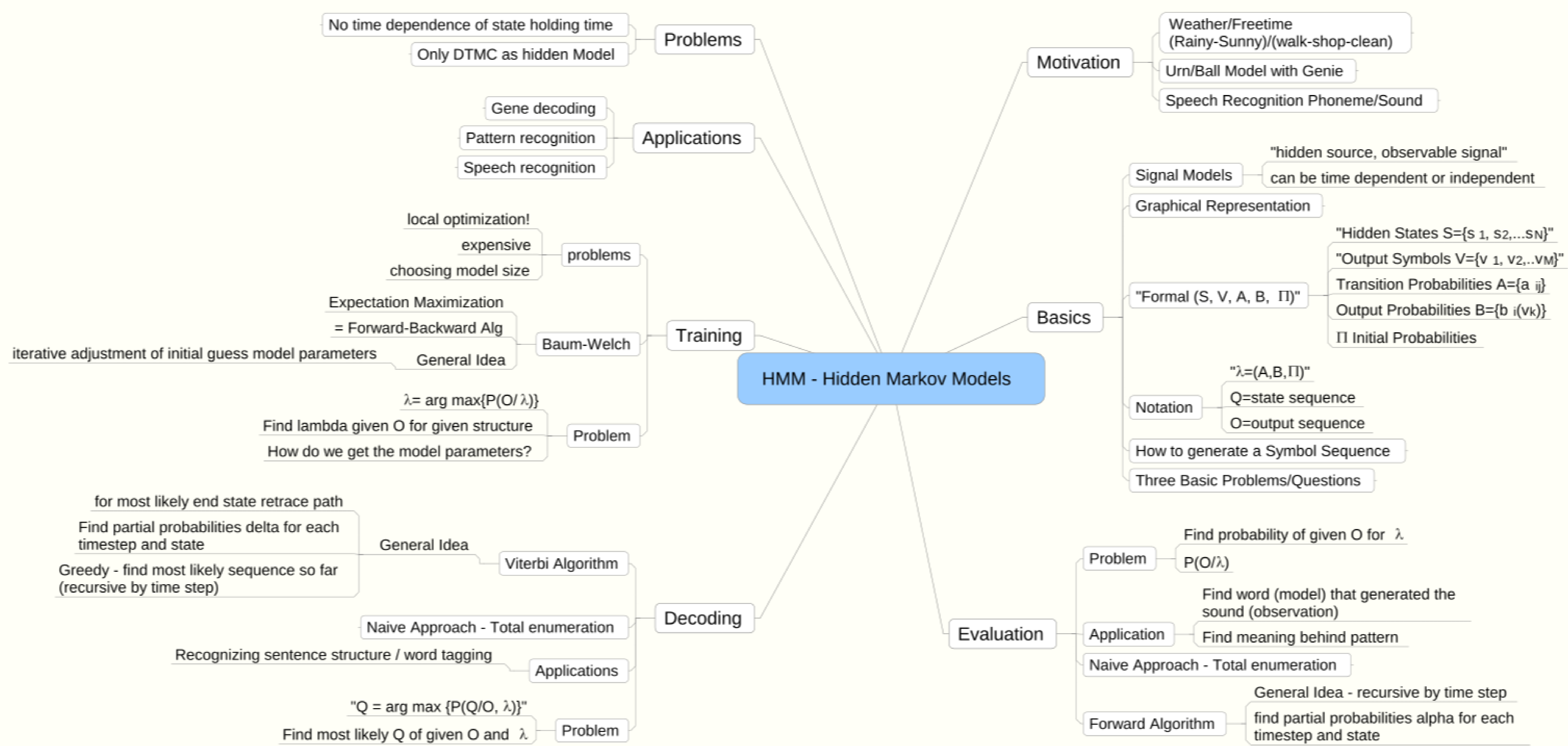
