UNIVERSITY OF LINCOLN

CMP9132M

Advanced Artificial Intelligence

Lecture 6: “Hidden Markov Models I”

Nicola Bellotto & Heriberto Cuayahuitl

Hidden Markov Models

- So far we assumed that we could directly observe the state of the weather, but what if this state is **hidden**?

- Example: I am in an office without windows. The only information I have about the weather ***w*** is through the umbrella ***u*** carried by a colleague.

$$P(w_t \mid u_t) = ?$$

HMM Assumptions

- (First-order) Markov assumption, plus…

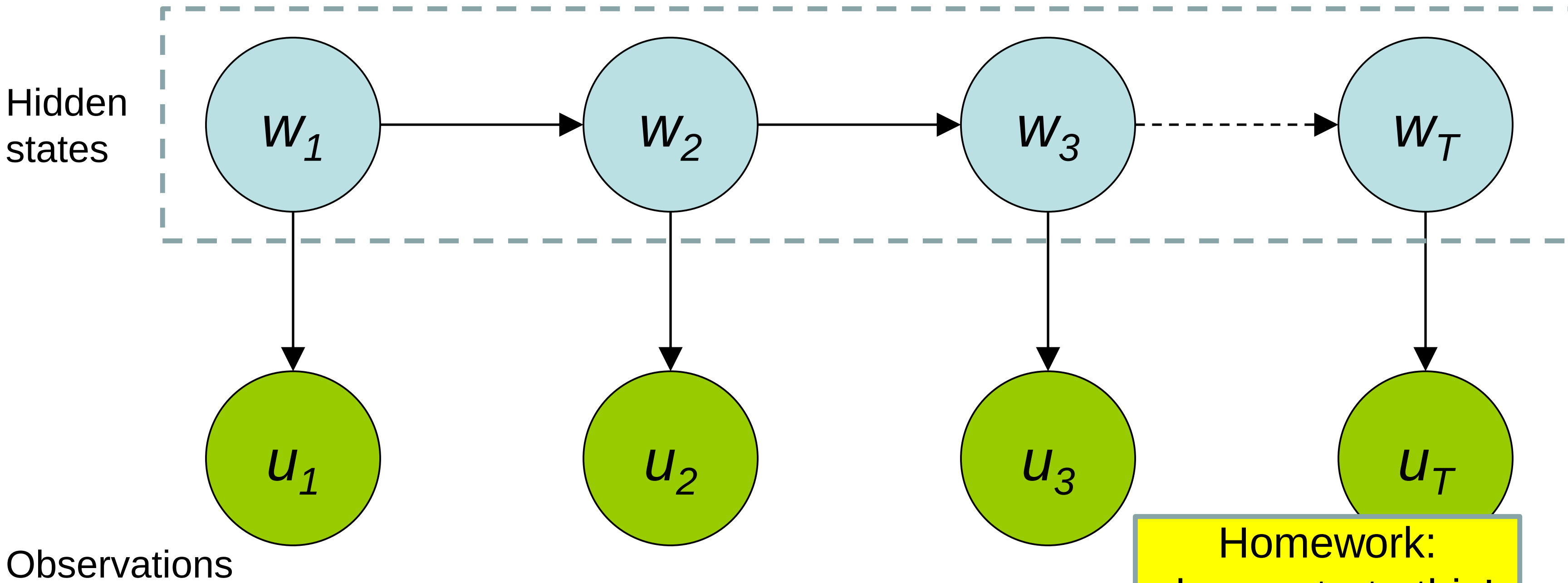
- **Sensor Markov Assumption**
  - The current observation depends only on the current state
    - E.g. – The conditional probability of observing an umbrella today (u_t), given today's weather (w_t) and previous information, depends only on w_t

- That is

$$\boxed{\mathbf{P}(u_t \mid w_t, w_{t-1}, w_{t-2}, \ldots, u_{t-1}, u_{t-2}, \ldots) = \mathbf{P}(u_t \mid w_t)}$$

HMM Graphical Model

UNIVERSITY OF LINCOLN



Homework: demonstrate this!

$$P(w_1,\ldots,w_T,u_1,\ldots,u_T) = P(w_1)P(u_1|w_1)\ \Pi_{i=2..T}[P(w_i|w_{i-1})P(u_i|w_i)]$$

Transition distribution

- Consider the table of conditional probabilities **P(w_t|w_{t-1})** – probability of weather tomorrow (w_t) given today's weather (w_{t-1}) – i.e. transition distribution

		Tomorrow's Weather		
		Sunny	Rainy	Foggy
Today's Weather	Sunny	0.8	0.05	0.15
	Rainy	0.2	0.6	0.2
	Foggy	0.2	0.3	0.5

Emission distribution

- Let's assume we can only get information indirectly about the weather by observing someone carrying an umbrella
- The conditional probability **P(u_t = True | w_t)** of observing the umbrella u_t given the weather w_t is expressed by the emission distribution

	Probability of Umbrella
Sunny	0.1
Rainy	0.8
Foggy	0.3

what are the values P(u_t = False | w_t) ?

