Introduction to Artificial Intelligence

English practicals 8: Probabilistic reasoning over time

Marika Ivanová

Department of Theoretical Computer Science and Mathematical Logic (KTIML)
Faculty of Mathematics and Physics

April 7th 2022

A small remainder from last practicals

Example: Covid-19 has the following symptoms: caugh (c), high temperature (t), loss of smell (l) and fatigue (f). Knowing the probabilities $P(symptom|Covid = true)$ calculate the probability that a patient exhibiting a combination of symptoms has Covid.

A small remainder from last practicals

Example: Covid-19 has the following symptoms: caugh (c), high temperature (t), loss of smell (l) and fatigue (f). Knowing the probabilities $P(symptom|Covid = true)$ calculate the probability that a patient exhibiting a combination of symptoms has Covid.

Chain rule (from lecture):

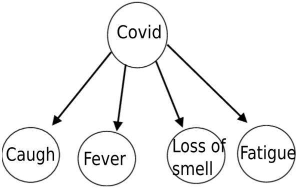
$$P(f, l, t, c, covid) = P(f|l, t, c, covid)P(l, t, c, ccovid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t, c, covid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t|c, covid)P(c, covid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t|c, covid)P(c|covid)P(covid)$$

A small remainder from last practicals

Example: Covid-19 has the following symptoms: caugh (c), high temperature (t), loss of smell (l) and fatigue (f). Knowing the probabilities $P(symptom|Covid = true)$ calculate the probability that a patient exhibiting a combination of symptoms has Covid.

Chain rule (from lecture):

$$P(f, l, t, c, covid) = P(f|l, t, c, covid)P(l, t, c, ccovid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t, c, covid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t|c, covid)P(c, covid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t|c, covid)P(c|covid)P(covid)$$

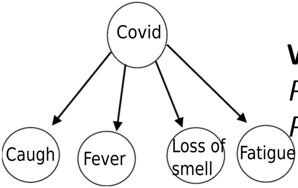


A small remainder from last practicals

Example: Covid-19 has the following symptoms: caugh (c), high temperature (t), loss of smell (l) and fatigue (f). Knowing the probabilities $P(symptom|Covid = true)$ calculate the probability that a patient exhibiting a combination of symptoms has Covid.

Chain rule (from lecture):

$$P(f, l, t, c, covid) = P(f|l, t, c, covid)P(l, t, c, ccovid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t, c, covid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t|c, covid)P(c, covid) =$$
$$= P(f|l, t, c, covid)P(l|t, c, covid)P(t|c, covid)P(c|covid)P(covid)$$



With BN we can use conditional independence:

$$P(c, t, l, f, covid) =$$
$$P(f|covid)P(l|covid)P(t|covid)P(c|covid)P(covid)$$

A small reminder from the lecture

- World viewed as a series of time slices

A small reminder from the lecture

- World viewed as a series of time slices
- **Hidden** (not observable) random variables X_t - describe the state at time t
- **Observable** random variables E_t - what we observe about the state at time t (e.g., sensors)

A small reminder from the lecture

- World viewed as a series of time slices
- **Hidden** (not observable) random variables X_t - describe the state at time t
- **Observable** random variables E_t - what we observe about the state at time t (e.g., sensors)

- **Transition model:** $P(X_t|X_{0:t-1})$
  - **Markov assumption:** $P(X_t|X_{0:t-1}) = P(X_t|X_{t-1})$
  - **Stationary process:** All transition tables $P(X_t|X_{t-1})$ are identical for each t

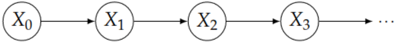
A small reminder from the lecture

- World viewed as a series of time slices
- **Hidden** (not observable) random variables X_t - describe the state at time t
- **Observable** random variables E_t - what we observe about the state at time t (e.g., sensors)

- **Transition model:** $P(X_t | X_{0:t-1})$
  - **Markov assumption:** $P(X_t | X_{0:t-1}) = P(X_t | X_{t-1})$
  - **Stationary process:** All transition tables $P(X_t | X_{t-1})$ are identical for each t
- **Observation model:** $P(E_t | X_{0:t}, E_{1:t-1})$
  - **sensor Markov assumption:** $P(E_t | X_{0:t}, E_{1:t-1}) = P(E_t | X_t)$