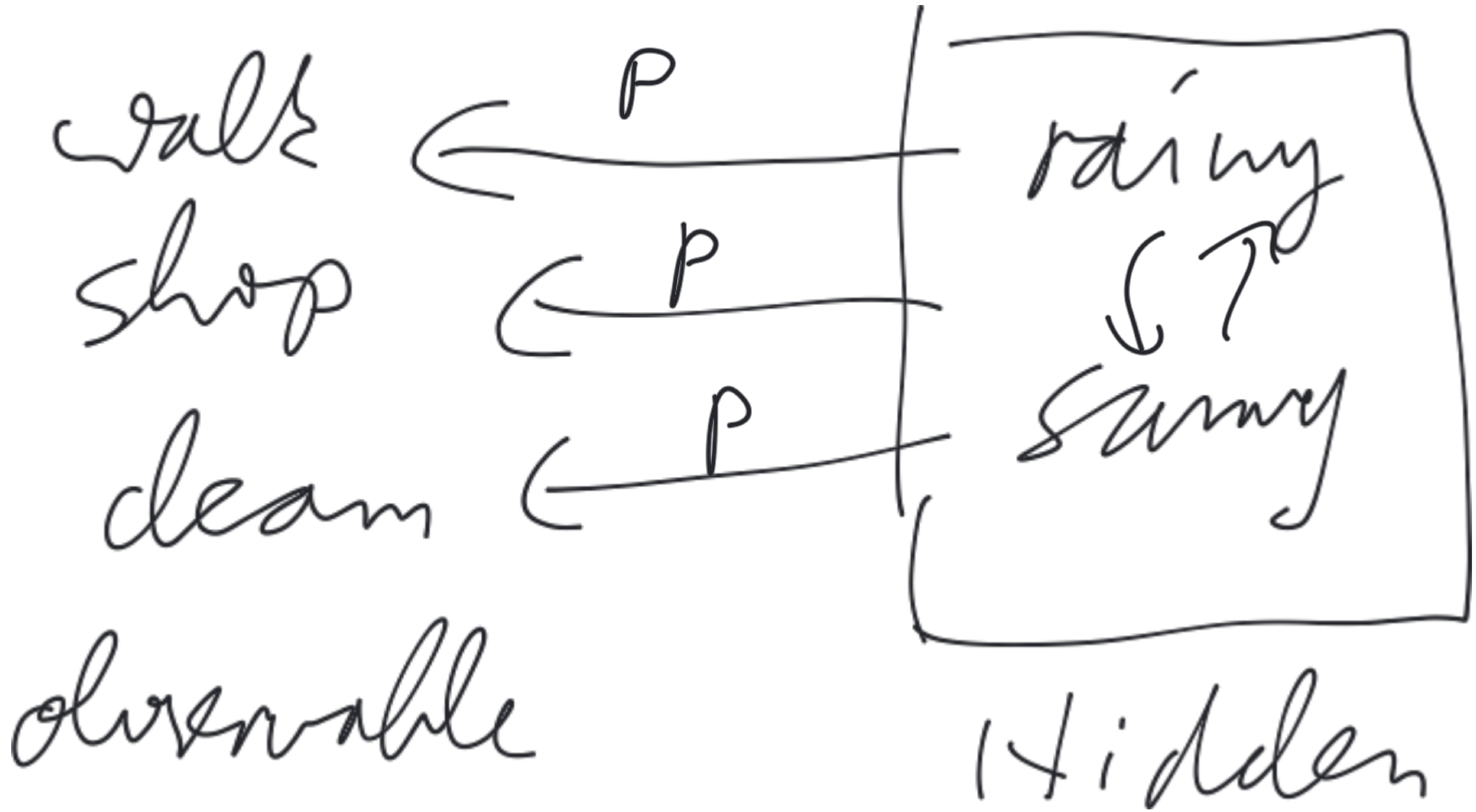