HMM Filtering

- A common problem is to compute the probability distribution of a set of (hidden) state variables x_t given a sequence of observations $e_1, e_2, \ldots, e_t$ (emissions)
- For example the probability distribution of today's weather given the umbrella observations of today, yesterday, the day before, etc.
- In general, that means

$$P(x_t \mid e_1, e_2, \ldots, e_t) \Rightarrow \underline{\text{filtering}}$$

Forward algorithm

$$
\begin{aligned}
P(x_t \mid e_1,e_2,\ldots,e_t) &= P(x_t \mid e_1,e_2,\ldots,e_{t-1},e_t) \\
&= P(x_t, e_t \mid e_1,e_2,\ldots,e_{t-1}) \,/\, P(e_t \mid e_1,e_2,\ldots,e_{t-1}) \\
&= \alpha_t\, P(e_t \mid x_t, \cancel{e_1,e_2,\ldots,e_{t-1}})\, P(x_t \mid e_1,e_2,\ldots,e_{t-1}) \\
&= \alpha_t\, P(e_t \mid x_t)\, P(x_t \mid e_1,e_2,\ldots,e_{t-1}) \\
&= \alpha_t\, P(e_t \mid x_t) \textstyle\sum_{x_{t-1}} [P(x_t \mid x_{t-1}, \cancel{e_1,e_2,\ldots,e_{t-1}}) P(x_{t-1} \mid e_1,e_2,\ldots,e_{t-1})] \\
&= \alpha_t\, P(e_t \mid x_t) \textstyle\sum_{x_{t-1}} [P(x_t \mid x_{t-1}) P(x_{t-1} \mid e_1,e_2,\ldots,e_{t-1})]
\end{aligned}
$$

See “marginalization”

- α_t is a normalization factor to guarantee that

$$\sum_{x_t} P(x_t \mid e_1,e_2,\ldots,e_t) = 1 \quad \Rightarrow \quad \alpha_t = 1 \,/ \sum_{x_t} P(x_t \mid e_1,e_2,\ldots,e_t)$$

Forward algorithm

- In summary, filtering can be done recursively by the forward algorithm using the emission and transition probabilities as follows:

$$P(x_t \mid e_1, \ldots, e_t) = \alpha_t \, P(e_t \mid x_t) \, \Sigma_{x_{t-1}} [P(x_t \mid x_{t-1}) \, P(x_{t-1} \mid e_1, \ldots, e_{t-1})]$$

$$f_t = \alpha_t \, \mathbf{FORWARD}(f_{t-1}, e_t)$$

where α_t is a normalization factor and the initial condition $f_0 = P(x_0)$ is known

Filtering: weather example

- Consider only two types of weather: Rainy and Sunny (or R and ¬R). What is the probability that tomorrow is Rainy given that today and tomorrow we observe an umbrella?

