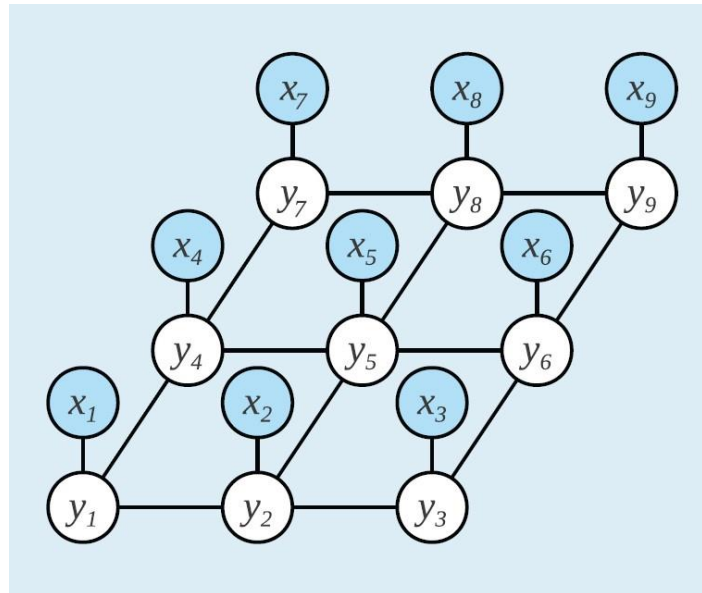


# Probabilistic Graphical Models in Bioinformatics

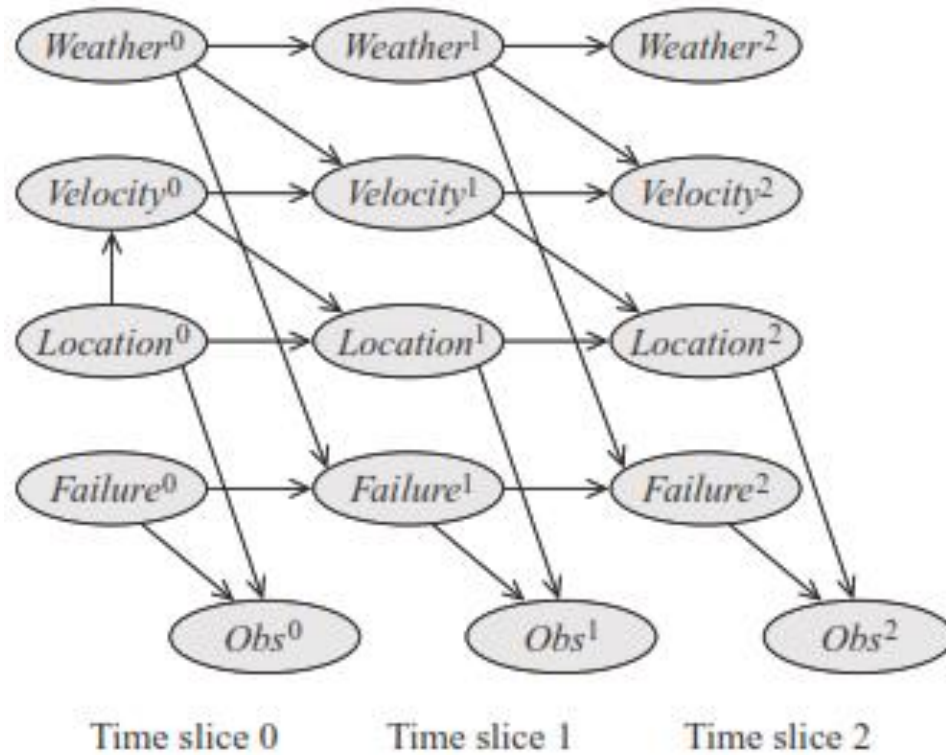
## Lecture 14: Bayesian networks for temporal data



# Temporal models

- Focus is to model dynamical settings in terms of a *system state* over time
  - System state: value at time  $t$  is a snapshot of the relevant attributes (hidden or observed).
- We use  $X_i^{(t)}$  to represent the instantiation of the variable  $X_i$  at time  $t$ 
  - Note that variable  $X_i$  is no longer a random variable rather a *template variable*.
- **Trajectory:** an assignment of values to each variable  $X_i^{(t)}$  for each relevant time  $t$ .
- Goal is to represent a joint probability distribution over such trajectories.

