1 Full joint parameters 1	2 Bayesian parameters 2	3 Conditionally independent 3
4 Prior probability 4	5 Conditional independence 5	6 Markov property 6
7 Markov chain 7	8 Absorbing state 8	9 Transient state 9
10 Ergodic chain 10	11 Ergodic transitions 11	12 Transient to transient 12
13 Transient to absorbing 13	14 Viterbi 14	15 Causal inference 15
16 Exact inference complexity 16		

1

$$2^N - 1$$

Full joint parameters

2

Number of nodes by 2^k, where k represents the number of incoming edges, giving us the total number of parameters needed

Bayesian parameters

3

We define X_1 to be conditionally independent of X_2 given X_3 if the probability of X_1 is not influenced by the value of X_2 when we have knowledge about X_3.:

$$\mathrm{P}(X_1|X_2, X_3) = \mathrm{P}(X_1|X_3)$$

Conditionally independent

4

A prior probability is a probability associated with a variable that has no incoming edges in a Bayesian network

Prior probability

5

Two sets of nodes, denoted as A and B, exhibit what is termed as conditional independence or are often referred to as being d-separated, given a set of nodes C if and only if all paths from A to B are effectively blocked by the presence of C

Conditional independence

6

The Markov property asserts that the transition from $X_{t-1} = x_i$ to $X_t = x_j$ depends only on the current state X_{t-1}:

$$\mathrm{P}\left(X_t|X_{t-1}, X_{t-2}, \ldots, X_0\right) =$$
$$= \mathrm{P}(X_t|X_{t-1})$$

Markov property

7

A discrete stochastic process is a first-order Markov chain when, for all t and for all N states, the following condition holds:

$$\mathrm{P}\left(X_t|X_{t-1}, X_{t-2}, \ldots, X_0\right) =$$
$$= \mathrm{P}(X_t|X_{t-1})$$

Markov chain

8

A state i is an absorbing state if $p_{ij} = 1$

Absorbing state

9

A state i is transient if there exists j reachable from i, but i is not reachable from j

Transient state

10

If all states in a Markov chain are recurrent, aperiodic, and communicate with each other, it is said to be ergodic

Ergodic chain

11

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} \cdot m_{kj}$$

Ergodic transitions

12

$$(I - Q)^{-1}$$

Transient to transient

13

$$(I - Q)^{-1} \cdot R$$

Transient to absorbing

14

Objective is to minimize:

$$-\log \mathrm{P}(X_0) +$$
$$\sum_{i=1:t} (-\log \mathrm{P}(X_i|X_{i-1}) - \log \mathrm{P}(e_i|X_i))$$

Viterbi

15

Causal inference is the process of identifying the independent and actual effect of a specific phenomenon within a larger system

Causal inference

16

$O(N|X_i|^k)$ where N is the number of nodes, X_i is the max arity, and k is the max variable connected to a factor

Exact inference complexity