Exercise 1: Markov chain

Assume a stationary first order Markov chain:

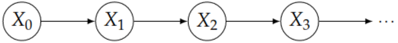


- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t|X_{t-1})$

Exercise 1: Markov chain

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:

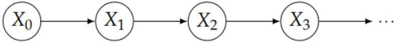


- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t \mid X_{t-1})$

Exercise 1: Markov chain

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



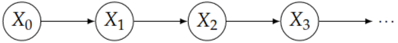
- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t \mid X_{t-1})$

What is the weather probability distribution at day 1, i.e., $P(X_1)$ given $P(X_0 = s) = 1$?

Exercise 1: Markov chain

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t|X_{t-1})$

What is the weather probability distribution at day 1, i.e., $P(X_1)$ given $P(X_0 = s) = 1$?

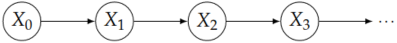
$$P(X_1 = s) = P(X_1 = s|X_0 = r)P(X_0 = r) + P(X_1 = s|X_0 = s)P(X_0 = s) =$$

$$= 0.3 * 0 + 0.9 * 1 = 0.9$$

Exercise 1: Markov chain

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t \mid X_{t-1})$

What is the weather probability distribution at day 1, i.e., $P(X_1)$ given $P(X_0 = s) = 1$?

$$P(X_1 = s) = P(X_1 = s \mid X_0 = r)P(X_0 = r) + P(X_1 = s \mid X_0 = s)P(X_0 = s) =$$

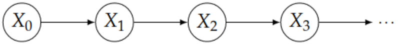
$$= 0.3 * 0 + 0.9 * 1 = 0.9$$

How about two steps, i.e., $P(X_2)$ given $P(X_1)$ from the previous step?

Exercise 1: Markov chain

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t \mid X_{t-1})$

What is the weather probability distribution at day 1, i.e., $P(X_1)$ given $P(X_0 = s) = 1$?

$$P(X_1 = s) = P(X_1 = s \mid X_0 = r)P(X_0 = r) + P(X_1 = s \mid X_0 = s)P(X_0 = s) =$$

$$= 0.3 * 0 + 0.9 * 1 = 0.9$$

How about two steps, i.e., $P(X_2)$ given $P(X_1)$ from the previous step?

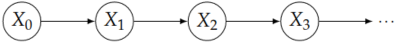
$$P(X_2 = s) = P(X_2 = s \mid X_1 = r)P(X_1 = r) + P(X_2 = s \mid X_1 = s)P(X_1 = s) =$$

$$= 0.3 * 0.1 + 0.9 * 0.9 = 0.84$$

Exercise 1: Markov chain contd.

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



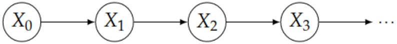
- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t \mid X_{t-1})$

What is the weather probability distribution at day "infinity", i.e., $P(X_\infty)$

Exercise 1: Markov chain contd.

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t \mid X_{t-1})$

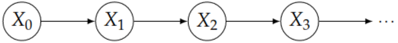
What is the weather probability distribution at day "infinity", i.e., $P(X_\infty)$

$$P(X_\infty = s) = P(X_\infty = s \mid X_{\infty-1} = r)P(X_{\infty-1} = r) + P(X_\infty = s \mid X_{\infty-1} = s)P(X_{\infty-1} = s)$$

Exercise 1: Markov chain contd.

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t|X_{t-1})$

What is the weather probability distribution at day "infinity", i.e., $P(X_\infty)$

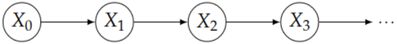
$$P(X_\infty = s) = P(X_\infty = s|X_{\infty-1} = r)P(X_{\infty-1} = r) + P(X_\infty = s|X_{\infty-1} = s)P(X_{\infty-1} = s)$$

$$= 0.3 * P(X_{\infty-1} = r) + 0.9 * P(X_{\infty-1} = s)$$

Exercise 1: Markov chain contd.