# Example- vehicle localization task



# Basic assumptions

- Discretize timeline into time slices
- We want to represent  $P(\mathcal{X}^{(0)}, \mathcal{X}^{(1)}, \dots, \mathcal{X}^{(T)})$  often abbreviated as  $P(\mathcal{X}^{(0:T)})$
- Using chain rule

$$P(\mathcal{X}^{(0:T)}) = P(\mathcal{X}^{(0)}) \prod_{t=0}^{T-1} P(\mathcal{X}^{(t+1)} \mid \mathcal{X}^{(0:t)}).$$

- **Question:** can you draw the corresponding Bayesian network?

# Markov assumption

- In order to simplify  $P(\mathcal{X}^{(0:T)})$ , a natural approach is to assume that future is conditionally independent of the past given the present.
- A dynamic system over the template variables  $\mathcal{X}$  satisfies the Markov assumption if

$$(\mathcal{X}^{(t+1)} \perp \mathcal{X}^{(0:(t-1))} \mid \mathcal{X}^{(t)}).$$

- Hence we have the following compact representation:

$$P(\mathcal{X}^{(0)}, \mathcal{X}^{(1)}, \dots, \mathcal{X}^{(T)}) = P(\mathcal{X}^{(0)}) \prod_{t=0}^{T-1} P(\mathcal{X}^{(t+1)} \mid \mathcal{X}^{(t)}).$$

# Stationary Markovian system

- We say that a Markovian dynamic system is stationary if  $P(\mathcal{X}^{(t+1)} \mid \mathcal{X}^{(t)})$  is the same for all  $t$ .
  - The system is also called time invariant or homogenous.

$$P(\mathcal{X}^{(t+1)} = \xi' \mid \mathcal{X}^{(t)} = \xi) = P(\mathcal{X}' = \xi' \mid \mathcal{X} = \xi).$$

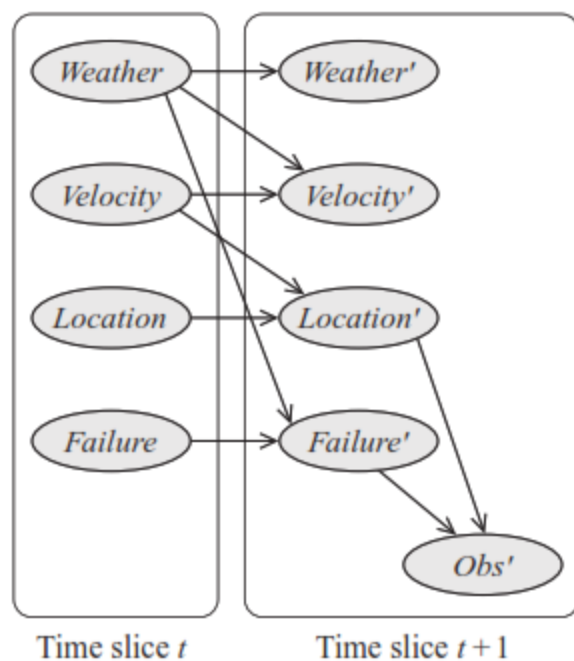
- hence, we can represent the temporal process using a transition model  $P(\mathcal{X}' \mid \mathcal{X})$

