Intrusion Detection System

- Signature based
- **Anomaly based**
- Host based
- **Network based**

Anomaly based Network Intrusion Detection System (A-NIDS)

- Statistical based
  - Univariate
  - Multivariate
- Knowledge based
- **Machine learning based**

Exploiting Communication Regularities

- Learn the normal sequences of messages on a network
- Build a model describing these sequences

Machine Learning

- Bayesian networks
- **Markov models**
- Neural networks
- Fuzzy logic
- Genetic algorithm
- Etc.

Hidden Markov Model

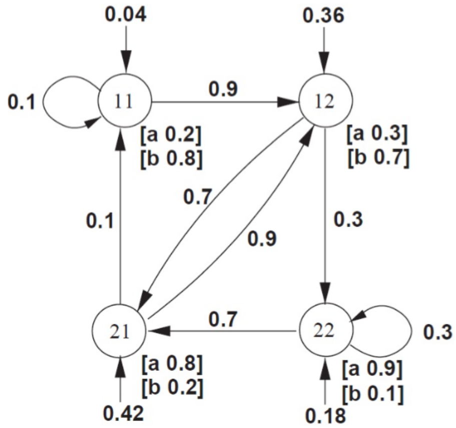


Figure : PAutomaC: a PFA/HMM Learning Competition, Sicco Verwer et al., 2012

Hidden Markov Model - Urn and Ball

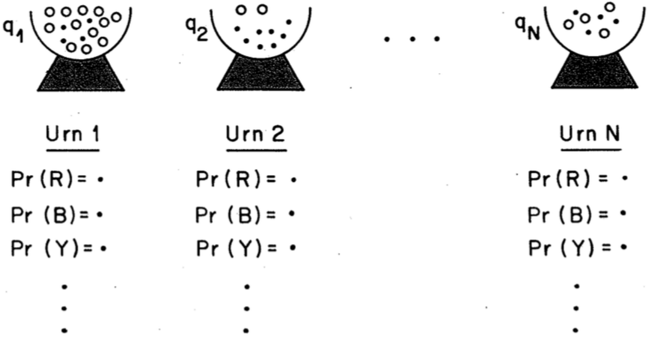
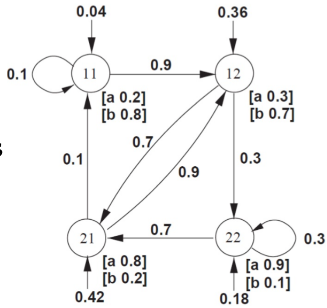


Figure : An Introduction to Hidden Markov Models, L. R. Rabiner B. H. juang, 1986

Hidden Markov Model



- T = length of observation sequence
- N = number of states in the model
- M = number of observation symbols
- $Q = \{q_1, q_2, ..., q_N\}$, states
- $V = \{v_1, v_2, ..., v_M\}$, observation symbols
- $A = \{a_{ij}\}$, $a_{ij} = Pr(q_j \text{ at } t+1 | q_i \text{ at } t)$, state transition probability distribution
- $B = \{b_j(k)\}$, $b_j(k) = Pr(v_k \text{ at } t | q_j \text{ at } t)$, observation symbol probability distribution
- $\pi = \{\pi_i\}$, $\pi_i = Pr(q_i \text{ at } t = 1)$, initial state distribution
- $\lambda = (A, B, \pi)$, the HMM

- Given a sequence of observations $O = O_1, O_2 \dots O_t$, the model moves through states $S = s_1, s_2 \dots s_t$
- Forward variable

$$\alpha_t(i) = Pr(O_1, O_2 \dots O_t, s_t = q_i | \lambda)$$

- Backward variable

$$\beta_t(i) = Pr(O_{t+1}, O_{t+2} \dots O_T | s_t = q_i, \lambda)$$

Forward variable

1. $\alpha_1(i) = \pi_i b_i(O_1)$ for $1 \leq i \leq N$
2. for $t = 1...T-1$ and $1 \leq i \leq N$

$$\alpha_{t+1}(i) = \left[\sum_{j=1}^{N} \alpha_t(j) a_{ji}\right] b_i(O_{t+1})$$

3. then $Pr(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$

Backward variable

1. $\beta_T(i) = 1$ for $1 \leq i \leq N$
2. for $t = T-1, T-2 \ldots 1$ and $1 \leq i \leq N$

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$$

- $\gamma_t(i) = Pr(s_t = q_i | O, \lambda)$

- a