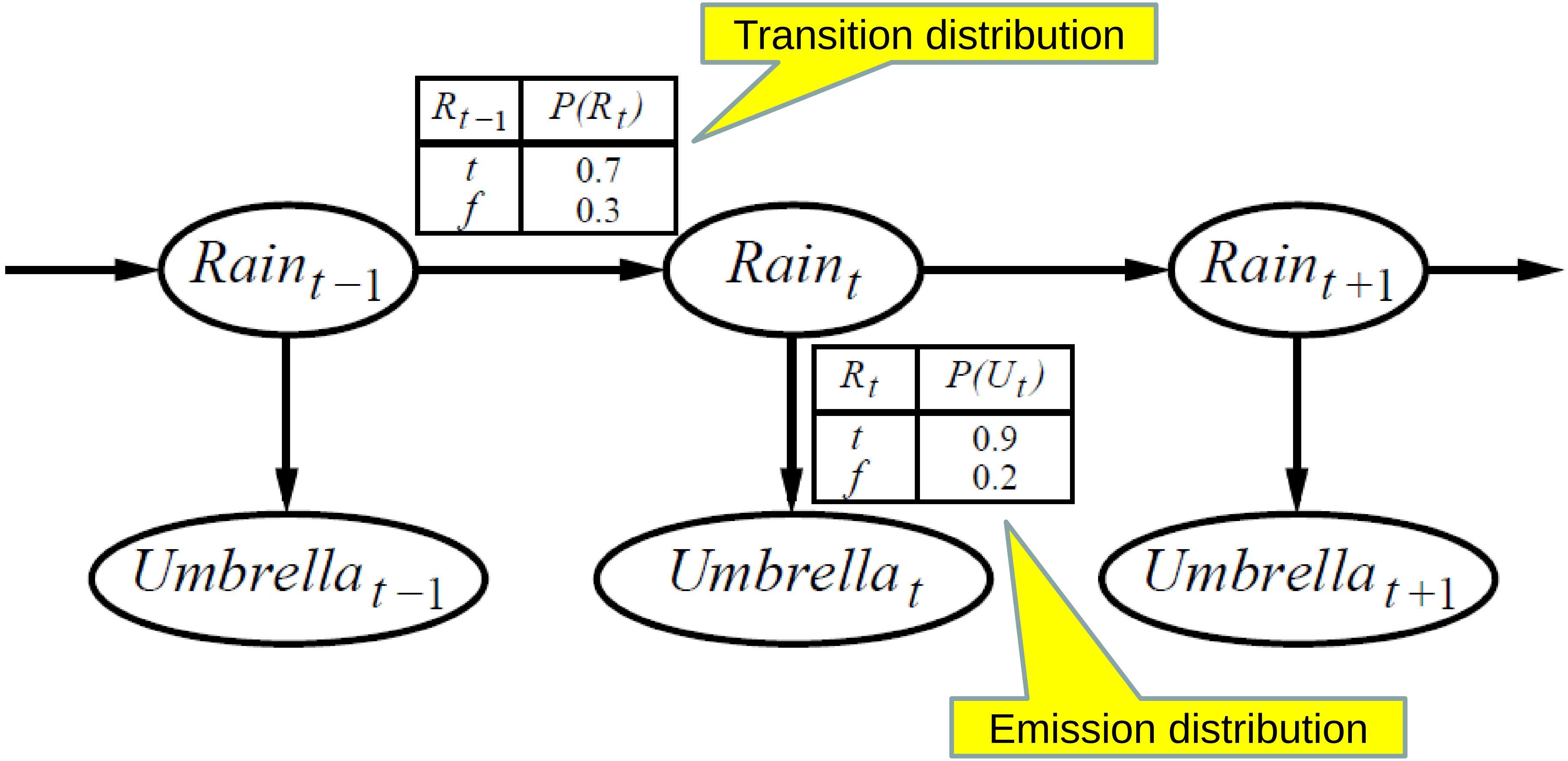
$$P(w_2 = R \mid u_1 = T, u_2 = T)$$

- Assuming equiprobability of the initial states:

$$P(w_0 = R) = P(w_0 = S) = 0.5$$

- $f_1 = \alpha_1$ **FORWARD**$(P(w_0), u_1 = \text{True})$
- $f_2 = \alpha_2$ **FORWARD**$(f_1, u_2 = \text{True})$

Filtering: weather example



R_{t-1}	$P(R_t)$
t	0.7
f	0.3

R_t	$P(U_t)$
t	0.9
f	0.2

Filtering: weather example

- $f_1 = \alpha_1$ **FORWARD**$(P(w_0), u_1{=}T)$

$P(w_1 \mid u_1{=}T)$

$= \alpha_1 P(u_1{=}T|w_1) [P(w_1|w_0{=}R) P(w_0{=}R) + P(w_1|w_0{=}S) P(w_0{=}S)]$

$= \alpha_1 P(u_1{=}T|w_1) [P(w_1|w_0{=}R) * 0.5 + P(w_1|w_0{=}S) * 0.5]$

(for $w_1 = R$)

$= \alpha_1 P(u_1{=}T|w_1{=}R) [P(w_1{=}R|w_0{=}R) * 0.5 + P(w_1{=}R|w_0{=}S) * 0.5]$

$= \alpha_1 * 0.9 * [0.7 * 0.5 + 0.3 * 0.5] = \alpha_1 * 0.9 * \mathbf{0.5} = \mathbf{0.818}$

(for $w_1 = S$)

$= \alpha_1 P(u_1{=}T|w_1{=}S) [P(w_1{=}S|w_0{=}R) * 0.5 + P(w_1{=}S|w_0{=}S) * 0.5]$

$= \alpha_1 * 0.2 * [0.3 * 0.5 + 0.7 * 0.5] = \alpha_1 * 0.2 * \mathbf{0.5} = \mathbf{0.182}$

Filtering: weather example

- $f_2 = \alpha_2$ **FORWARD**$(f_1, u_2=T)$

$P(w_2 \mid u_1=T, u_2=T)$

(for w_2 = R)