# Dynamic Bayesian networks

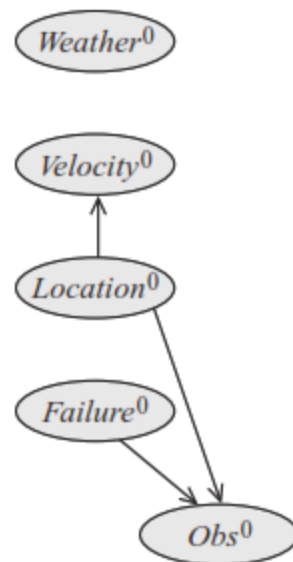
- The Markov and stationary assumptions allow us to represent the probability distribution over infinite trajectories compactly.
- We need only represent the initial state distribution and the transition model  $P(\mathcal{X}' | \mathcal{X})$ .
- The transition model can be represented using a *conditional Bayesian network* referred to as a 2-time-slice Bayesian network (2-TBN) (Section 5.6).

$$P(\mathcal{X}' | \mathcal{X}) = \prod_{i=1}^n P(X'_i | \text{Pa}_{X'_i}).$$

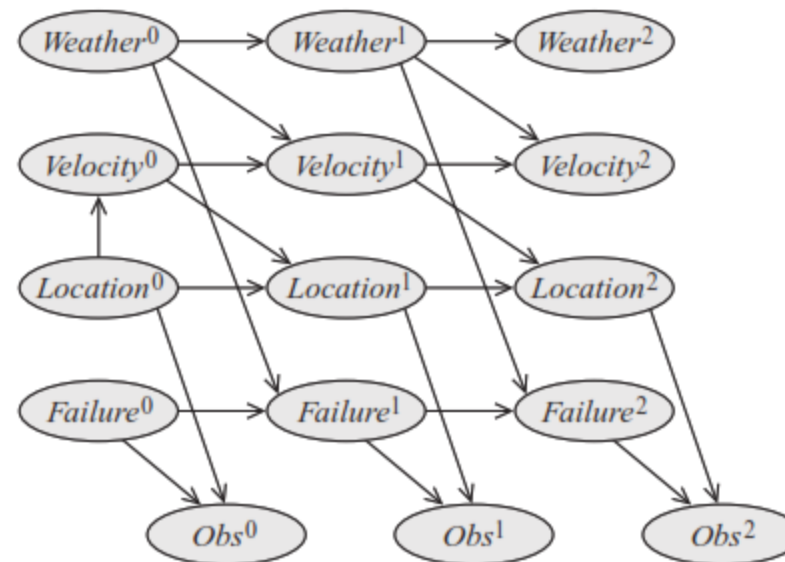
# Example: a simplified DBN for monitoring a vehicle



