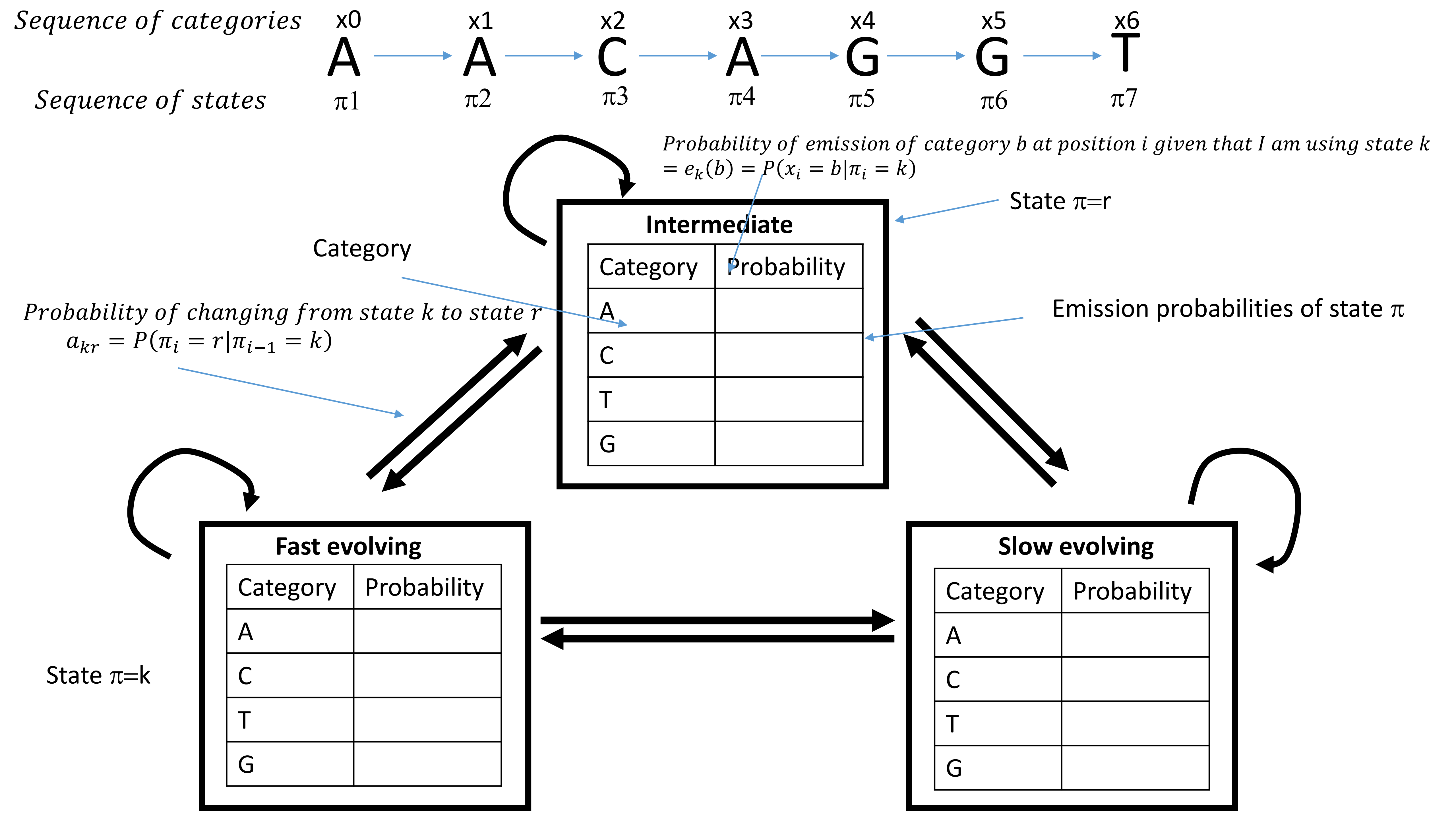
Practical Session 2

MM, HMM and sequences

REMEMBER!