HMM - Hidden Markov Models
Problems
No time dependence of state holding time
Only DTMC as hidden Model
Applications
Gene decoding
Pattern recognition
Speech recognition
Training
problems
local optimization!
expensive
choosing model size
Baum-Welch
Expectation Maximization
= Forward-Backward Alg
General Idea
iterative adjustment of initial guess model parameters
Problem
λ= arg max{P(O/ λ)}
Find lambda given O for given structure
How do we get the model parameters?
Decoding
Viterbi Algorithm
General Idea
for most likely end state retrace path
Find partial probabilities delta for each timestep and state
Greedy - find most likely sequence so far (recursive by time step)
Naive Approach - Total enumeration
Applications
Recognizing sentence structure / word tagging
Problem
"Q = arg max {P(Q/O, λ)}"
Find most likely Q of given O and λ
Motivation
Weather/Freetime (Rainy-Sunny)/(walk-shop-clean)
Urn/Ball Model with Genie
Speech Recognition Phoneme/Sound
Basics
Signal Models
"hidden source, observable signal"
can be time dependent or independent
Graphical Representation
"Formal (S, V, A, B, Π)"
"Hidden States S={s 1, s2,...sN}"
"Output Symbols V={v 1, v2,..vM}"
Transition Probabilities A={a ij}
Output Probabilities B={b i(vk)}
Π Initial Probabilities
Notation
"λ=(A,B,Π)"
Q=state sequence
O=output sequence
How to generate a Symbol Sequence
Three Basic Problems/Questions
Evaluation
Problem
Find probability of given O for λ
P(O/λ)
Application
Find word (model) that generated the sound (observation)
Find meaning behind pattern
Naive Approach - Total enumeration
Forward Algorithm
General Idea - recursive by time step
find partial probabilities alpha for each timestep and state