(a)  $\mathcal{B}_{\rightarrow}$



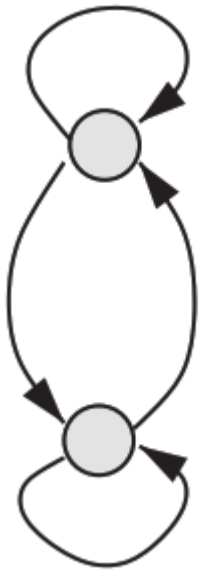
(b)  $\mathcal{B}_0$



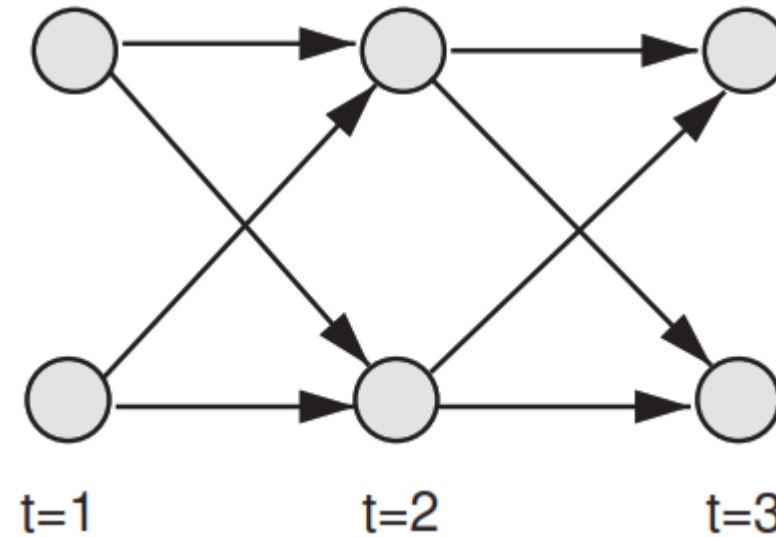
(c) DBN unrolled over 3 steps



# Unfolding a recurrent network can resolve feedback loops

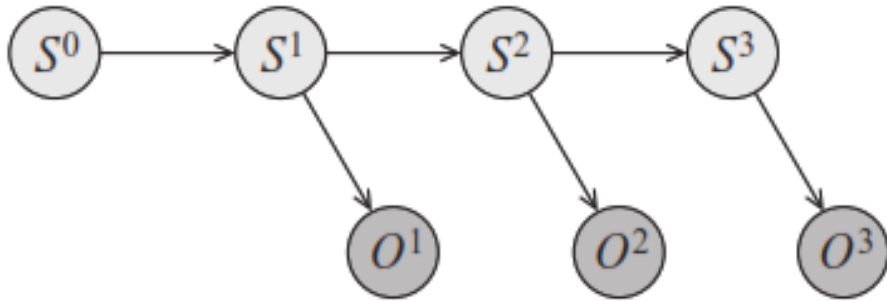


A recurrent network

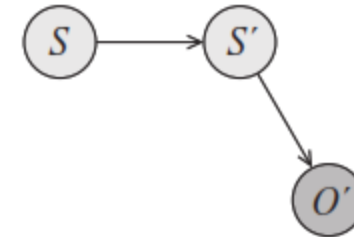


Equivalent DBN

# The hidden Markov model is a DBN



The unrolled DBN for four time slices

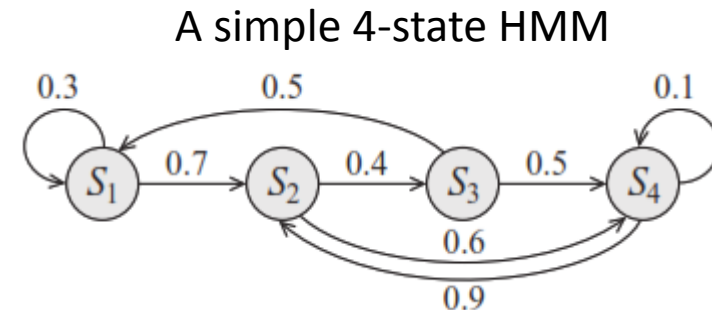


The 2-TBN for a generic HMM

# An alternative viewpoint to a temporal process

- State-observation model
  - We view the system as evolving naturally on its own, with the observations of it occurring in a separate process
  - This view separates out the system dynamics from the observational model.
- HMM is a state-observation model
  - The transition model  $P(S' | S)$  is assumed to be sparse.

	$s_1$	$s_2$	$s_3$	$s_4$
$s_1$	0.3	0.7	0	0
$s_2$	0	0	0.4	0.6
$s_3$	0.5	0	0	0.5
$s_4$	0	0.9	0	0.1