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t | X_{t-1})$

What is the weather probability distribution at day "infinity", i.e., $P(X_\infty)$

$$P(X_\infty = s) = P(X_\infty = s | X_{\infty-1} = r) P(X_{\infty-1} = r) + P(X_\infty = s | X_{\infty-1} = s) P(X_{\infty-1} = s)$$

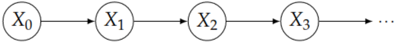
$$= 0.3 * P(X_{\infty-1} = r) + 0.9 * P(X_{\infty-1} = s)$$

$$\Rightarrow P(X_\infty = s) = 3 * P(X_{\infty-1} = r)$$

Exercise 1: Markov chain contd.

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
r	r	0.7
r	s	0.3
s	r	0.1
s	s	0.9

Assume a stationary first order Markov chain:



- States $X = \{(r)ain, (s)un\}$
- Initial state: sun
- Transition model $P(X_t|X_{t-1})$

What is the weather probability distribution at day "infinity", i.e., $P(X_\infty)$

$$P(X_\infty = s) = P(X_\infty = s|X_{\infty-1} = r)P(X_{\infty-1} = r) + P(X_\infty = s|X_{\infty-1} = s)P(X_{\infty-1} = s)$$

$$= 0.3 * P(X_{\infty-1} = r) + 0.9 * P(X_{\infty-1} = s)$$

$$\Rightarrow P(X_\infty = s) = 3 * P(X_{\infty-1} = r)$$

$$\Rightarrow P(X_\infty) = (0.75, 0.25)$$

Basic inference tasks

- **Filtering** $P(X_t|e_{1:t})$ (Where am I now?)
- **Prediction** $P(X_{t+k}|e_{1:t}), k > 0$ (Where will I be in future?)
- **Smoothing** $P(X_k|e_{1:t}), k < t$ (Where was I in the past?)
- **Most likely explanation** $\arg\max_{x_{1:t}} P(X_{1:t}|e_{1:t})$ (What path did I go through?)

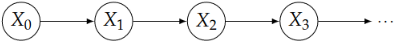
Basic inference tasks

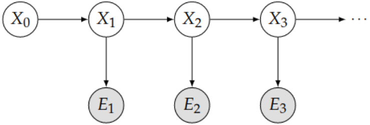
- **Filtering** $P(X_t|e_{1:t})$ (Where am I now?)
- **Prediction** $P(X_{t+k}|e_{1:t}), k > 0$ (Where will I be in future?)
- **Smoothing** $P(X_k|e_{1:t}), k < t$ (Where was I in the past?)
- **Most likely explanation** $\arg\max_{x_{1:t}} P(X_{1:t}|e_{1:t})$ (What path did I go through?)

Poll

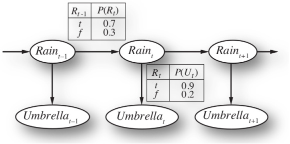
Markov chain (MC) vs Hidden Markov Model (HMM)

- MC: all variables are observable
- HMM: some variables are observable, some are hidden



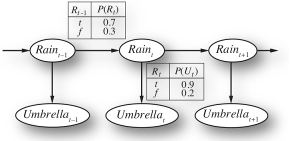


Exercise 2



Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

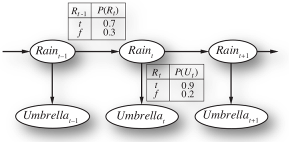
Exercise 2



Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

Exercise 2

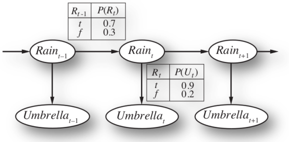


Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$

Exercise 2

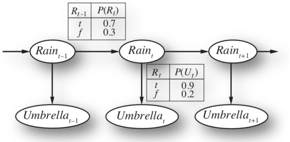


Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$
- Day 1: observation $u_1 = true$

Exercise 2

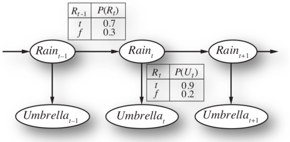


Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t) P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$
- Day 1: observation $u_1 = true$
  - Prediction:
    $P(r_1 = true) = \sum_{r_0} P(r_1 = true|r_0) * P(r_0) = 0.7 * 0.5 + 0.3 * 0.5 = 0.5$