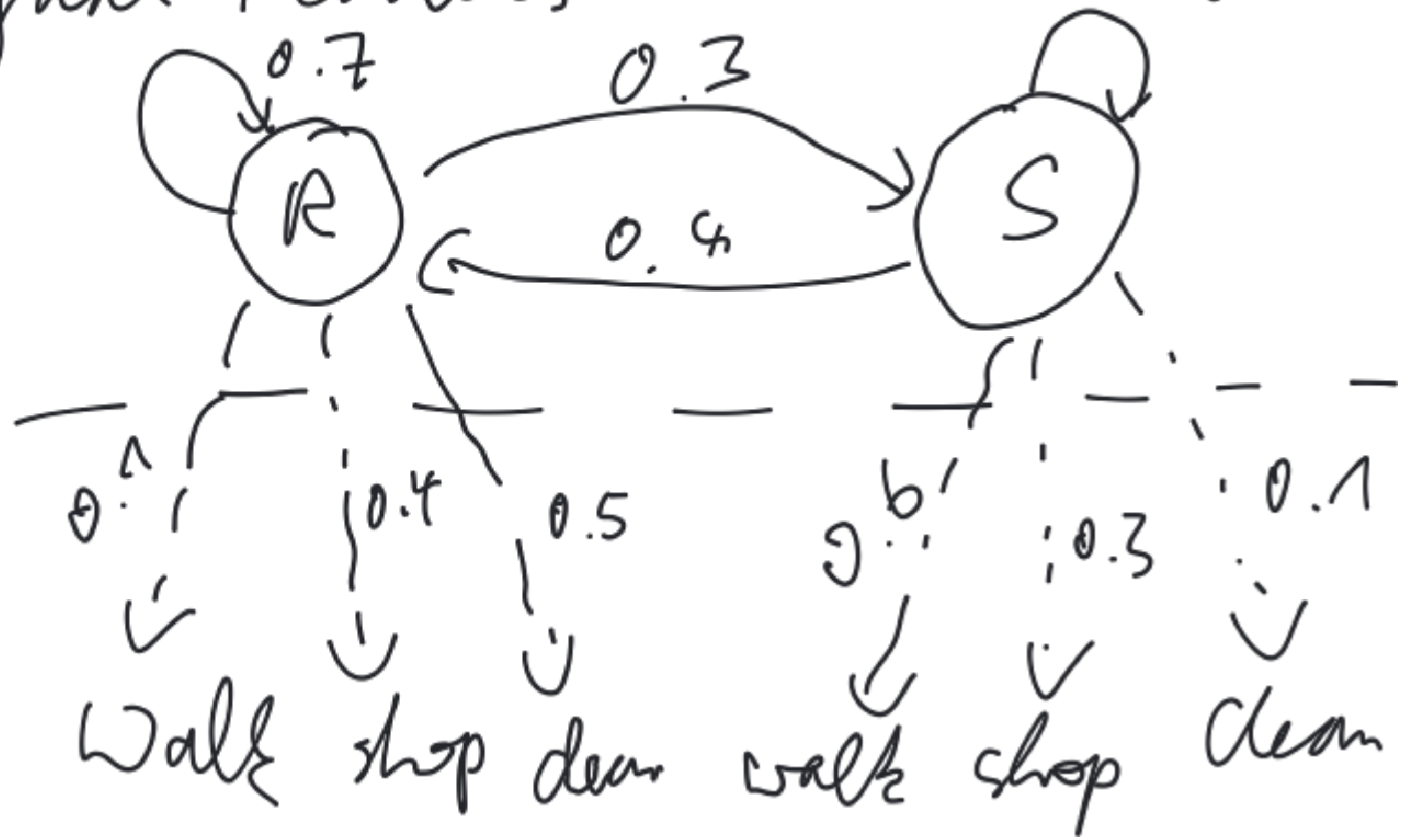

walk
shop
clean
observable
p
p
p
rainy
sunny
Hidden

It

meaning
Berlin

Signal Models

HMM



$HMM = (S, V, A, B, \pi)$

$S = (s_1, s_2 \dots s_N)$

$V = (v_1, v_2 \dots v_M)$

$A = \{a_{ij}\}_{N \times N}$

$B = \{b_i(v_k)\}$

$\pi = (\pi_1 \dots \pi_N)$

$\lambda = (A, B, \pi)$

$O = (o_1\ o_2 \dots o_T)$

trace

$Q = (q_1\ q_1 \dots q_T)$

path

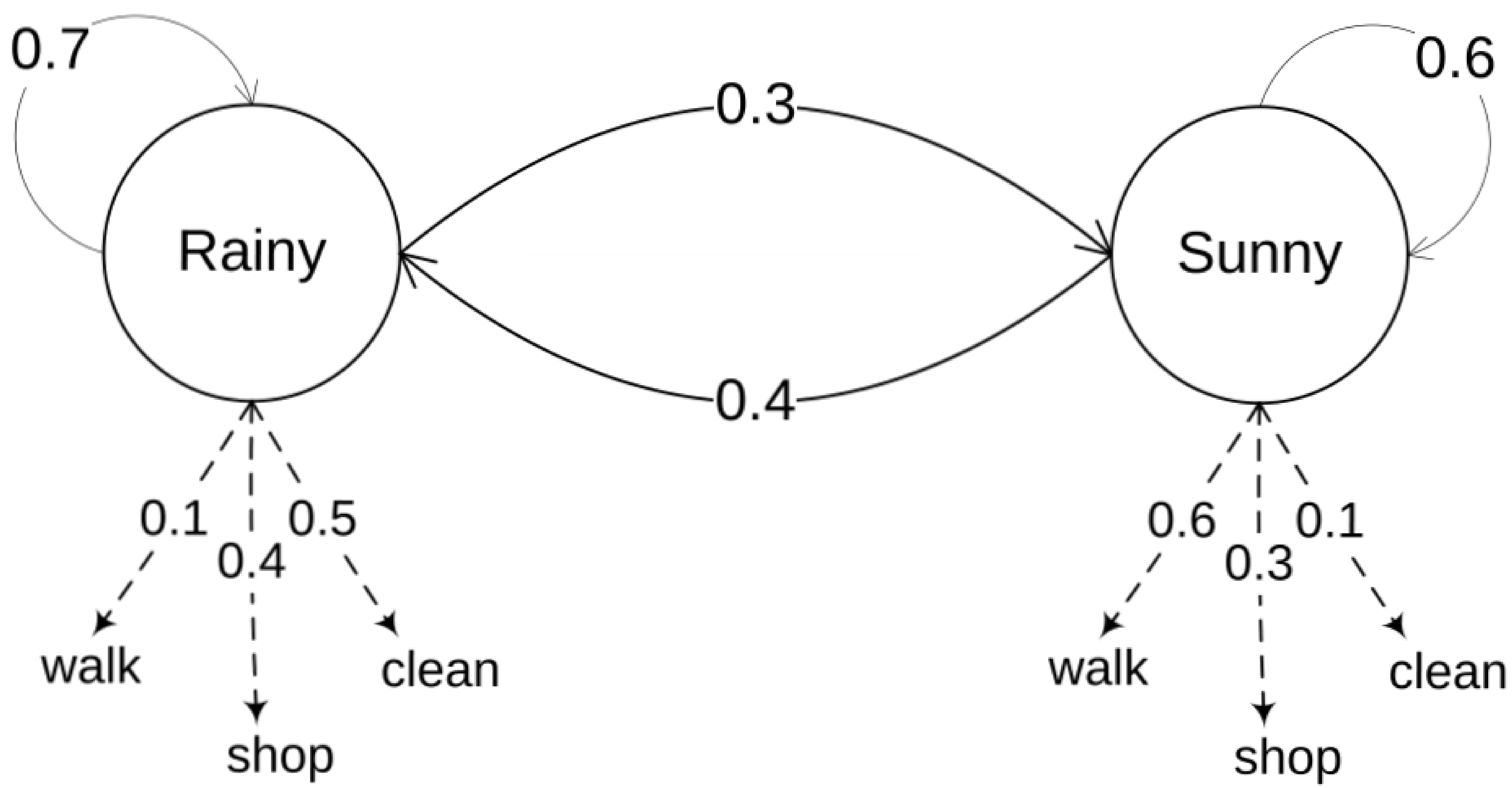
$$S = \{R, S\}$$

$$V = \{w, s, c\} \qquad \pi = (0.6, 0.4)$$

$$A = \begin{bmatrix} 0{,}7 & 0{,}3 \\ 0.4 & 0.6 \end{bmatrix}$$

$$B = \begin{array}{c c} & \begin{array}{ccc} w & s & c \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.6 & 0.3 & 0.1 \end{bmatrix} \end{array}$$

ADM - Hidden Markov Models



$$\begin{array}{ccccc} R & \to & S & \to & S \to \\ \downarrow & & \downarrow & & \downarrow \\ s & & w & & s \end{array}$$

$$O = (c, s, w)$$

(S,V,A,B,Π) = ({Rainy,Sunny},{walk,shop,clean},A,B,(0.6,0.4))

$$A = \begin{array}{c} \\ R \\ S \end{array} \begin{array}{c} \begin{array}{cc} R & S \end{array} \\ \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \end{array} \qquad B = \begin{array}{c} \\ R \\ S \end{array} \begin{array}{c} \begin{array}{ccc} w & s & c \end{array} \\ \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.6 & 0.3 & 0.1 \end{bmatrix} \end{array}$$

$\lambda = (A, B, \pi)$, O

① Evaluation $\lambda, O \rightarrow P(O \mid \lambda)$

② Decoding $\lambda, O \rightarrow Q$

③ Training $O \rightarrow \lambda$

(1) $\lambda, O \rightarrow P(O/\lambda)$

$\lambda_1, \lambda_2, \lambda_3$

(c, s, w)