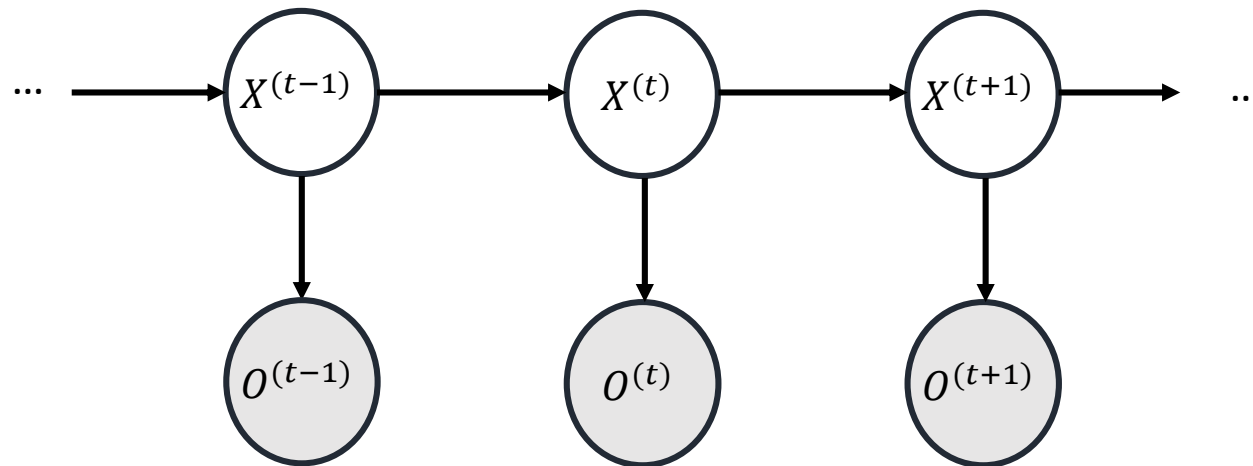


Different representation than BN

- The nodes are state
- The directed edges represent possible transitions between states.

# Linear dynamical systems

- Linear dynamical system (LDS) represents a system of one or more real-valued variables that evolve linearly over time with some Gaussian noise.
- Often referred to as *Kalman filters*
- An LDS is a DBN where
  - the variables are all continuous
  - All dependencies are linear Gaussian



$$P(\mathbf{X}^{(t)} | \mathbf{X}^{(t-1)}) = \mathcal{N}(\mathbf{A}\mathbf{X}^{(t-1)}; \mathbf{Q}),$$
$$P(O^{(t)} | \mathbf{X}^{(t)}) = \mathcal{N}(H\mathbf{X}^{(t)}; R),$$

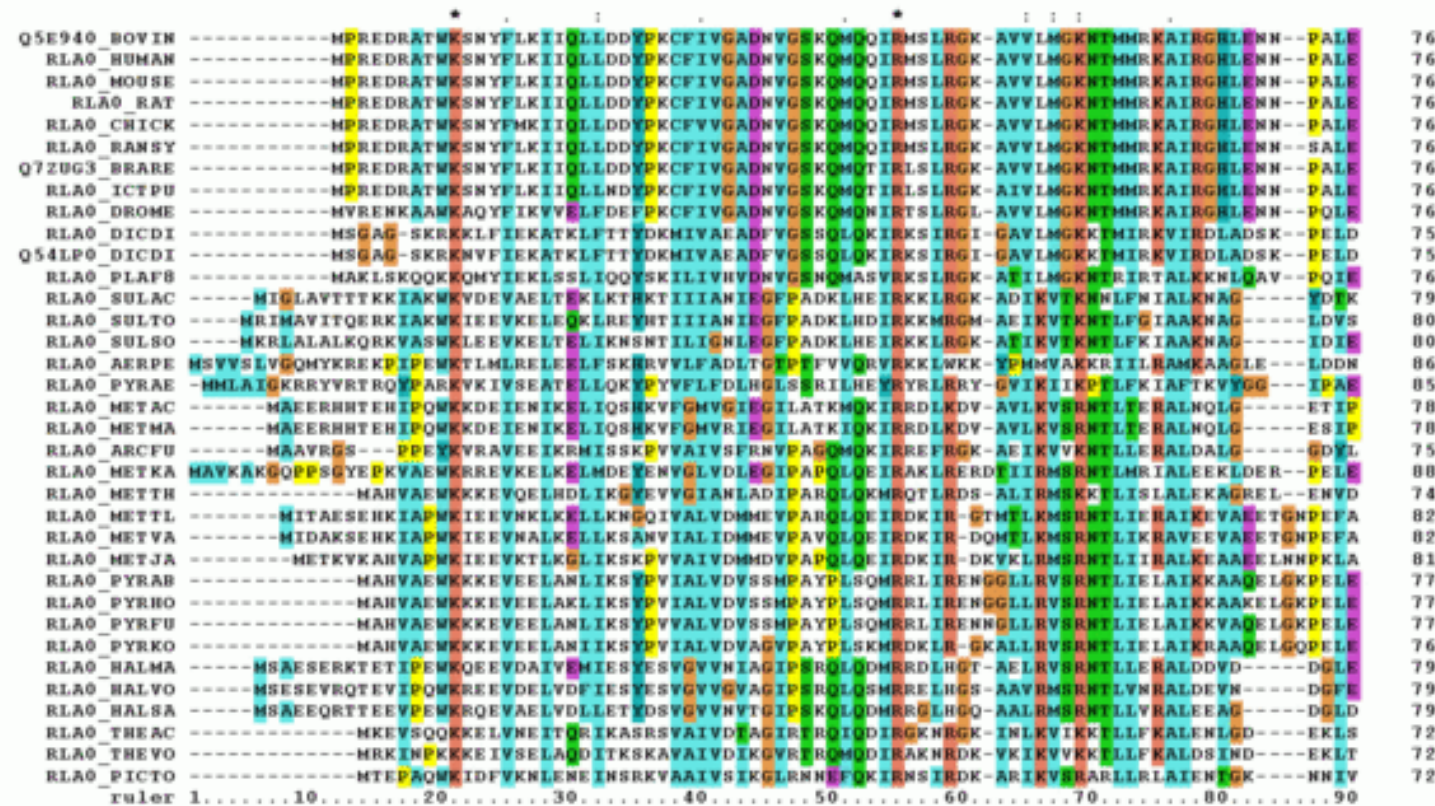
- If we allow discrete variables, the model is called *switching linear dynamical systems*.

