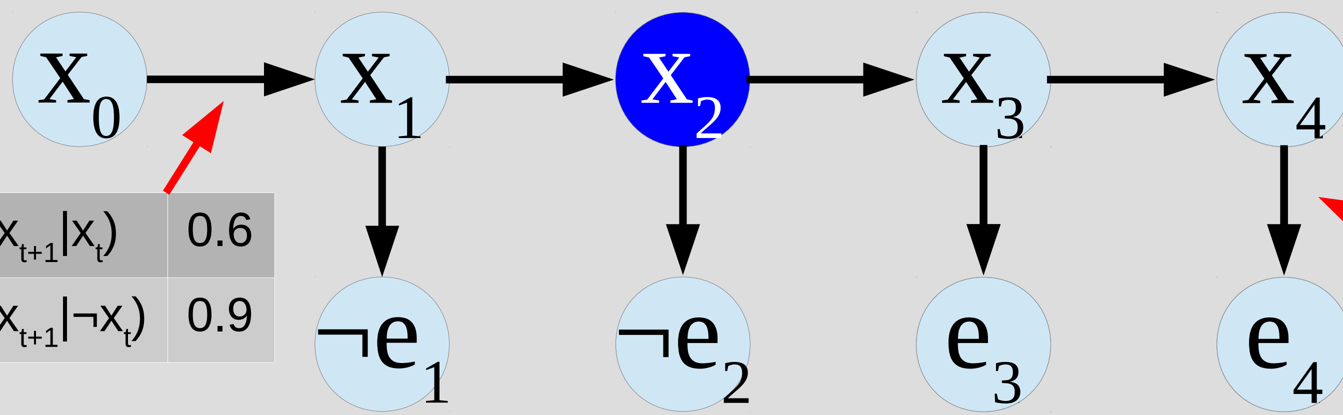
unlike examples only need to normalize at end

... where you take the point-wise product of f(t) and h(t) (i.e. $<f_{true}*h_{true}, f_{false}*h_{false}>$

Smoothing in HMMs

$P(x_0)$	0.5

$h(4) =< 1, 1 >$



$P(x_{t+1}\|x_t)$	0.6
$P(x_{t+1}\|\neg x_t)$	0.9

$P(e_t\|x_t)$	0.3
$P(e_t\|\neg x_t)$	0.8

Note: $P(x_2|\neg e_1,\neg e_2) = 0.854$

Find: $P(x_2|\neg e_1,\neg e_2,e_3,e_4) = 0.8788$ (below)

$$h(3) = 0.6 \cdot h(4) \cdot 0.3 + (1-0.6) \cdot h(\neg 4) \cdot 0.8 = 0.6 \cdot 1 \cdot 0.3 + (1-0.6) \cdot 1 \cdot 0.8 = 0.5$$

$$h(\neg 3) = 0.9 \cdot h(4) \cdot 0.3 + (1-0.9) \cdot h(\neg 4) \cdot 0.8 = 0.9 \cdot 1 \cdot 0.3 + (1-0.9) \cdot 1 \cdot 0.8 = 0.35$$

$$h(2) = 0.6 \cdot h(3) \cdot 0.3 + (1-0.6) \cdot h(\neg 3) \cdot 0.8 = 0.6 \cdot 0.5 \cdot 0.3 + (1-0.6) \cdot 0.35 \cdot 0.8 = 0.202$$

$$h(\neg 2) = 0.9 \cdot h(3) \cdot 0.3 + (1-0.9) \cdot h(\neg 3) \cdot 0.8 = 0.9 \cdot 0.5 \cdot 0.3 + (1-0.9) \cdot 0.35 \cdot 0.8 = 0.163$$

$$P(x_2|\neg e_1, \neg e_2, e_3, e_4) = \alpha 0.202 \cdot 0.854$$

$$P(\neg x_2|\neg e_1, \neg e_2, e_3, e_4) = \alpha 0.163 \cdot (1 - 0.854)$$

Smoothing in HMMs

Side note: for smoothing it takes O(n) to compute for a single x_n

If you wanted to compute for all days (n of them) it would take $O(n^2)$

However you can get it in O(n) (≈2n) if you compute all backwards messages: h(n) ... h(1) and all forward: f(1), ... f(n)
Then do on day i you have: $\alpha * f(i) * h(i)$

Most Likely Explanation

TO DO:

- $P(x_t|e_{1:t})$ Filtering
- $P(x_{t+k}|e_{1:t})$ Prediction
- $P(x_k|e_{1:t}), k < t$ Smoothing
- $P(x_{1:t}|e_{1:t})$ Most likely-explanation

One more to go....

Most Likely Explanation

So far we have been looking at probabilities of individual x_ns being true/false

What if we wanted to know the most likely explanation on a single day/x_n, but for all?

Most Likely Explanation

So far we have been looking at probabilities of individual x_ns being true/false

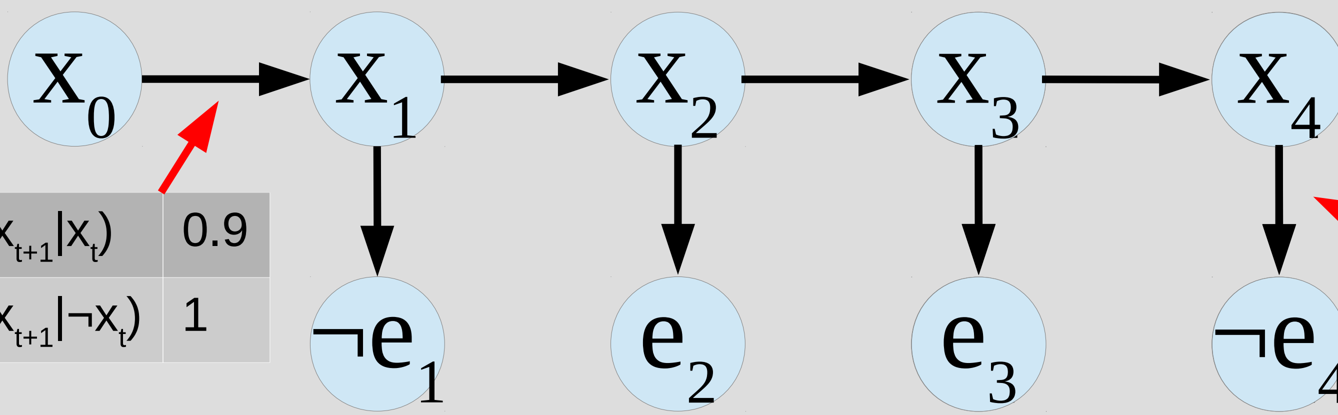
What if we wanted to know the most likely explanation on a single day/x_n, but for all?