$P(RRR, csw, \lambda) = \sum \rightarrow P(O/\lambda)$

RRS

RSR

RSS

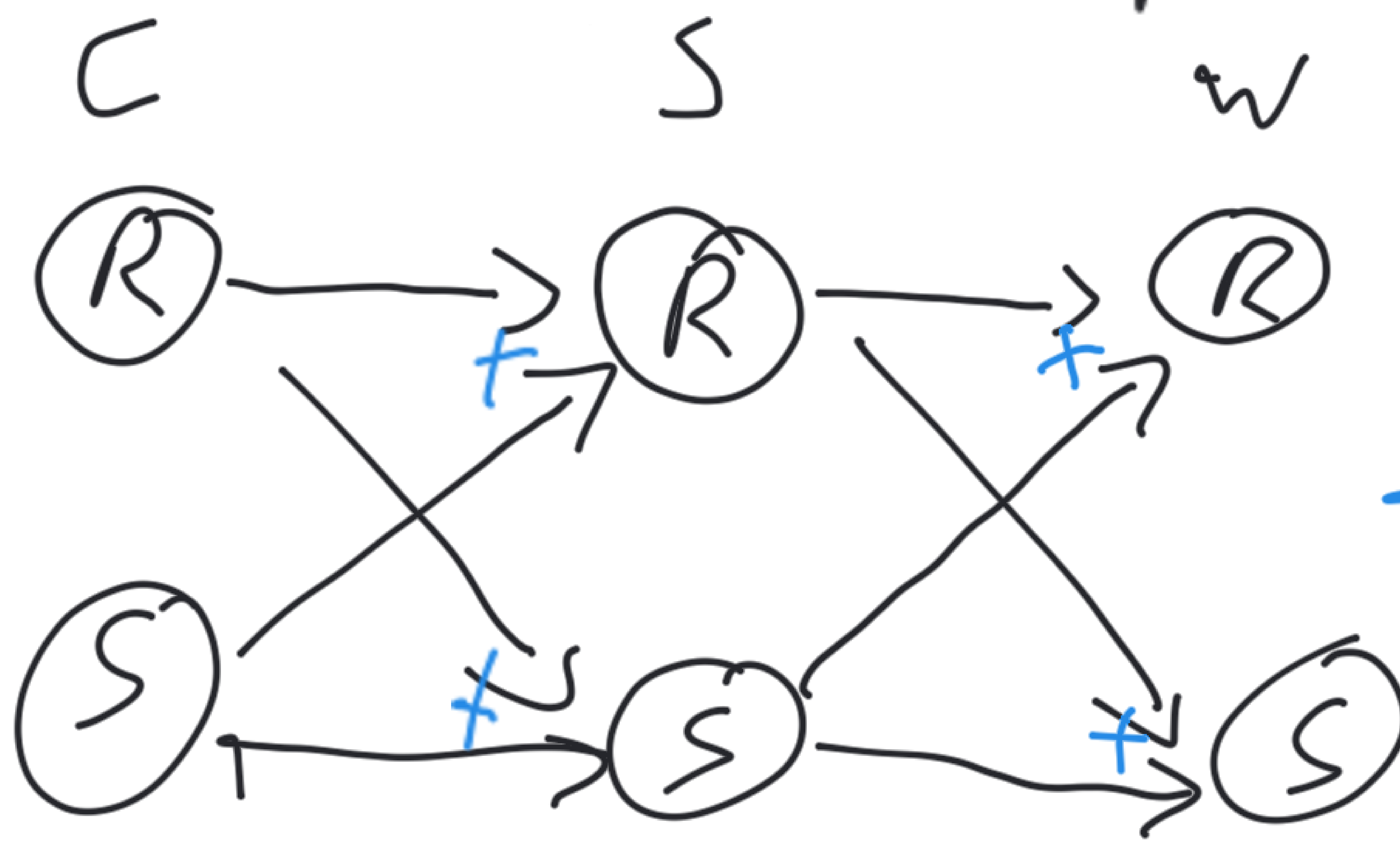
$\vdots$

① $\lambda, O \rightarrow P(O/\lambda)$

Forward Alg.

C S W

partial prob.