Exercise 2

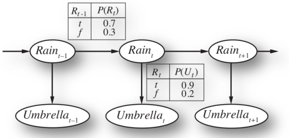


Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$
- Day 1: observation $u_1 = true$
  - Prediction: $P(r_1 = true) = \sum_{r_0} P(r_1 = true|r_0) * P(r_0) = 0.7 * 0.5 + 0.3 * 0.5 = 0.5$
  - Update by the new evidence: $P(r_1 = true|u_1 = true) = \alpha P(u_1 = true|r_1 = true)P(r_1 = true) = \alpha * 0.9 * 0.5 = \alpha * 0.45 \approx 0.818$

Exercise 2

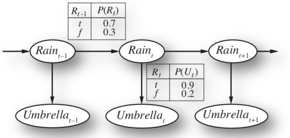


Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$
- Day 1: observation $u_1 = true$
  - Prediction:
    $P(r_1 = true) = \sum_{r_0} P(r_1 = true|r_0) * P(r_0) = 0.7 * 0.5 + 0.3 * 0.5 = 0.5$
  - Update by the new evidence: $P(r_1 = true|u_1 = true) = \alpha P(u_1 = true|r_1 = true)P(r_1 = true) = \alpha * 0.9 * 0.5 = \alpha * 0.45 \approx 0.818$
- Day 2: observation $u_2 = true$

Exercise 2

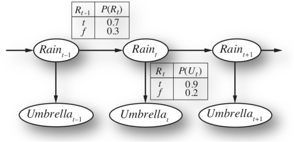


Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$
- Day 1: observation $u_1 = true$
  - Prediction: $P(r_1 = true) = \sum_{r_0} P(r_1 = true|r_0) * P(r_0) = 0.7 * 0.5 + 0.3 * 0.5 = 0.5$
  - Update by the new evidence: $P(r_1 = true|u_1 = true) = \alpha P(u_1 = true|r_1 = true)P(r_1 = true) = \alpha * 0.9 * 0.5 = \alpha * 0.45 \approx 0.818$
- Day 2: observation $u_2 = true$
  - Prediction: $P(r_2 = true|u_1 = true) = \sum_{r_1} P(r_2 = true|r_1) * P(r_1|u_1 = true) = 0.7 * 0.818 + 0.3 * 0.182 \approx 0.627$

Exercise 2



Assume that the probability of rain at day 0 is 0.5. What is the probability of rain at day 2, given that we observed an umbrella at day 1 and 2?

Filtering: Where am I now? $P(X_{t+1}|E_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$

- Day 0: no observation, only the prior belief $P(R_0) = (0.5, 0.5)$
- Day 1: observation $u_1 = true$
  - Prediction:
    $P(r_1 = true) = \sum_{r_0} P(r_1 = true|r_0) * P(r_0) = 0.7 * 0.5 + 0.3 * 0.5 = 0.5$
  - Update by the new evidence: $P(r_1 = true|u_1 = true) = \alpha P(u_1 = true|r_1 = true)P(r_1 = true) = \alpha * 0.9 * 0.5 = \alpha * 0.45 \approx 0.818$
- Day 2: observation $u_2 = true$
  - Prediction: $P(r_2 = true|u_1 = true) = \sum_{r_1} P(r_2 = true|r_1) * P(r_1|u_1 = true) = 0.7 * 0.818 + 0.3 * 0.182 \approx 0.627$
  - Update by the new evidence:
    $P(r_2 = true|u_1 = u_2 = true) = \alpha P(u_2 = true|r_2 = true)P(r_2 = true|u_1 = true) = \alpha * 0.9 * 0.627 = \alpha * 0.564 \approx 0.883$

Exercise 3: Hidden Markov model

Professor has a 3 shirts: red (R), green (G) and blue (B). The color of shirt he wears gives a hint about his mood (happy and sad) on that day. What is the most likely sequence of her mood in 3 days, if he wars G, B, R? Assume that on day 1 there is a probability 0.4 that she is sad.

