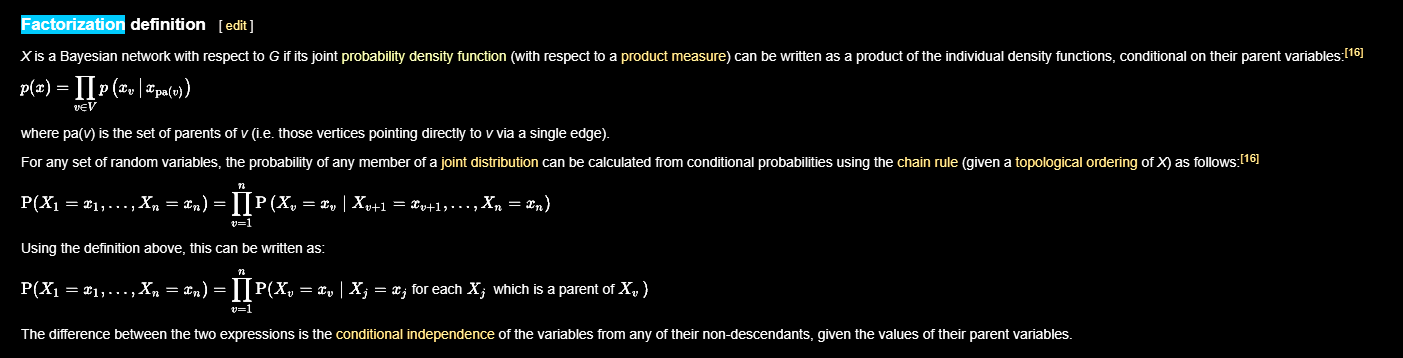
CS 344 – Guide 5

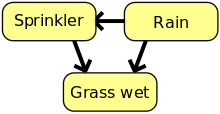
* Bayesian Networks:
  + Definition: a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (DAG).
  + Be familiar with the structure of Bayesian networks:
    - Nodes represent variables – observable quantities, latent variables, unknown parameters or hypotheses.
    - Edges represent conditional dependencies.
    - Unconnected nodes represent variables that are conditionally independent of each other.
    - Directed acyclic graphs:
      * A finite directed graph with no directed cycles.
      * Consists of finitely many vertices and edges such that there is no way to start at any vertex v and follow a consistently-directed sequence of edges that eventually loops back to v again.
    - Semantics (see “factorization definition” section)
      * Factorization definition:
        + X is a Bayesian network with respect to G if its joint probability density function (with respected to a product measure) can be written as a product of the individual density functions, conditional on their parent variables.
        + p(x) = ITP (xv I Xpa{v))

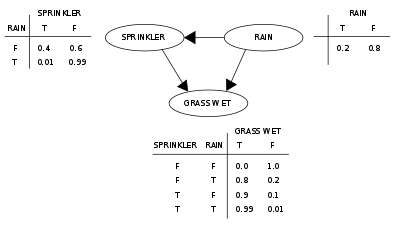
vEV

* + - * + where pa(v) is the set of parents of v (i.e. those vertices pointing directly to v via a single edge)



* + Compare and contrast the conditional probability tables in Bayes Networks with the full joint probability distribution.



* + E.g. see the rain-sprinkler example:
    - How many values must it store?
      * 14 values.
    - How many would the full joint have to store?
      * 2^n = 2^3 = 8
  + 
  + What benefit is there to using Bayesian networks as opposed to the probabilistic mechanism discussed in the previous unit?
    - Computationally far less expensive the more variables you have as the full joint requires n-dimensions for n-variables, leading to 2^n exponential run-time.
* Scikit Learn
  + ^\_^