$+ \; P(O/\lambda)$

② $\lambda, O \rightarrow Q$ $\quad$ C S W

$$Q = \arg\max_{Q} \{ P(Q \mid O, \lambda) \}$$

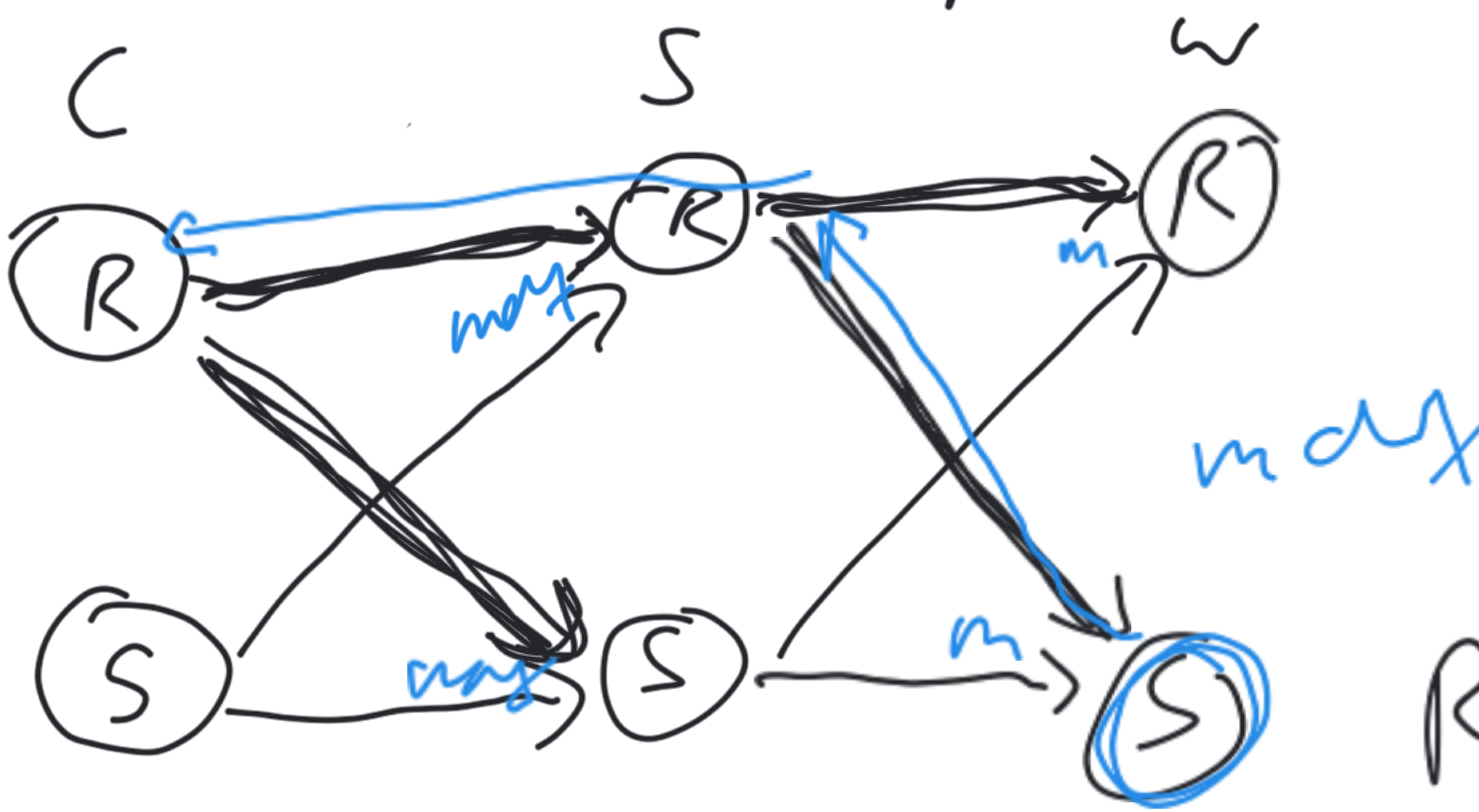
$$\left.\begin{array}{l} P(RRR, CSW, \lambda) = \\ P(RRS, CSW, \lambda) = \\ \vdots \end{array}\right\} \max$$

O

② $\lambda, O \rightarrow Q$ C S W

$$Q = \underset{Q}{\operatorname{argmax}} \{ P(Q \mid O, \lambda) \}$$

Viterbi Alg. partial prob



RRS

(3) $O_1 O \ldots \longrightarrow \lambda$

$$\lambda = \arg\max_{\lambda} \{ P(O \mid \lambda) \}$$

Baum-Welch $\quad \lambda_0 \to \lambda_1 \to \lambda_2 \ldots\ldots$

F/B-Alg