Exercise 3: Hidden Markov model

Professor has a 3 shirts: red (R), green (G) and blue (B). The color of shirt he wears gives a hint about his mood (happy and sad) on that day. What is the most likely sequence of her mood in 3 days, if he wars G, B, R? Assume that on day 1 there is a probability 0.4 that she is sad.

	H	**S**
H	0.7	0.3
S	0.5	0.5

	R	**G**	**B**
H	0.8	0.1	0.1
S	0.2	0.3	0.5

Exercise 3: Hidden Markov model

Professor has a 3 shirts: red (R), green (G) and blue (B). The color of shirt he wears gives a hint about his mood (happy and sad) on that day. What is the most likely sequence of her mood in 3 days, if he wars G, B, R? Assume that on day 1 there is a probability 0.4 that she is sad.

$$\max_{m_1, m_2, m_3} P(C_1 = G, C_2 = B, C_3 = R, M_1 = m_1, M_2 = m_2, M_3 = m_3)$$

	H	**S**
H	0.7	0.3
S	0.5	0.5

	R	**G**	**B**
H	0.8	0.1	0.1
S	0.2	0.3	0.5

Exercise 3: Hidden Markov model

Professor has a 3 shirts: red (R), green (G) and blue (B). The color of shirt he wears gives a hint about his mood (happy and sad) on that day. What is the most likely sequence of her mood in 3 days, if he wars G, B, R? Assume that on day 1 there is a probability 0.4 that she is sad.

	H	S
H	0.7	0.3
S	0.5	0.5

	R	G	B
H	0.8	0.1	0.1
S	0.2	0.3	0.5

$$\max_{m_1,m_2,m_3} P(C_1 = G, C_2 = B, C_3 = R, M_1 = m_1, M_2 = m_2, M_3 = m_3)$$

Chain rule:

- $P(C_3|C_2, C_1, M_3, M_2, M_1)\times$
- $P(C_2|C_1, M_3, M_2, M_1)\times$
- $P(C_1|M_3, M_2, M_1)\times$
- $P(M_3|M_2, M_1)\times$
- $P(M_2|M_1)\times$
- $P(M_1)\times$

Exercise 3: Hidden Markov model

Professor has a 3 shirts: red (R), green (G) and blue (B). The color of shirt he wears gives a hint about his mood (happy and sad) on that day. What is the most likely sequence of her mood in 3 days, if he wars G, B, R? Assume that on day 1 there is a probability 0.4 that she is sad.

	H	S
H	0.7	0.3
S	0.5	0.5

	R	G	B
H	0.8	0.1	0.1
S	0.2	0.3	0.5

$$\max_{m_1,m_2,m_3} P(C_1 = G, C_2 = B, C_3 = R, M_1 = m_1, M_2 = m_2, M_3 = m_3)$$

Chain rule:

- $P(C_3|C_2, C_1, M_3, M_2, M_1)\times$
- $P(C_2|C_1, M_3, M_2, M_1)\times$
- $P(C_1|M_3, M_2, M_1)\times$
- $P(M_3|M_2, M_1)\times$
- $P(M_2|M_1)\times$
- $P(M_1)\times$
- $=$
- $P(C_3|M_3)P(C_2|M_2)P(C_1|M_1)P(M_3|M_2)P(M_2|M_1)P(M_1|S)$

Exercise 3: Hidden Markov model

Professor has a 3 shirts: red (R), green (G) and blue (B). The color of shirt he wears gives a hint about his mood (happy and sad) on that day. What is the most likely sequence of her mood in 3 days, if he wars G, B, R? Assume that on day 1 there is a probability 0.4 that she is sad.

	H	S
H	0.7	0.3
S	0.5	0.5

	R	G	B
H	0.8	0.1	0.1
S	0.2	0.3	0.5

$$\max_{m_1, m_2, m_3} P(C_1 = G, C_2 = B, C_3 = R, M_1 = m_1, M_2 = m_2, M_3 = m_3)$$

Chain rule:

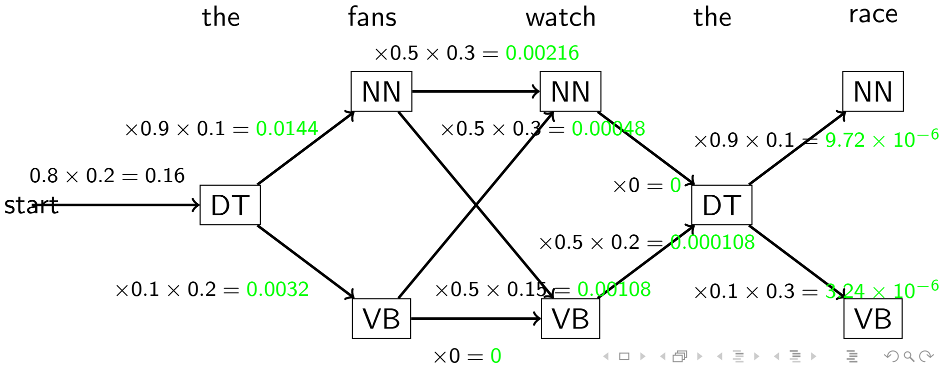
- $P(C_3|C_2, C_1, M_3, M_2, M_1)\times$
- $P(C_2|C_1, M_3, M_2, M_1)\times$
- $P(C_1|M_3, M_2, M_1)\times$
- $P(M_3|M_2, M_1)\times$
- $P(M_2|M_1)\times$
- $P(M_1)\times$
- $=$
- $P(C_3|M_3)P(C_2|M_2)P(C_1|M_1)P(M_3|M_2)P(M_2|M_1)P(M_1|S)$

- HHH: $0.1 * 0.1 * 0.8 * 0.7^3 = 0.0027$
- SSH: $0.3 * 0.5 * 0.8 * 0.5 * 0.5 * 0.6 = 0.018$
- ...

Viterbi algorithm: NLP example

	DT	NN	VB
(Start)	0.8	0.2	0
DT	0	0.9	0.1
NN	0	0.5	0.5
VB	0.5	0.5	0

	THE	FANS	WATCH	RACE
DT	0.2	0	0	0
NN	0	0.1	0.3	0.1
VB	0	0.2	0.15	0.3



Selected quiz questions

Selected quiz questions

- How is Markov assumption used in the transition model

Selected quiz questions

- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC

Selected quiz questions

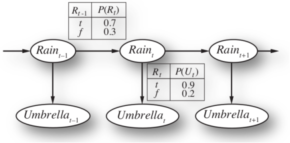
- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?

Selected quiz questions

- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?
  Static: state does not change. Stationary: probability distribution $P(X_t|X_{t-1})$ is eaqual for all time steps t.

Selected quiz questions

- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?
  Static: state does not change. Stationary: probability distribution $P(X_t|X_{t-1})$ is eaqual for all time steps t.
- What is the probability of the transition Rain=true → Rain= false?

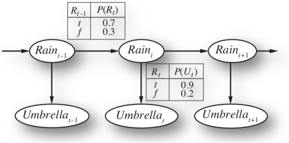


Selected quiz questions

- How is Markov assumption used in the transition model

  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?

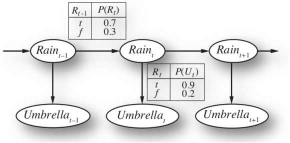
  Static: state does not change. Stationary: probability distribution $P(X_t|X_{t-1})$ is eaqual for all time steps t.
- What is the probability of the transition Rain=true → Rain= false?



0.3

Selected quiz questions

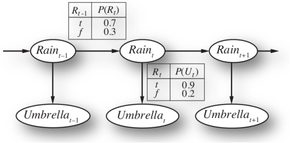
- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?
  Static: state does not change. Stationary: probability distribution $P(X_t|X_{t-1})$ is eaqual for all time steps t.
- What is the probability of the transition Rain=true $\rightarrow$ Rain= false?



  0.3
- Assume that you want to translate a speach that you hear. What type of inference task is it? (Poll)

Selected quiz questions

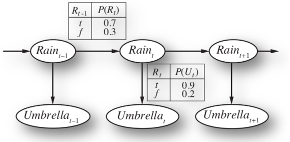
- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?
  Static: state does not change. Stationary: probability distribution $P(X_t|X_{t-1})$ is eaqual for all time steps t.
- What is the probability of the transition Rain=true → Rain= false?