Sequence of categories
x0 x1 x2 x3 x4 x5 x6
A A C A G G T
Sequence of states
π1 π2 π3 π4 π5 π6 π7
Probability of emission of category b at position i given that I am using state k = $e_k(b) = P(x_i = b|\pi_i = k)$
State π=r
Intermediate
Category
Probability
A
C
T
G
Category
Probability of changing from state k to state r $a_{kr} = P(\pi_i = r|\pi_{i-1} = k)$
Emission probabilities of state π
Fast evolving
Category
Probability
A
C
T
G
Slow evolving
Category
Probability
A
C
T
G
State π=k

NOTATION HMM

$State\ at\ position\ i = \pi_i$

$Sequence\ of\ states = \pi$

$Category\ at\ position\ i = x_i$

$Sequence\ of\ categories\ = x$

$$P(x, \pi) = a_{0\pi} \prod_{i}^{N} e_{\pi_i}(x_i) a_{\pi_i \pi_i + 1}$$

Initial state ("Probability that I started the chain at state π")

$$Probability\ of\ changing\ from\ state\ k\ to\ state\ r = a_{kr} = P(\pi_i = r|\pi_{i-1} = k)$$

I am using state *r*

Given that I was in the previous position using state *k*

Transmission matrix

	Fast Evolving	Intermediate Evolving	Slow Evolving
Fast Evolving			
Intermediate Evolving			
Slow Evolving			

Emission Probability matrix for each possible state

Fast evolving

Category	Probability
A	
C	
T	
G	

Intermediate

Category	Probability
A	
C	
T	
G	

Slow evolving

Category	Probability
A	
C	
T	
G	

$$Probability\ of\ emission\ of\ category\ b\ at\ position\ i\ given\ that\ I\ am\ using\ state\ k = e_k(b) = P(x_i = b|\pi_i = k)$$

HMM: Probability of x given model? Forward

$P(1,3,4,5,1,2,6,6,3,6,6,2,1,2,4,5|HMM)$

	1	3	4	…
F	$0.5*1/6$			
L	$0.5*1/10$			

x	$P(x\|F)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

$\boldsymbol{P(C|F) = 0.01}$

$\boldsymbol{P(F|C) = 0.1}$

x	$P(x\|C)$
1	$\frac{1}{10}$
2	$\frac{1}{10}$
3	$\frac{1}{10}$
4	$\frac{1}{10}$
5	$\frac{1}{10}$
6	$\frac{1}{2}$

HMM: Probability of x given model? Forward

P(1,3,4,5,1,2,6,6,3,6,6,2,1,2,4,5|HMM)

F,F

	1	3	4	…
F	0.5*1/6	0.5*1/6*0.99*1/6		
L	0.5*1/10			

x	$P(x\|F)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

$\boldsymbol{P(C|F) = 0.01}$

$\boldsymbol{P(F|C) = 0.1}$

x	$P(x\|C)$
1	$\frac{1}{10}$
2	$\frac{1}{10}$
3	$\frac{1}{10}$
4	$\frac{1}{10}$
5	$\frac{1}{10}$
6	$\frac{1}{2}$

HMM: Probability of x given model? Forward

P(1,3,4,5,1,2,6,6,3,6,6,2,1,2,4,5|HMM)

F,F

C,F

	1	3	4	…
F	0.5*1/6	0.5*1/6*0.99*1/6+ 0.5*1/10*0.1*1/6		
L	0.5*1/10			

x	$P(x\|F)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

$P(C|F) = 0.01$

$P(F|C) = 0.1$

x	$P(x\|C)$
1	$\frac{1}{10}$
2	$\frac{1}{10}$
3	$\frac{1}{10}$
4	$\frac{1}{10}$
5	$\frac{1}{10}$
6	$\frac{1}{2}$

HMM: Probability of x given model? Forward

P(1,3,4,5,1,2,6,6,3,6,6,2,1,2,4,5|HMM)