Unfortunately... you cannot use smoothing on each day individually (as we summed over other days in individual calculation)

Most Likely Explanation

Consider this example:

$P(x_0)$	0.5



$P(x_{t+1} \mid x_t)$	0.9
$P(x_{t+1} \mid \neg x_t)$	1

$P(e_t \mid x_t)$	0.01
$P(e_t \mid \neg x_t)$	1

bad rounding (not enough space)

Using smoothing this would give:

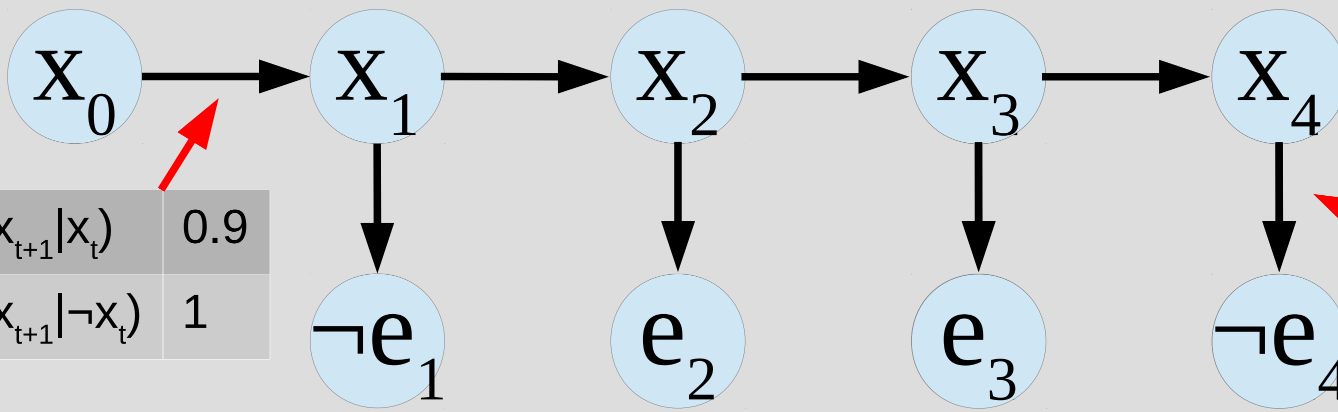
F: <1,0> <0.08, 0.92> <0.55,0.45> <1, 0>

B:<something> <0.92,0.08> <0.47, 0.53>

S: <1,0> <0.52, 0.48> <0.52, 0.48> <1,0>

Most Likely Explanation

$P(x_0)$	0.5



$P(x_{t+1}\|x_t)$	0.9
$P(x_{t+1}\|\neg x_t)$	1

$P(e_t\|x_t)$	0.01
$P(e_t\|\neg x_t)$	1

S: <1,0> <0.52, 0.48> <0.52, 0.48> <1,0>

So using smoothing we get:

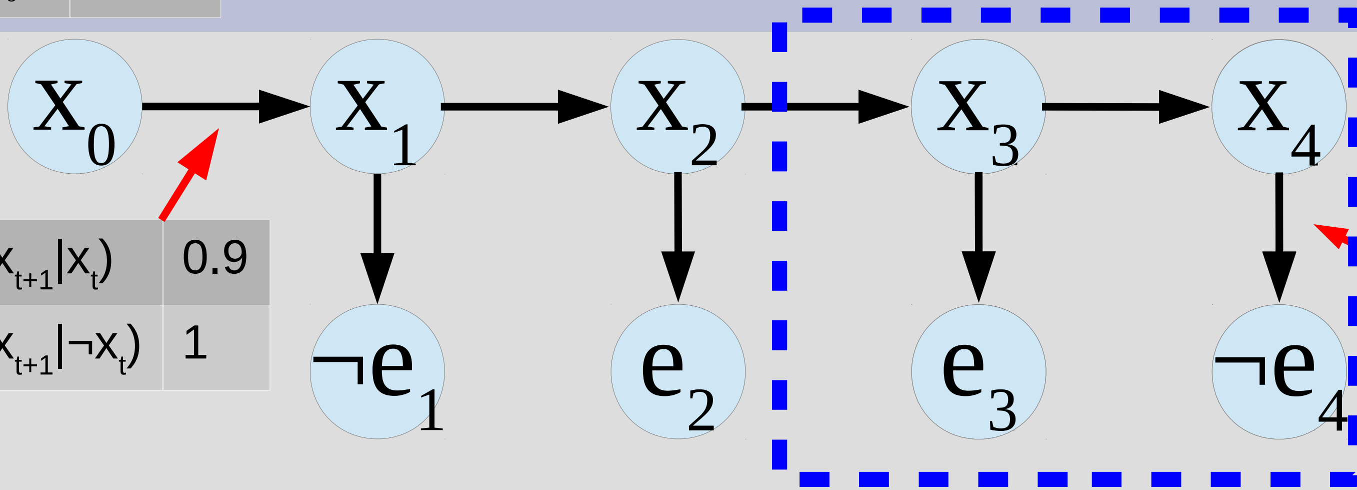
x_1=true, x_2=true, x_3=true, x_4=true...

... This is very wrong (x_2 or x_3 should be false)

Most Likely Explanation

$P(x_0)$	0.5

$P(x_{t+1} \mid x_t)$	0.9
$P(x_{t+1} \mid \neg x_t)$	1



$P(e_t \mid x_t)$	0.01
$P(e_t \mid \neg x_t)$	1

ignore for a second

Instead of a sum:

$$P(x_2 \mid \neg e_1, e_2) = \alpha \sum_{x_0} \sum_{x_1} P(x_0) P(x_1 \mid x_0) P(\neg e_1 \mid x_1) P(x_2 \mid x_1) P(e_2 \mid x_2)$$

We want to max (all variables):

$$\max_{x_1, x_2} P(x_1, x_2 \mid \neg e_1, e_2) = \alpha \max_{x_0} \max_{x_1} P(x_0) P(x_1 \mid x_0) P(\neg e_1 \mid x_1) P(x_2 \mid x_1) P(e_2 \mid x_2)$$

Most Likely Explanation

This setup should look very familiar

$$
\begin{aligned}
\max_{x_1,x_2} P(x_1, x_2 | \neg e_1, e_2) &= \alpha \max_{x_0} \max_{x_1} P(x_0)P(x_1|x_0)P(\neg e_1|x_1)P(x_2|x_1)P(e_2|x_2) \\
&= \alpha \max_{x_1} \max_{x_0} P(e_2|x_2)P(x_2|x_1)P(\neg e_1|x_1)P(x_1|x_0)P(x_0) \\
&= \alpha P(e_2|x_2) \max_{x_1} P(x_2|x_1)P(\neg e_1|x_1) \max_{x_0} P(x_1|x_0)P(x_0)
\end{aligned}
$$

Most Likely Explanation

This setup should look very familiar 

$$\begin{aligned}\max_{x_1,x_2} P(x_1, x_2|\neg e_1, e_2) &= \alpha \max_{x_0} \max_{x_1} P(x_0)P(x_1|x_0)P(\neg e_1|x_1)P(x_2|x_1)P(e_2|x_2) \\ &= \alpha \max_{x_1} \max_{x_0} P(e_2|x_2)P(x_2|x_1)P(\neg e_1|x_1)P(x_1|x_0)P(x_0) \\ &= \alpha P(e_2|x_2) \max_{x_1} P(x_2|x_1)P(\neg e_1|x_1) \max_{x_0} P(x_1|x_0)P(x_0)\end{aligned}$$

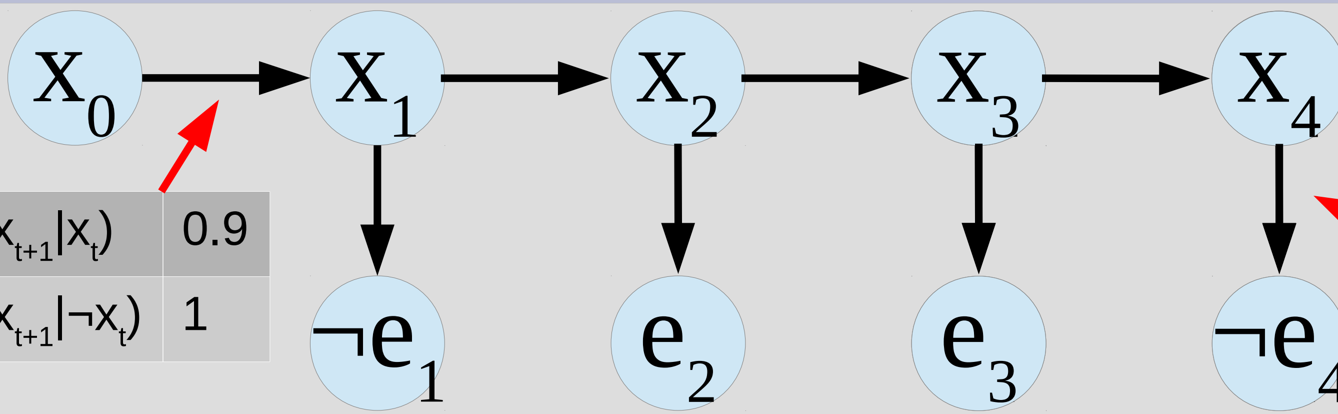
It's just filtering with max instead of sum!