  0.3
- Assume that you want to translate a speach that you hear. What type of inference task is it? (Poll)

Selected quiz questions

- How is Markov assumption used in the transition model
  $P(X_t|X_{0:t-1}) = P(X_t|X_{t-k})$, where $k > 0$ is the order of MC
- What is the difference between stationary and static process?
  Static: state does not change. Stationary: probability distribution $P(X_t|X_{t-1})$ is eaqual for all time steps t.
- What is the probability of the transition Rain=true → Rain= false?



  0.3
- Assume that you want to translate a speach that you hear. What type of inference task is it? (Poll)
  Most likely explanation.

- Does mixing time mean that after that time, the state will not change?

- Does mixing time mean that after that time, the state will not change?
  No.

- Does mixing time mean that after that time, the state will not change?

  No.

- Assume that state variable X depends on two previous state variables X_{t-1} and X_{t-2}. Can we encode this transition using dependence just between subsequent states?

- Does mixing time mean that after that time, the state will not change?

  No.

- Assume that state variable X depends on two previous state variables X_{t-1} and X_{t-2}. Can we encode this transition using dependence just between subsequent states?

  Yes.

  Example: in 2nd order MC with possible states$A$, G, C, T. transition table has 4x4 values. We construct 1st MC with 16 states $AA, AC, \ldots, CA, CC, \ldots, GG$. The new transition table has 16^2 entries.

- Does mixing time mean that after that time, the state will not change?

  No.

- Assume that state variable X depends on two previous state variables X_{t-1} and X_{t-2}. Can we encode this transition using dependence just between subsequent states?

  Yes.

  Example: in 2nd order MC with possible states$A$, G, C, T. transition table has 4x4 values. We construct 1st MC with 16 states $AA, AC, \dots, CA, CC, \dots, GG$. The new transition table has 16^2 entries.

- Assume that we do full smoothing, that is we smooth every past variable (think about an efficient method how to do it), and for each past variable we select the most probable value. Will we get the most likely explanation of a sequence of observations? Why?

- Does mixing time mean that after that time, the state will not change?

  No.

- Assume that state variable X depends on two previous state variables X_{t-1} and X_{t-2}. Can we encode this transition using dependence just between subsequent states?

  Yes.

  Example: in 2nd order MC with possible states$A$, G, C, T. transition table has 4x4 values. We construct 1st MC with 16 states $AA, AC, \ldots, CA, CC, \ldots, GG$. The new transition table has 16^2 entries.

- Assume that we do full smoothing, that is we smooth every past variable (think about an efficient method how to do it), and for each past variable we select the most probable value. Will we get the most likely explanation of a sequence of observations? Why?

  Use smoothing to find posterior distribution of rain $P(R_k|u_{1:t})$ for all time steps.

  Then, construct a sequence of most likely states

$$(\arg\max_{r_1} P(r_1|u_{1:t}), \arg\max_{r_2} P(r_2|u_{1:t}), \ldots, \arg\max_{r_t} P(r_t|u_{1:t}))$$

- Does mixing time mean that after that time, the state will not change?

  No.

- Assume that state variable X depends on two previous state variables X_{t-1} and X_{t-2}. Can we encode this transition using dependence just between subsequent states?

  Yes.

  Example: in 2nd order MC with possible states$A$, G, C, T. transition table has 4x4 values. We construct 1st MC with 16 states $AA, AC, \ldots, CA, CC, \ldots, GG$. The new transition table has 16^2 entries.

- Assume that we do full smoothing, that is we smooth every past variable (think about an efficient method how to do it), and for each past variable we select the most probable value. Will we get the most likely explanation of a sequence of observations? Why?

  Use smoothing to find posterior distribution of rain $P(R_k|u_{1:t})$ for all time steps.

  Then, construct a sequence of most likely states

$$(\arg\max_{r_1} P(r_1|u_{1:t}), \arg\max_{r_2} P(r_2|u_{1:t}), \ldots, \arg\max_{r_t} P(r_t|u_{1:t}))$$

  However, the most likely sequence is different:

$$\arg\max_{r_{1:t}} P(r_{1:t}|u_{1:t})$$