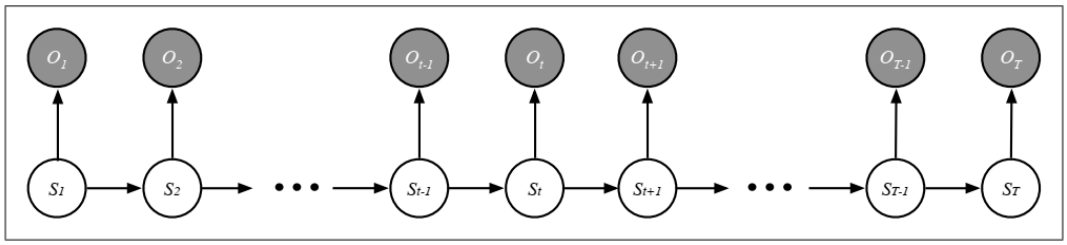
7.1 Viterbi algorithm



$$S_t \in \{1,2,\ldots.27\}$$

$$O_t \in \{0,1\}$$

$$\pi_i = P(S_1 = i) \qquad i \in \{1,2,\ldots.27\}$$

$$a_{ij} = P(S_{t+1} = j | S_t = i) \qquad i,j \in \{1,2,\ldots.27\}$$

$$b_{ik} = P(O_t = k | S_t = i) \qquad k \in \{0,1\} \qquad i \in \{1,2,\ldots.27\}$$

Define:

$$l^*_{it} = \max_{S_1,S_2\ldots S_{t-1}} logP(S_1, S_2 \ldots S_{t-1}, S_t = i, O_1, O_2, \ldots O_t)$$

t=1

$$l^*_{i1} = logP(S_1 = i, O_1) = logP(S_1 = i) + logP(O_1 | S_1 = i)$$

$$= \log \pi_i + \log b_i(O_1)$$

t>=1

$$l^*_{j,t+1} = \max_i \left(l^*_{it} + loga_{ij}\right) + \log b_j(O_{t+1})$$

One more matrix to record the maximum choice of state at t by $S_{t+1} = j$

$$\Phi_{t+1}(j) = argmax_i(l^*_{it} + loga_{ij})$$

Backtracking:

For t=T

$$s_T{}^* = argmax_i(l^*_{iT})$$

For 1<=t<=T-1

$$s_t^* = \Phi_{t+1}(s_{t+1}^*)$$

Programing result:

```
C:\Users\HP\AppData\Local\Programs\Python\Python37\python.exe D:/python/test/CSE250A07.py
total length:  430000
process indicator, how many has been processed
0
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000
130000
140000
150000
160000
170000
180000
190000
200000
```

```
220000
230000
240000
250000
260000
270000
280000
290000
300000
310000
320000
330000
340000
350000
360000
370000
380000
390000
400000
410000
420000
the hidden sentence is:
a house divided against itself canot stand
plot
```

(b)

Result sentence ignoring the repeated elements is" a house divided against itself canot stand". Because 2 "n" are considered as one mistakenly, it should be:

a house divided against itself cannot stand