# HMMs-main questions

- Whiteboard notes

# Application: profile HMMs for sequence alignment

- Assume given a multiple alignment of a family of proteins, the aim is to realistically model propensities of symbols in the input alignment.



- HMM models can be used to classify unknown proteins.

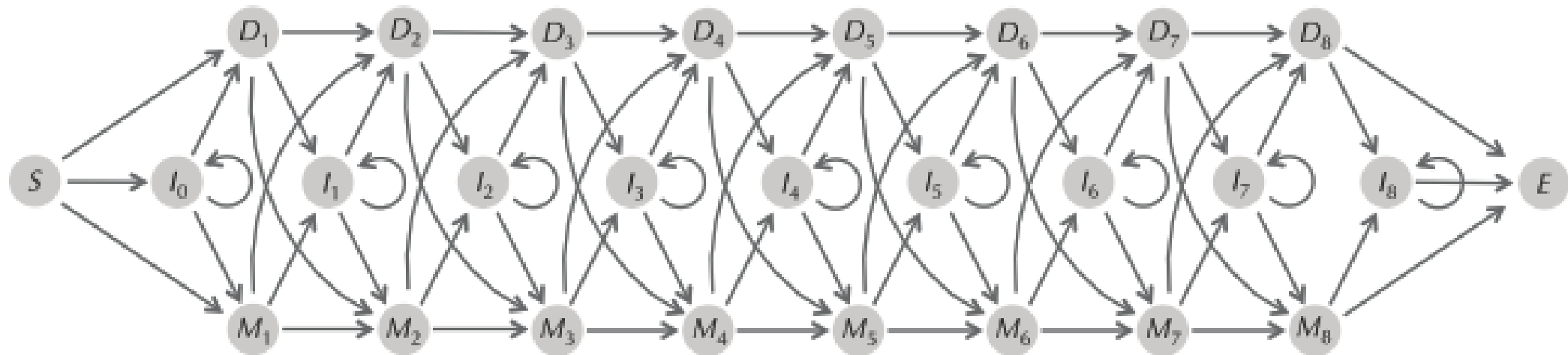
# A simple HMM

		1	2	3	4	5	6	7	8	
Alignment		A	C	D	E	F	A C	A	D	F
		A	F	D	A	-	- - -	C	C	F
		A	-	-	E	F	D -	F	D	C
		A	C	A	E	F	- - -	A	-	C
		A	D	D	E	F	A A	A	D	F
Alignment*		A	C	D	E	F		A	D	F
		A	F	D	A	-		C	C	F
		A	-	-	E	F		F	D	C
		A	C	A	E	F		A	-	C
		A	D	D	E	F		A	D	F
PROFILE(Alignment*)	A	1	0	0	1/5	0	3/5	0	0	
	C	0	2/4	0	0	0	1/5	1/4	2/5	
	D	0	1/4	3/4	0	0	0	3/4	0	
	E	0	0	0	4/5	0	0	0	0	
	F	0	1/4	0	0	1	1/5	0	3/5	
<div><div>M<sub>1</sub></div><div>→</div><div>M<sub>2</sub></div><div>→</div><div>M<sub>3</sub></div><div>→</div><div>M<sub>4</sub></div><div>→</div><div>M<sub>5</sub></div><div>→</div><div>M<sub>6</sub></div><div>→</div><div>M<sub>7</sub></div><div>→</div><div>M<sub>8</sub></div></div>										
Text		A	D	D	A	F	F	D	F	
emission probability		1	1/4	3/4	1/5	1	1/5	3/4	3/5	