So we can re-use our forward message trick, only slightly modified

Side note: max functions a lot like sum(linear)

Most Likely Explanation

$P(x_0)$	0.5



$P(x_{t+1}\|x_t)$	0.9
$P(x_{t+1}\|\neg x_t)$	1

$P(e_t\|x_t)$	0.01
$P(e_t\|\neg x_t)$	1

First find the best explanation for $x_0 \rightarrow x_1$

$$x_1 : P(\neg e_1|x_1) \max_{x_0} P(x_0)P(x_1|x_0)$$

$$0.99 \max(\underbrace{0.5 \cdot 0.9}_{x_0}, \underbrace{0.5 \cdot 1}_{\neg x_0})$$

$$0.99(0.5) = 0.495$$

$$\neg x_1 : P(\neg e_1|\neg x_1) \max_{x_0} P(x_0)P(\neg x_1|x_0)$$

$$0 \max(\underbrace{0.5 \cdot 0.1}_{x_0}, \underbrace{0.5 \cdot 0}_{\neg x_0})$$

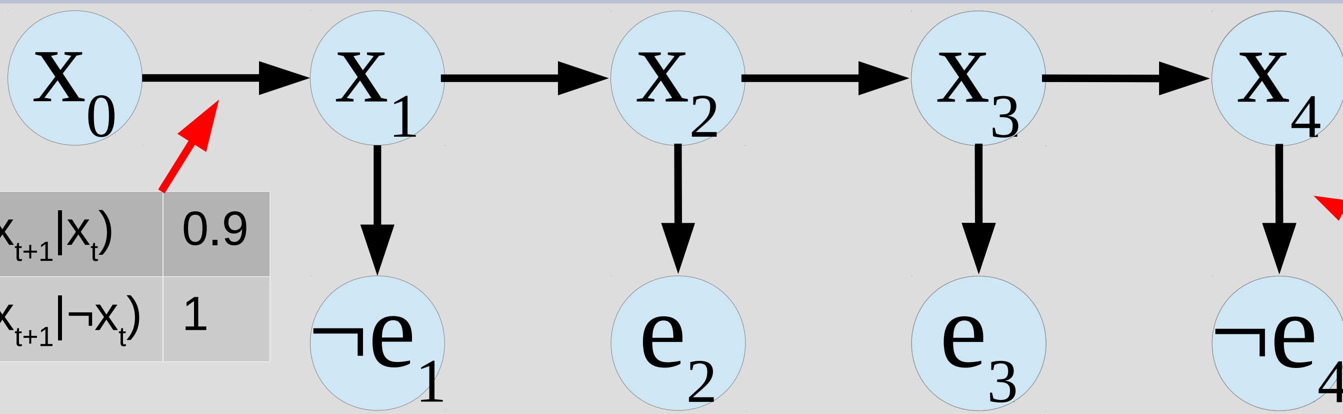
$$0(0.05) = 0$$

sorta...

So, best way to x_1 is pos $\neg x_0$, way to $\neg x_1$ is x_0

Most Likely Explanation

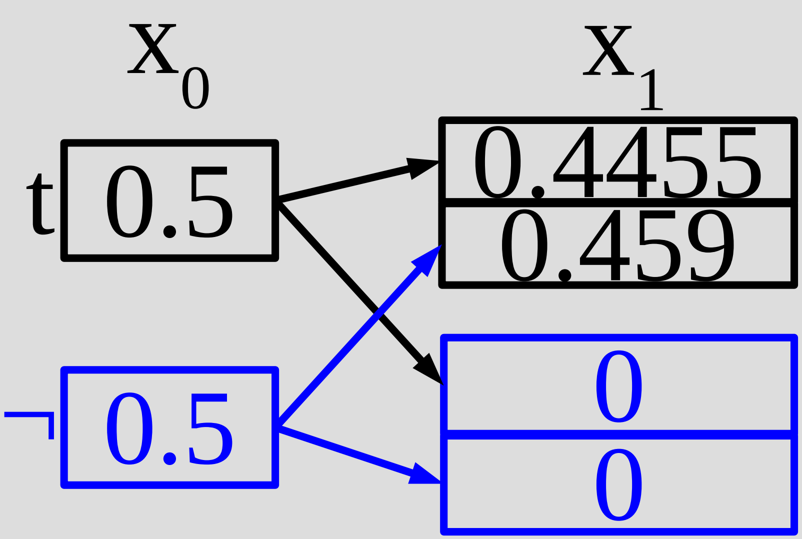
$P(x_0)$	0.5



$P(x_{t+1}\|x_t)$	0.9
$P(x_{t+1}\|\neg x_t)$	1

$P(e_t\|x_t)$	0.01
$P(e_t\|\neg x_t)$	1

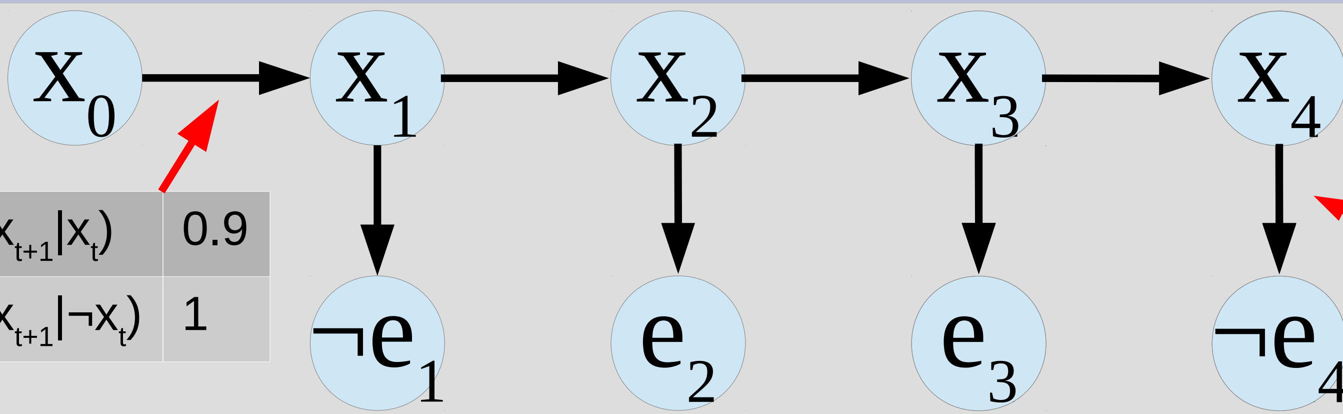
I will actually represent this more graphically:



Then we mark best way to x_1 (in red on next slide)