$= \alpha_2 P(u_2=T|w_2=R)[P(w_2=R|w_1=R)*0.818+P(w_2=R|w_1=S)*0.182]$

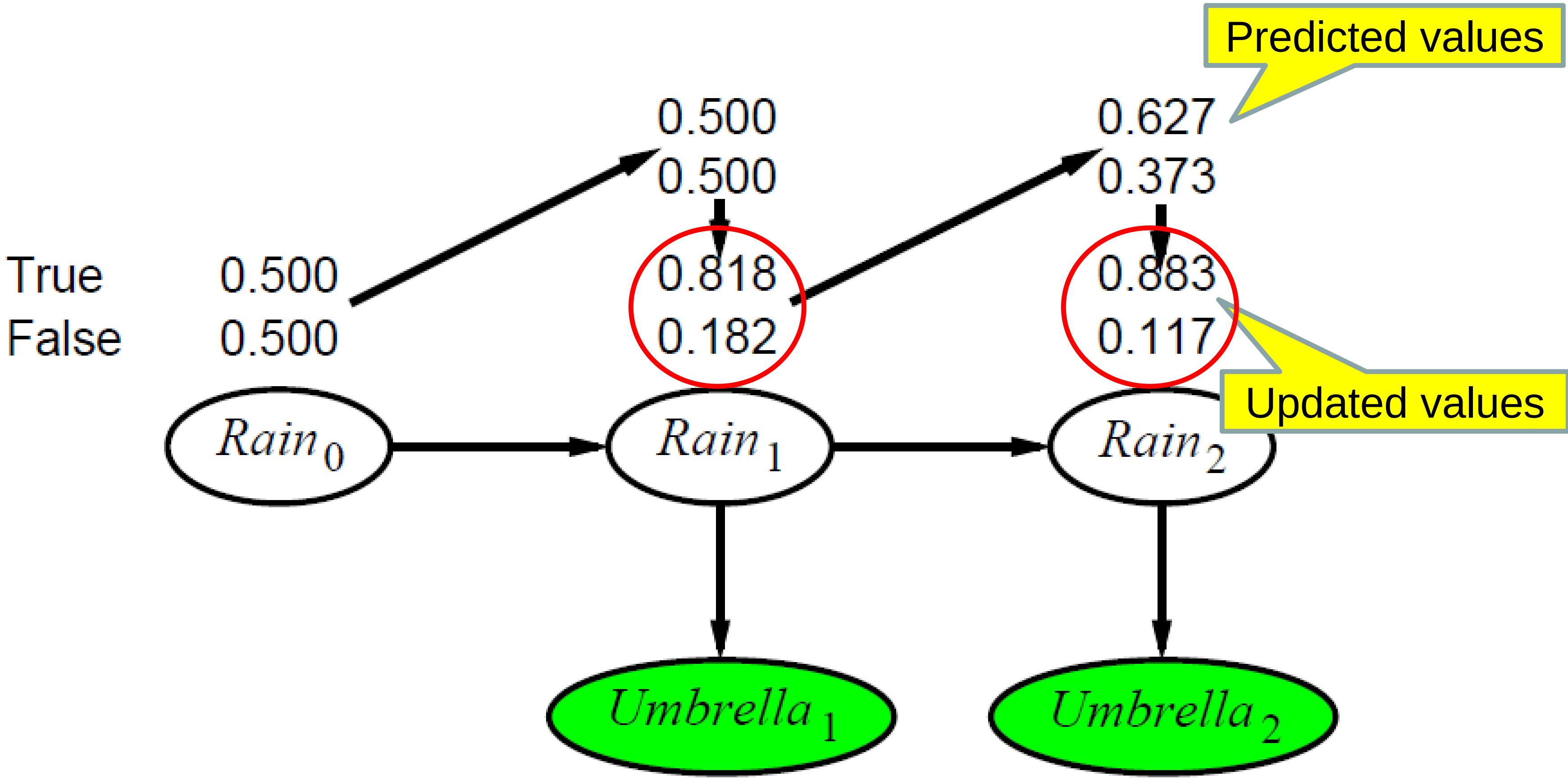
$= \alpha_2\, 0.9 * [0.7 * 0.818 + 0.3 * 0.182] = \alpha_2 * 0.9 * \mathbf{0.627} = \mathbf{0.883}$

(for w_2 = S)

$= \alpha_2 P(u_2=T|w_2=S)[P(w_2=S|w_1=R)*0.818+P(w_2=S|w_1=S)*0.182]$

$= \alpha_2\, 0.2 * [0.3 * 0.818 + 0.7 * 0.182] = \alpha_2 * 0.2 * \mathbf{0.373} = \mathbf{0.117}$

Filtering: weather example



Conclusions

- Suggested reading
  - Tutorial on Blackboard (try the exercises!)
  - Chapter 15 of Russell & Norvig
- Next week
  - Hidden Markov Models II
- Any question?

*Important Rules

- **Summing out rule** (or **marginalization**)

$$P(A) = \sum_B P(A, B)$$

- Can also be written

$$P(A) = \sum_B P(A \mid B)\, P(B)$$