This HMM is limited because we are not able to align a string of length other than 8

# Profile HMM diagram

- We can solve this issue by adding insertion and deletion states.





# Estimating transition and emission probabilities

Alignment

A	C	D	E	F	AC	A	D	F
A	F	D	A	-	-	C	C	F
A	-	-	E	F	D	-	D	C
A	C	A	E	F	-	A	-	C
A	D	D	E	F	AA	A	D	F

$$\text{transition}_{\text{MATCH}(5), \text{INSERTION}(5)} = 3/4$$

$$\text{transition}_{\text{MATCH}(5), \text{MATCH}(6)} = 1/4$$

$$\text{transition}_{\text{MATCH}(5), \text{DELETION}(6)} = 0$$

$$\text{emission}_{\text{INSERTION}(5)}(\text{A}) = 3/5$$

$$\text{emission}_{\text{INSERTION}(5)}(\text{C}) = 1/5$$

$$\text{emission}_{\text{INSERTION}(5)}(\text{D}) = 1/5$$

$$\text{emission}_{\text{INSERTION}(5)}(\text{E}) = 0$$

$$\text{emission}_{\text{INSERTION}(5)}(\text{F}) = 0$$

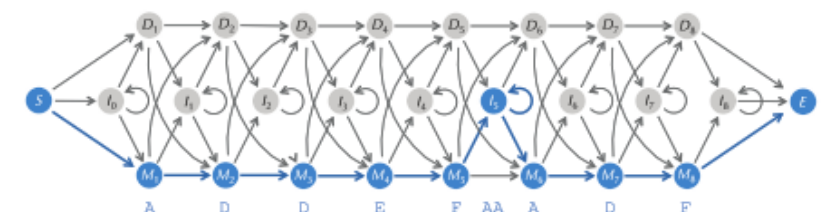
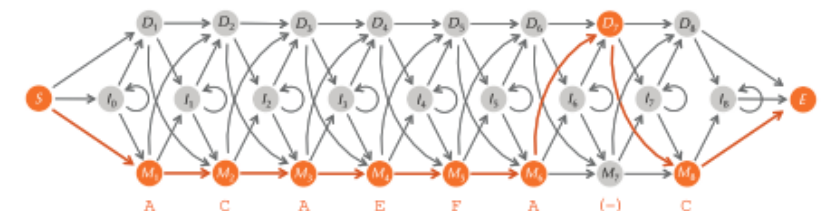
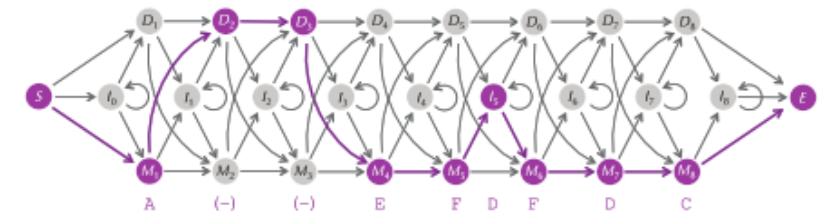
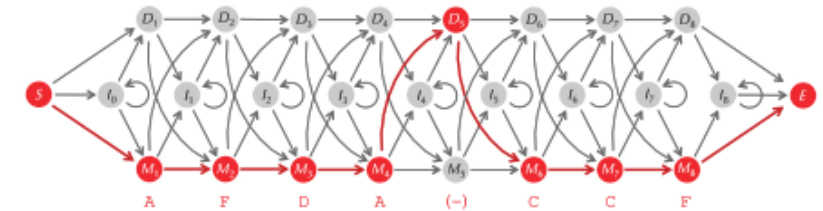
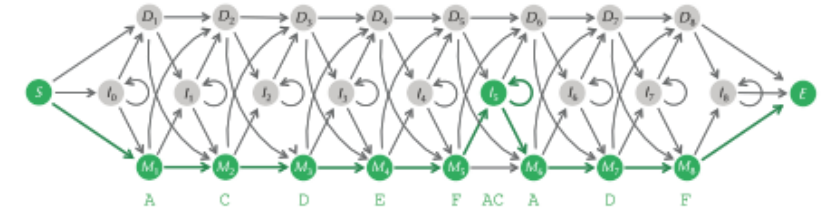
$$\text{emission}_{\text{MATCH}(2)}(\text{A}) = 0$$