Most Likely Explanation

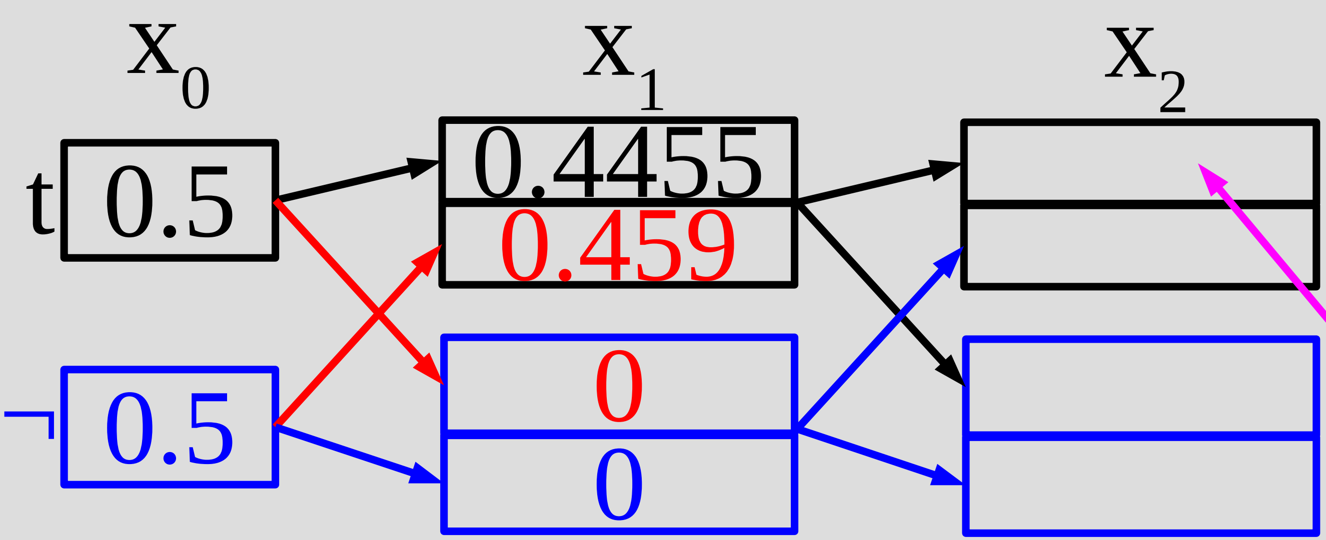
$P(x_0)$	0.5



$P(x_{t+1} \mid x_t)$	0.9
$P(x_{t+1} \mid \neg x_t)$	1

$P(e_t \mid x_t)$	0.01
$P(e_t \mid \neg x_t)$	1

I will actually represent this more graphically:

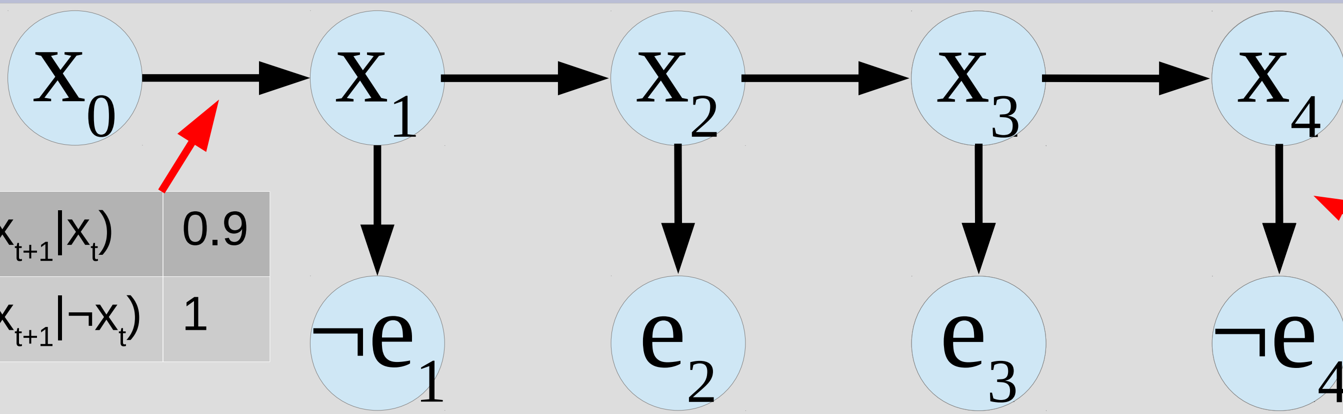


$x_1 \to x_2$:

$$P(e_2|x_2)P(x_2|x_1)P(x_1|e_1)$$
$$= 0.01 \cdot 0.9 \cdot 0.459$$
$$= 0.004131$$

Most Likely Explanation

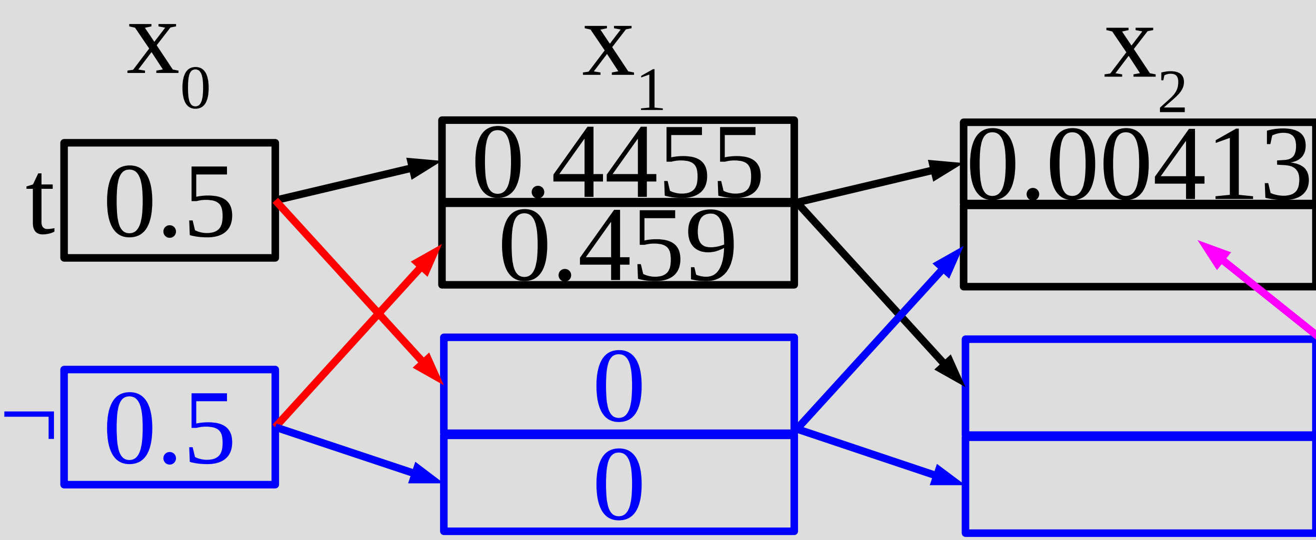
$P(x_0)$	0.5



$P(x_{t+1}\vert x_t)$	0.9
$P(x_{t+1}\vert \neg x_t)$	1

$P(e_t\vert x_t)$	0.01
$P(e_t\vert \neg x_t)$	1

I will actually represent this more graphically:



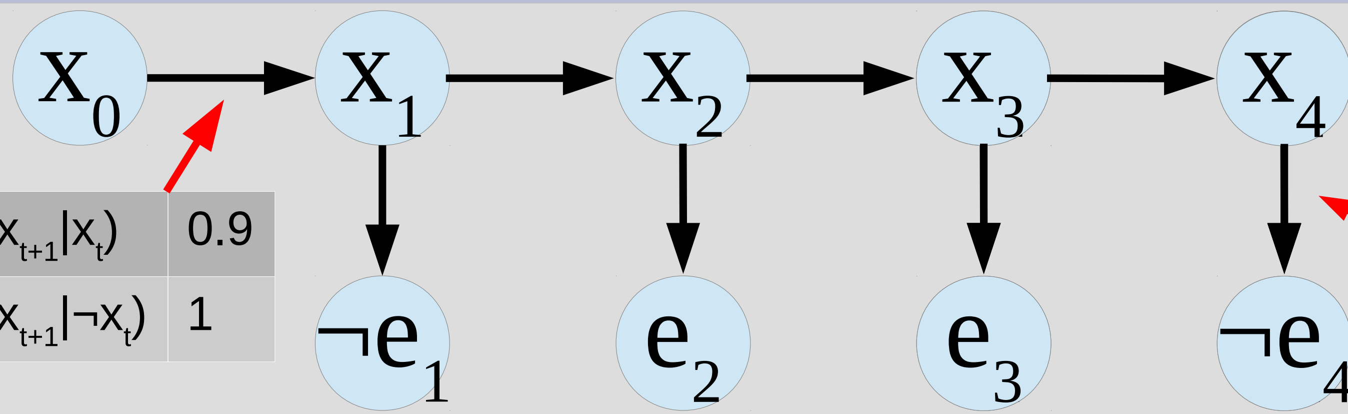
$\neg x_1 \rightarrow x_2$:

$$P(e_2|x_2)P(x_2|\neg x_1)P(\neg x_1|e_1)$$
$$= 0.01 \cdot 1 \cdot 0$$
$$= 0$$

Most Likely Explanation

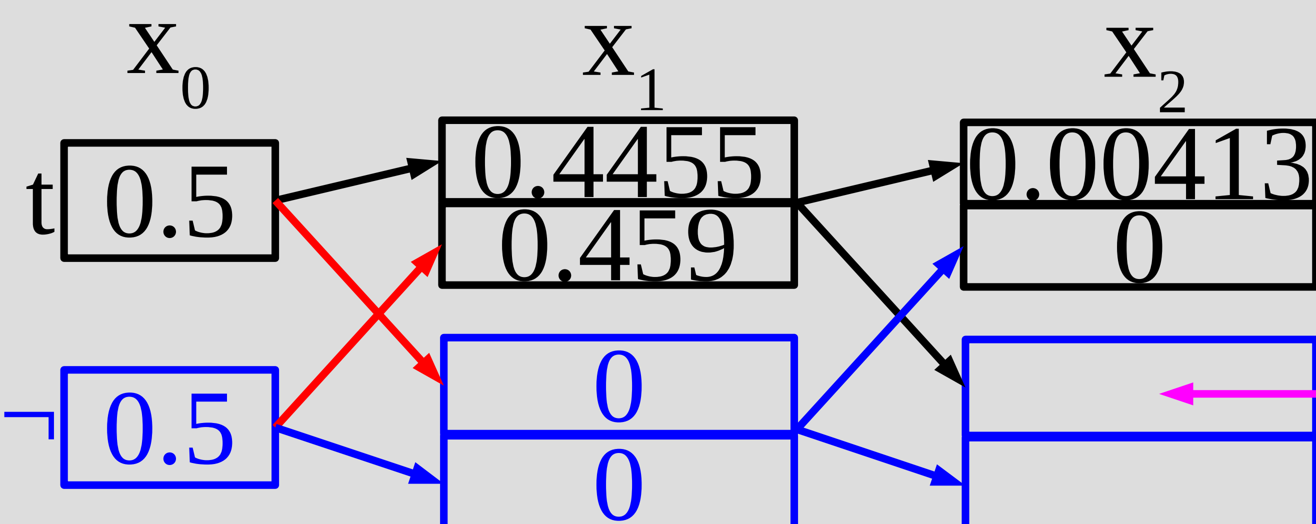
$P(x_0)$	0.5

$P(x_{t+1}\|x_t)$	0.9
$P(x_{t+1}\|\neg x_t)$	1



$P(e_t\|x_t)$	0.01
$P(e_t\|\neg x_t)$	1

I will actually represent this more graphically:

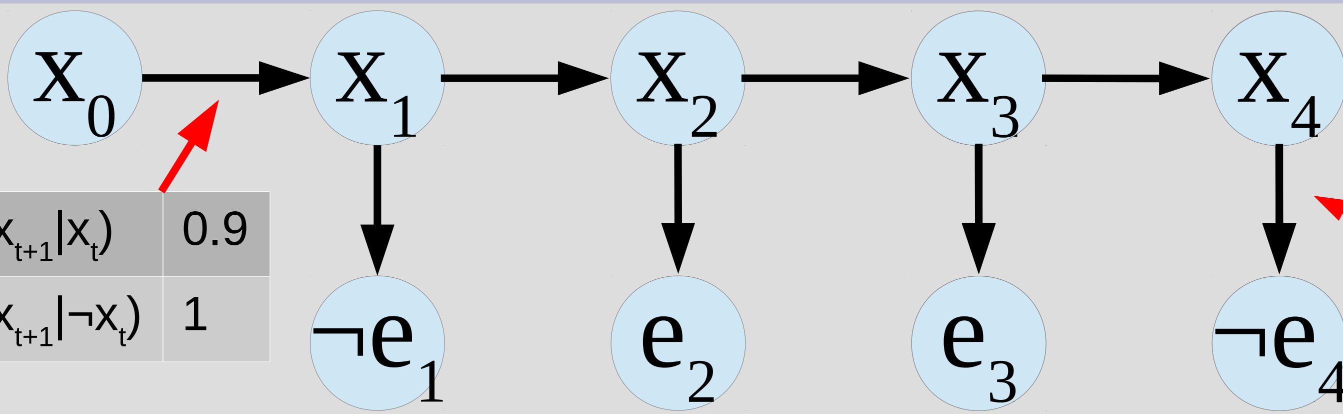


$x_1 \rightarrow \neg x_2$:

$$P(e_2|\neg x_2)P(\neg x_2|x_1)P(x_1|e_1)$$
$$= 1 \cdot 0.01 \cdot 0.459$$
$$= 0.00459$$

Most Likely Explanation

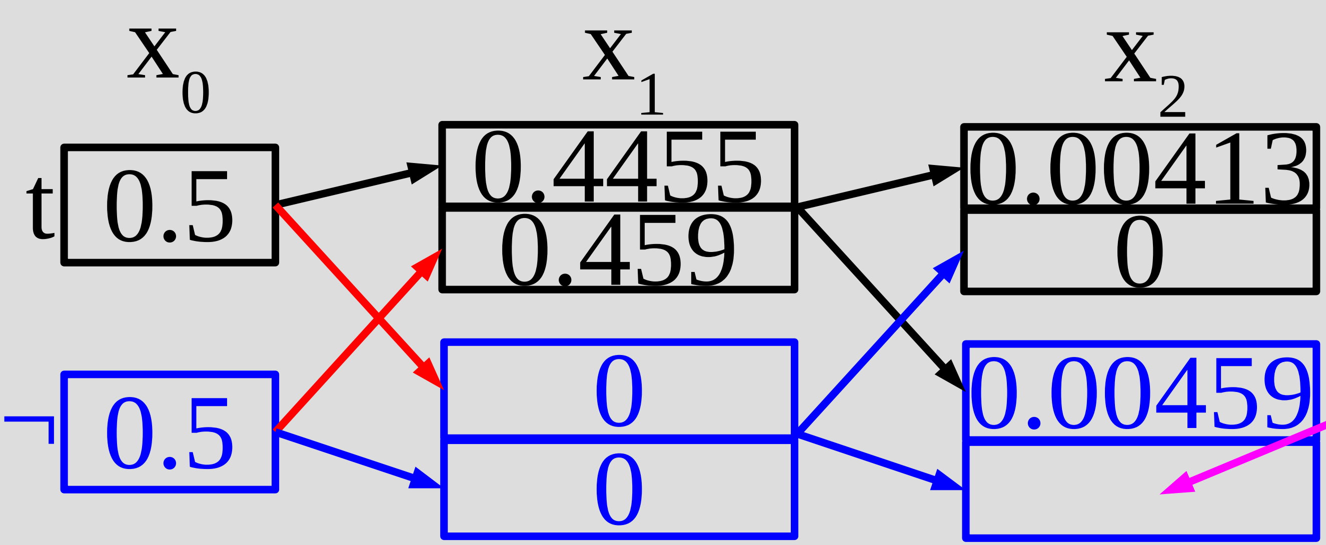
$P(x_0)$	0.5



$P(x_{t+1} \mid x_t)$	0.9
$P(x_{t+1} \mid \neg x_t)$	1

$P(e_t \mid x_t)$	0.01
$P(e_t \mid \neg x_t)$	1

I will actually represent this more graphically:



$\neg x_1 \rightarrow \neg x_2$:

$$P(e_2 \mid \neg x_2) P(\neg x_2 \mid \neg x_1) P(\neg x_1 \mid e_1)$$

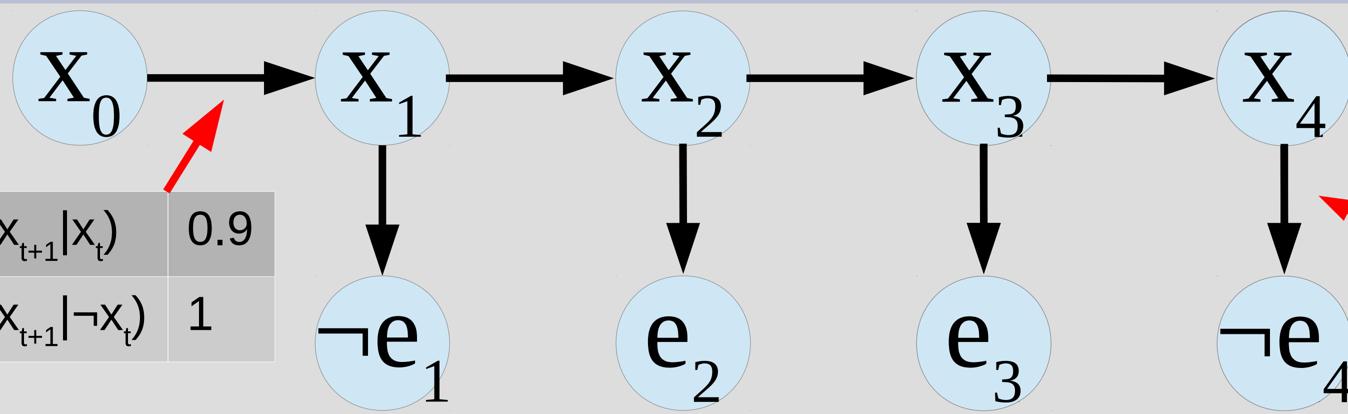
$$= 1 \cdot 0 \cdot 0$$

$$= 0$$

Most Likely Explanation

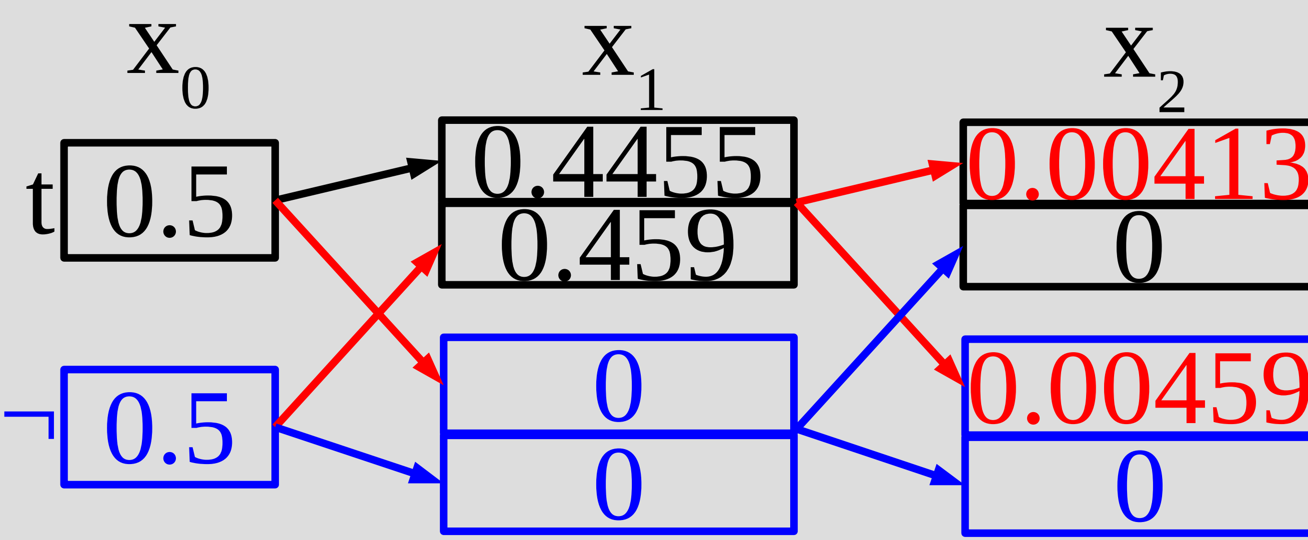
$P(x_0)$	0.5

$P(x_{t+1} \mid x_t)$	0.9
$P(x_{t+1} \mid \neg x_t)$	1



$P(e_t \mid x_t)$	0.01
$P(e_t \mid \neg x_t)$	1

I will actually represent this more graphically:

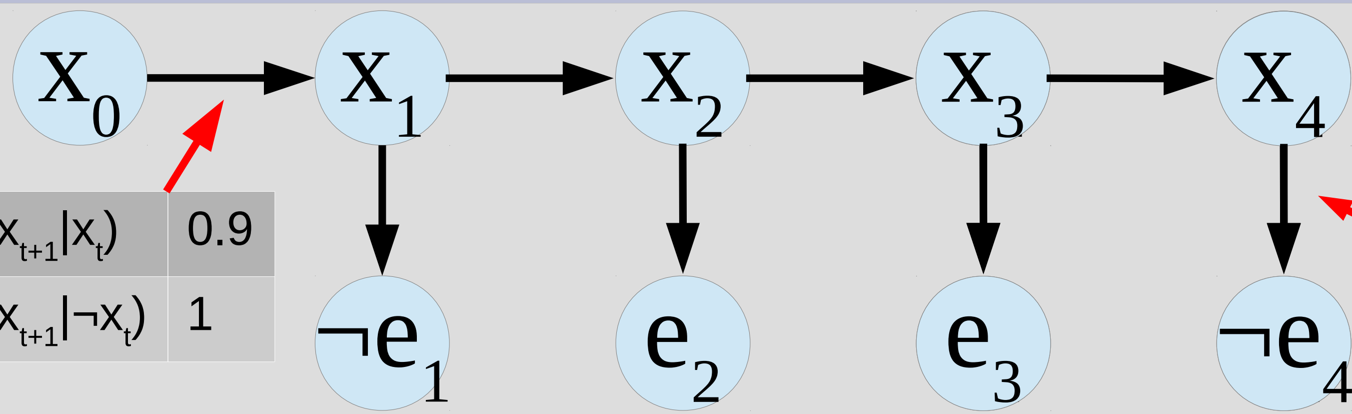


Then we find the best options (max shown red) for x_2

Most Likely Explanation

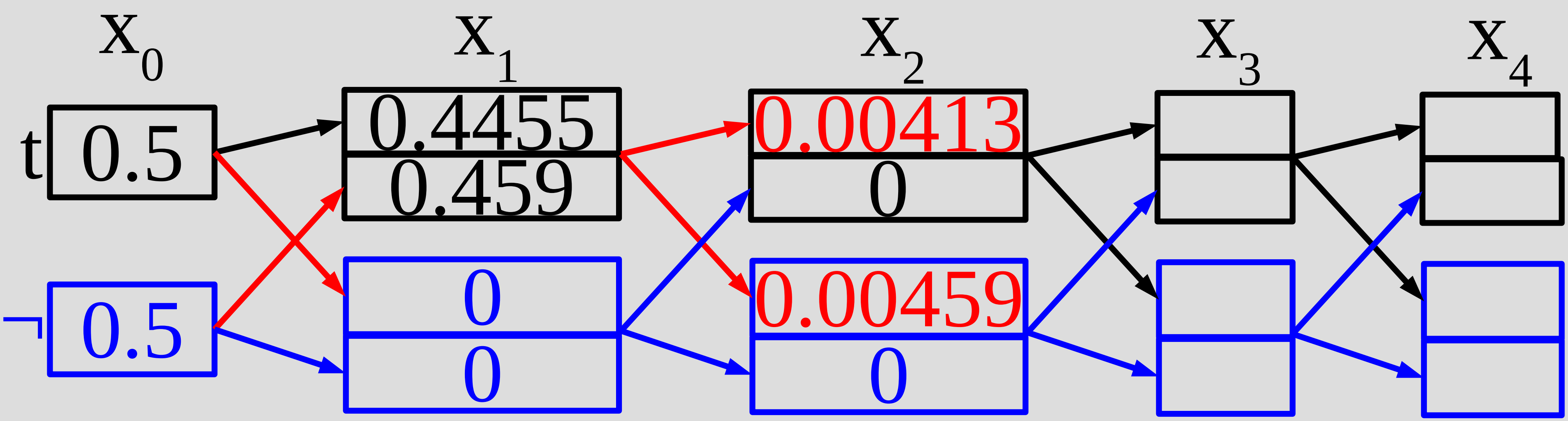
$P(x_0)$	0.5

$P(x_{t+1}\|x_t)$	0.9
$P(x_{t+1}\|\neg x_t)$	1



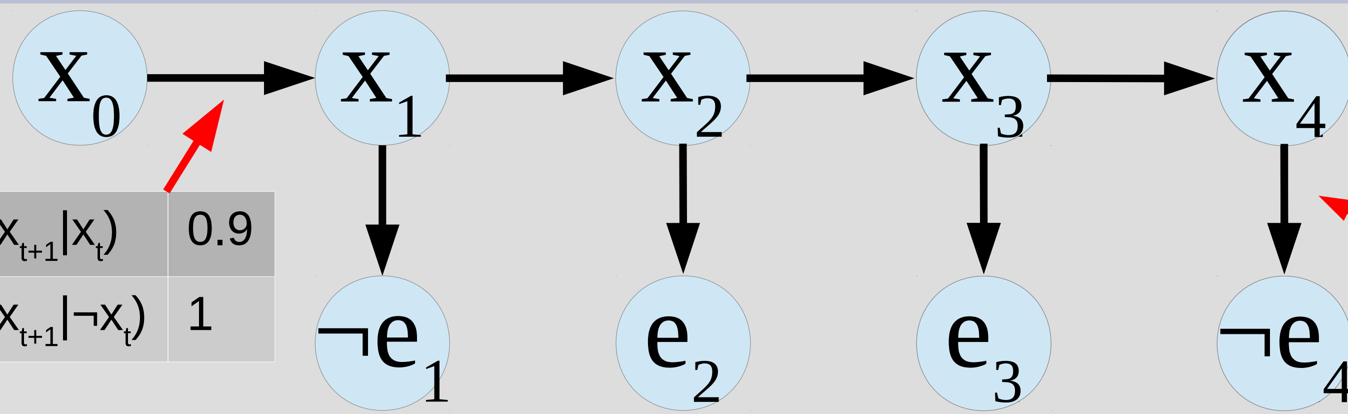
$P(e_t\|x_t)$	0.01
$P(e_t\|\neg x_t)$	1

Finish this example:



Most Likely Explanation

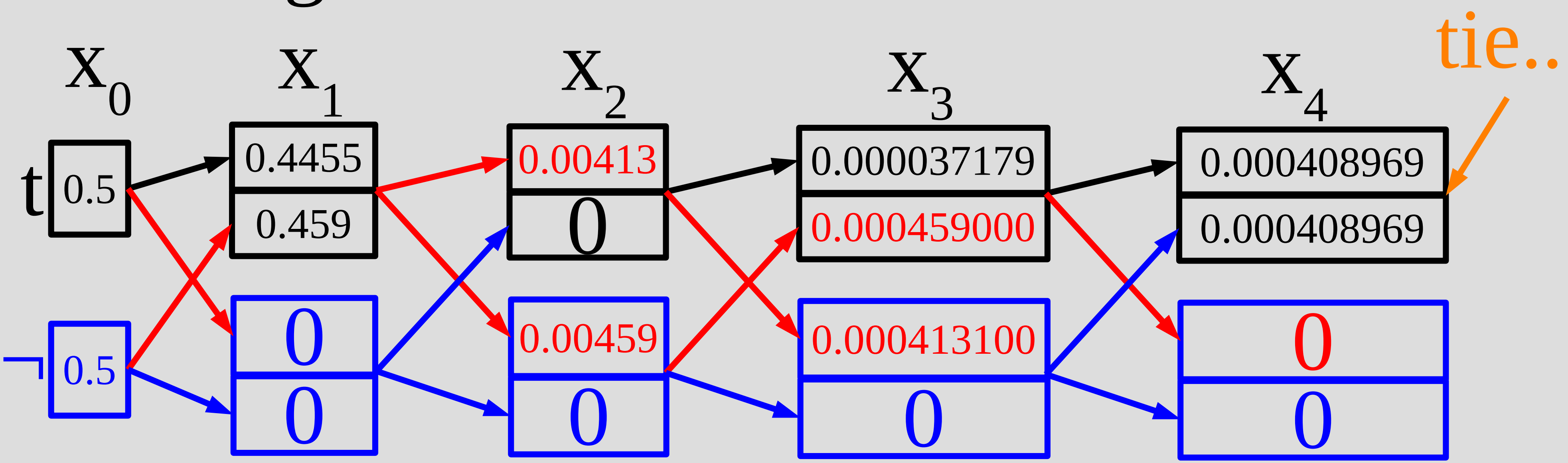
$P(x_0)$	0.5



$P(x_{t+1} \mid x_t)$	0.9
$P(x_{t+1} \mid \neg x_t)$	1

$P(e_t \mid x_t)$	0.01
$P(e_t \mid \neg x_t)$	1

Should get:

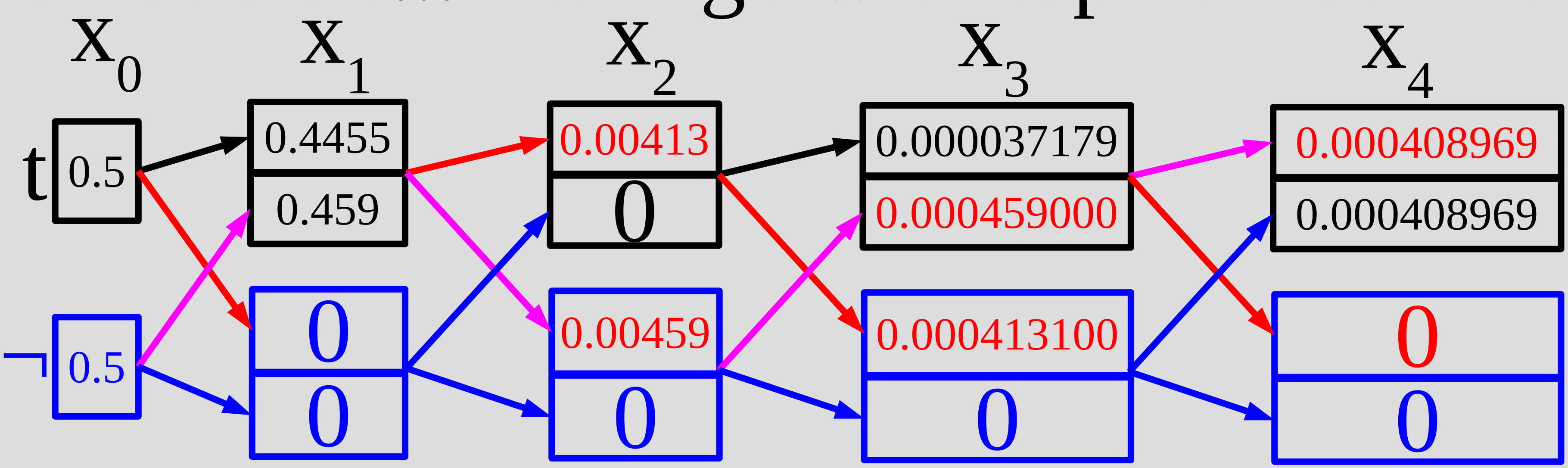


Most Likely Explanation

From here, you just find whether x_4 or $\neg x_4$ has a larger number (here it is x_4 in black)

trace in pink

Then trace the path back (two options here since a tie... I will go with top number max)