F,F

C,F

	1	3	4	…
F	0.5*1/6	0.5*1/6*0.99*1/6+ 0.5*1/10*0.1*1/6		
L	0.5*1/10	0.5*1/10*0.9*1/10		

C,C

x	$P(x\|F)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

$\boldsymbol{P(C|F) = 0.01}$

$\boldsymbol{P(F|C) = 0.1}$

x	$P(x\|C)$
1	$\frac{1}{10}$
2	$\frac{1}{10}$
3	$\frac{1}{10}$
4	$\frac{1}{10}$
5	$\frac{1}{10}$
6	$\frac{1}{2}$

HMM: Probability of x given model? Forward

P(1,3,4,5,1,2,6,6,3,6,6,2,1,2,4,5|HMM)

F,F

C,F

	1	3	4	…
F	0.5*1/6	0.5*1/6*0.99*1/6+ 0.5*1/10*0.1*1/6		
L	0.5*1/10	0.5*1/10*0.9*1/10 + 0.5*1/6*0.01*1/6		

C,C

F,C

x	$P(x\|F)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

$\boldsymbol{P(C|F) = 0.01}$

$\boldsymbol{P(F|C) = 0.1}$

x	$P(x\|C)$
1	$\frac{1}{10}$
2	$\frac{1}{10}$
3	$\frac{1}{10}$
4	$\frac{1}{10}$
5	$\frac{1}{10}$
6	$\frac{1}{2}$

We have searched all the combinations

HMM: Probability of x given model? Forward

P(1,3,4,5,1,2,6,6,3,6,6,2,1,2,4,5|HMM)

	1	3	4	...
F	0.5*1/6	0.5*1/6*0.99*1/6+ 0.5*1/10*0.1*1/6		
L	0.5*1/10	0.5*1/10*0.9*1/10 + 0.5*1/6*0.01*1/6		

$\left.\begin{array}{l}P1\\ \\ P2\end{array}\right\}$ P1+P2 =P(S|HMM)

HMM

- Scale the probability at each step (Rabiner 1989)

State	$\widehat{f_l}(i)$	$\widehat{f_l}(i+1)$
State1		
State2		
StateK		

Probability of moving from state r to state l in i+1

$$\widehat{f_l}(i+1) = \frac{1}{s_{i+1}} e_l(x_{i+1}) \sum_{r=1}^{K} \widehat{f_r}(i) a_{rl}$$

Scaled forward

Emission probability of category at xi+1 using state l

HMM

- Scale the probability at each step (Rabiner 1989)

State	$\hat{f}_l(i)$	$\hat{f}_l(i+1)$
State1		
State2		
StateK		

Total of the scaled value

If we impose that

$$\sum_l \hat{f}_l(i) = 1$$

Then si+1 is

$$s_{i+1} = \sum_l e_l(x_{i+1}) \sum_k \hat{f}_k(i) a_{kl}$$

$$P(X|HMM) = \prod_i^L s_i$$

Task 1

- Create in Python a class called HMM
- The constructor takes as input the transition probability matrix between states and a list of emission matrices, one for each state. Code these objects as a dictionary.

Task 1

```python
from AliasVose import RandomMultinomial

'''
A class to code methods that are used in HMM.
'''
class HiddenMarkovModel(object):

    '''
    Create an object HiddenMarkovModel with transition states and emission probabilities for each state
    Transition is a dictionary where, for each state, we count the probability to move to another state
    Emission is a dictionary where, for each state, we store the probabilities of each category
    '''
    def __init__(self, transition, emission):
        if transition.__len__()!=emission.__len__():
            raise Exception("for each state, we must have an emission probability vector, but found " + transition.__len__() + " " + emission.__len__())
        self.n = transition.__len__()
        self.transition = transition
        self.emission = emission
        '''
        dictionary to store the random multinomial
        '''
        self.random_transition = {}
        self.random_emission = {}
```

Task 2

- Implement a function that takes as input an observed sequence of categories and returns the log-likelihood. The scaled version is already implemented. Implement the unscaled version.

Task 3

- Implement a method that returns a sequence and the hidden states given the transition and emission probabilities of the HMM