$$\text{emission}_{\text{MATCH}(2)}(\text{C}) = 2/4$$

$$\text{emission}_{\text{MATCH}(2)}(\text{D}) = 1/4$$

$$\text{emission}_{\text{MATCH}(2)}(\text{E}) = 0$$

$$\text{emission}_{\text{MATCH}(2)}(\text{F}) = 1/4$$



# The complete matrix of transmission probabilities

	S	I <sub>0</sub>	M <sub>1</sub>	D <sub>1</sub>	I <sub>1</sub>	M <sub>2</sub>	D <sub>2</sub>	I <sub>2</sub>	M <sub>3</sub>	D <sub>3</sub>	I <sub>3</sub>	M <sub>4</sub>	D <sub>4</sub>	I <sub>4</sub>	M <sub>5</sub>	D <sub>5</sub>	I <sub>5</sub>	M <sub>6</sub>	D <sub>6</sub>	I <sub>6</sub>	M <sub>7</sub>	D <sub>7</sub>	I <sub>7</sub>	M <sub>8</sub>	D <sub>8</sub>	I <sub>8</sub>	E	
S			1																									
I <sub>0</sub>																												
M <sub>1</sub>						.8	.2																					
D <sub>1</sub>																												
I <sub>1</sub>																												
M <sub>2</sub>									1																			
D <sub>2</sub>										1																		
I <sub>2</sub>																												
M <sub>3</sub>											1																	
D <sub>3</sub>											1																	
I <sub>3</sub>																												
M <sub>4</sub>															.8	.2												
D <sub>4</sub>																												
I <sub>4</sub>																												
M <sub>5</sub>																	.75	.25										
D <sub>5</sub>																		1										
I <sub>5</sub>																	.4	.6										
M <sub>6</sub>																					.8	.2						
D <sub>6</sub>																												
I <sub>6</sub>																												
M <sub>7</sub>																								1				
D <sub>7</sub>																									1			
I <sub>7</sub>																												
M <sub>8</sub>																												1
D <sub>8</sub>																												
I <sub>8</sub>																												
E																												

Profile HMM is  
not stationary

# Membership detection

- **Given**
  - a profile HMM  $\mathcal{M}$  and a sequence  $x$
- We want to decide whether  $x$  belongs to the set of sequences represented by the profile HMM
- We may use the most probable alignment  $P(x, z^* \mid \mathcal{M})$  or the marginal probability  $P(x \mid \mathcal{M})$
- Another possibility is to use log-odds scores (*See Durbin 1998, chapter 5 for the details*)