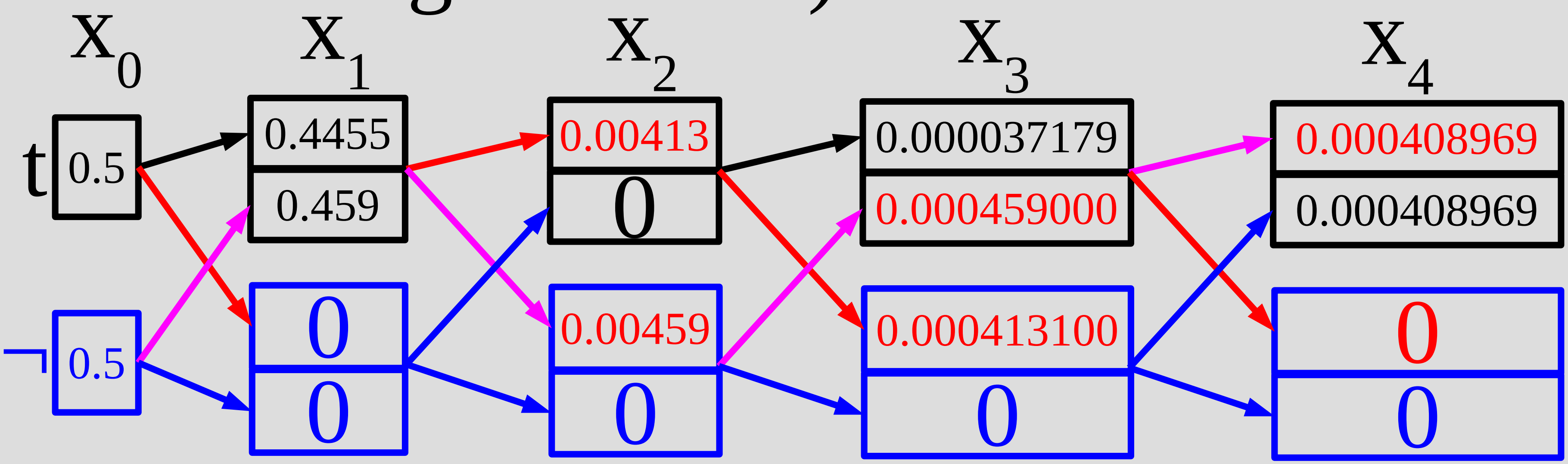
Most Likely Explanation

So the most likely sequence is:

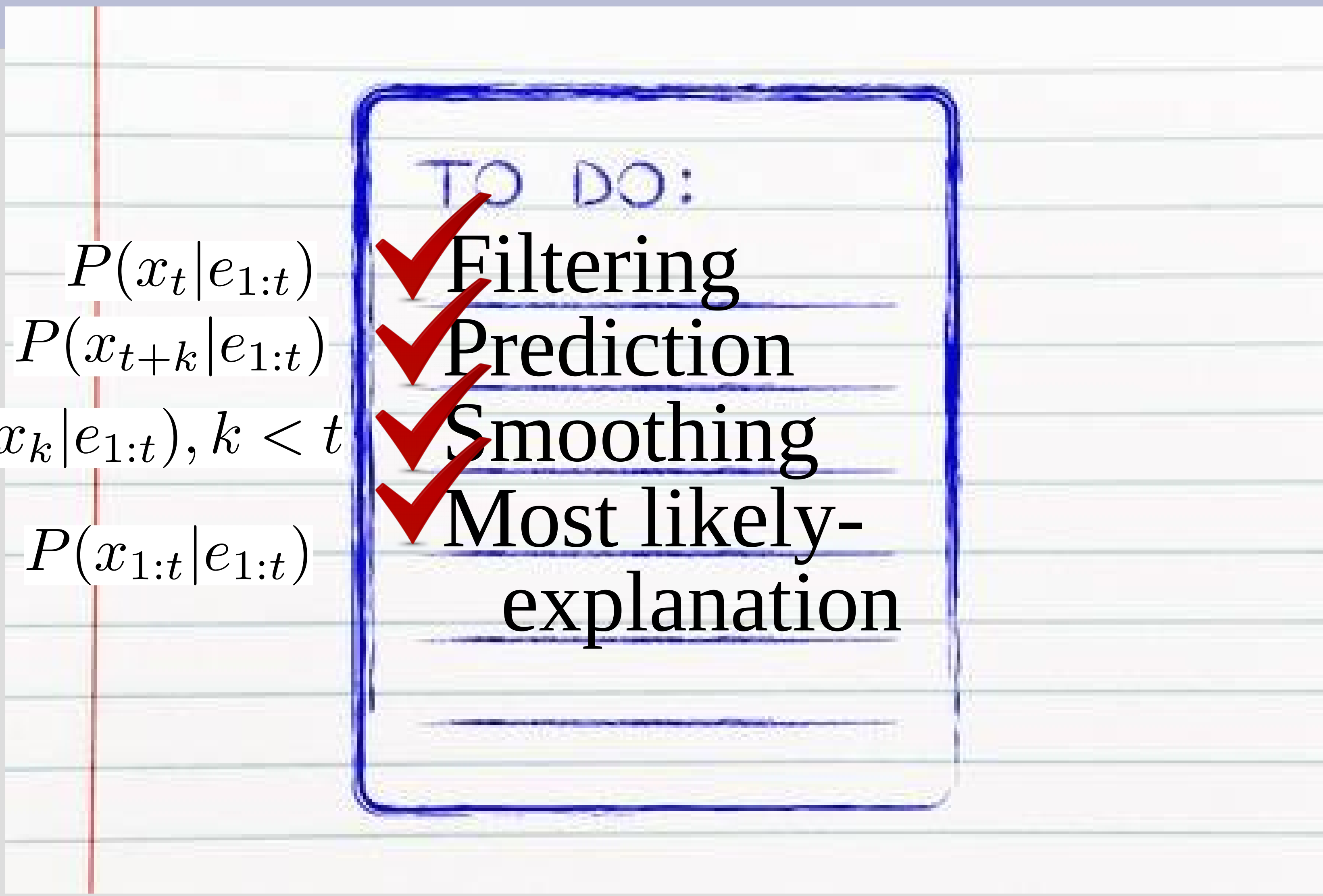
$[\neg x_0, x_1, \neg x_2, x_3, x_4]$

(tied with the sequence: $[\neg x_0, x_1, x_2, \neg x_3, x_4]$)

(Side note: this algorithm is called the "Viterbi algorithm"...)



Most Likely Explanation



- $P(x_t|e_{1:t})$ ✓ Filtering
- $P(x_{t+k}|e_{1:t})$ ✓ Prediction
- $P(x_k|e_{1:t}), k < t$ ✓ Smoothing
- $P(x_{1:t}|e_{1:t})$ ✓ Most likely-explanation

Done and done!