BMI/CS 576 – Day 25

- Today
  - Sparse candidate algorithm
- Tuesday
  - ~20 min concluding remarks lecture (IN CLASS!)
  - Final notebook
    - Computing the best network structure
    - Computing measures of confidence in network structures

HW6

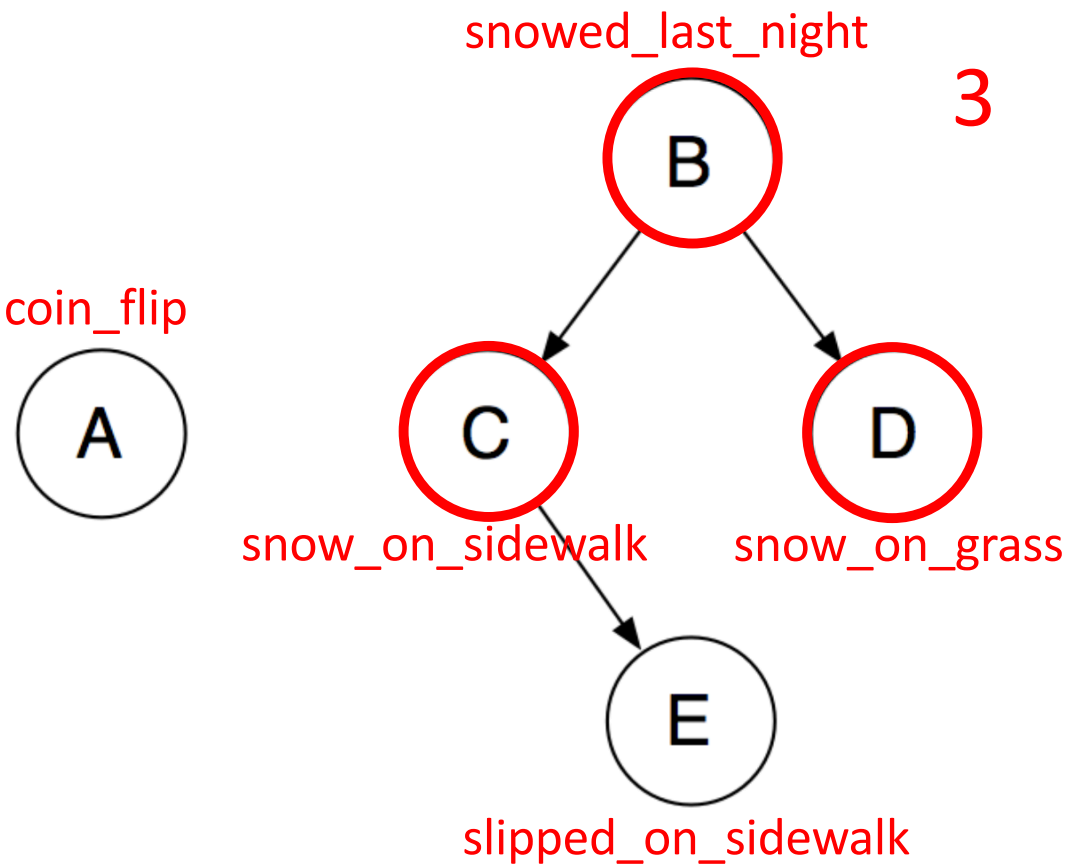
- Assigned by Saturday 12/7
- Due Monday 12/16
- NO LATE DAYS ALLOWED!
  - Solutions will need to be released on Tuesday 12/17, so that you have time to look at them before the final (on 12/18)
- Minimal to no programming required
- Topic: Networks

Course evaluation survey

- Please respond!
- Deadline: Dec 11th (Wed)
- Your feedback is critical to the future development of this course
- aefis.wisc.edu

Quiz

Suppose that the true Bayesian network relating the random variables A, B, C, D, and E is the one shown below. What is the maximum number of variables that could have positive (i.e., non-zero) mutual information with variable E?



Questions – Network scores

$$\text{score}(G:D) = \log \Pr(G \mid D)$$

$$= \log \Pr(D \mid G) + \log \Pr(G) + C$$

constant (depends on D)

log probability of data D given graph G

log prior probability of graph G

Treating the last two terms as constants, we get:

$$\text{score}(G:D) = \sum_i Score(X_i, Parents(X_i) : D)$$

Each term is the logarithm of the probability of the data at that node (X_i) and is negative

Questions

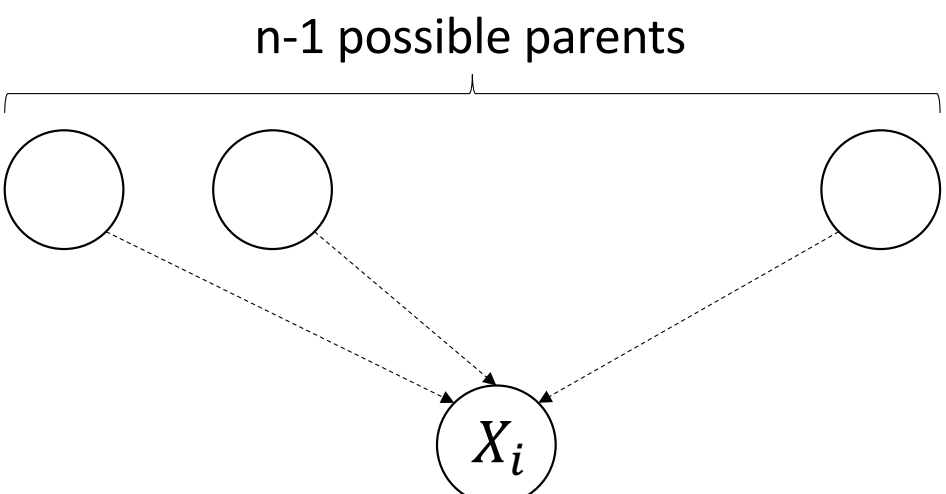
- How does the network score penalize more complex models?
  - Integrating out the model parameters (to compute the model evidence) results in penalizing models with more parameters
  - We will explore a simple example of this in HW6

Questions

- How do we get the joint probability between two variables from the data?
  - We will refer to this as the “empirical” distribution
  - Recall that we implemented a function `estimate_joint_dist` in the Day 10 notebook
  - In today’s notebook we will compute marginal distributions of pairs of variables from joint distributions

Questions

- Without the Sparse Candidate Algorithm, why is the number of possible parent sets for a variable $O(2^n)$?



2 possibilities (is/is not parent) for each possible parent $\rightarrow$ $2^{n-1} \rightarrow O(2^n)$