CS6462

Probabilistic and Explainable AI

Lesson 25

Causal Bayesian Networks

by Emil Vassev

March 24, 2025

Causal Networks and Interventions

Parent-child relationship :

- stable and autonomous
- can be changed locally

Revision: Structural Causal Model

- $X_i = f_i(PA_i, U_i)$
- *interventions*: modify subset of X_i - change U_i or set f_i to a constant
- changing (or intervening upon) one mechanism $p(X_i|PA_i)$ does not change the other mechanisms $p\left(X_j|PA_j\right),\ i \neq j$

Bayesian network model: joint distribution tells how probable events are and how probability would change with subsequent observations

Causal model: Bayesian network model + tells how probabilities would change as a result of interventions (such changes cannot be deduced from a joint distribution)

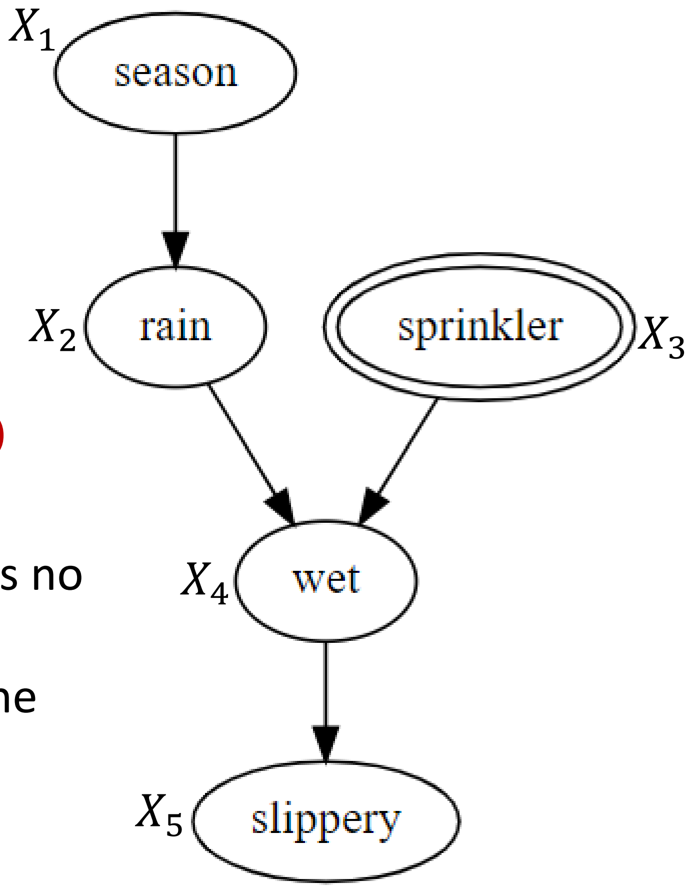
Bayesian Network modularity and interventions:

- specify the immediate change implied by an intervention – the change is local, because the parent-child relationship is autonomous
- an intervention can be predicted by modifying the corresponding factors and using the modified product to compute a new probability function

Causal Networks and Interventions (cont.)

Bayesian Network and interventions: Example

- consider the sprinkler example: $X_3 = \{on, off\}$
- action: "turning the sprinkler on"
- result*:*
  - remove the link $X_1 \rightarrow X_3$
  - $X_3 = on$
  - $P_{X_3=on}(X_1, X_2, X_4, X_5) = P(X_1) * P(X_2|X_1) * P(X_4|X_2, X_3 = on) * P(X_5|X_4)$
- conclusion:
  - removal of $P(X_3|X_1)$ - the prior relationship between season and sprinkler is no longer in effect while we perform the action
  - a new mechanism (in which the season is ignored) determines the state of the sprinkler
- difference between the action $do(X_3 = on)$ and the observation $X_3 = on$:
  - effect of observation $X_3 = \text{on}$ is obtained by ordinary Bayesian conditioning $P(X_1, X_2, X_4, X_5, X_3 = on)$
  - effect of action $do(X_3 = on)$ is conditioning a graph, with the link $X_1 \rightarrow X_3$ removed



Causal Bayesian Networks - Definition

Mathematical Definition:

- $X = \{X_1, X_2, \dots, X_n\}$ – set of observables (modelled as random variables)
- G – DAG, X associated with the vertices of G
- $P(X)$ – probability distribution on that set X of variables
- X_i – variables subject to interventions, $X_i \subseteq X$
- $P_{X_i}(X)$ – denotes the distribution resulting from the intervention $do(X_i = x)$ that sets a subset X_i of variables to constants x
- P_* – set of all interventional distributions $P_{X_i}(X)$, $X_i \subseteq X$, including $P(X)$, which represents no intervention (i.e.,$X_i \equiv \emptyset$)
- DAG G is said to be a Causal Bayesian Network, compatible with P_* iff $\forall P_{X_i} \in P_*$ following conditions hold:
  1) $P_{X_i}(X)$ is Markov relative to G
  2) $P_{X_i}(X_{i,j}) = 1$ for $\forall X_{i,j} \in X_i$, $X_i = \{X_{i,1}, X_{i,2}, \dots, X_{i,k}\}$, whenever $X_{i,j}$ is consistent with $X_i = x$
  3) $P_{X_i}(X_l|PA_l) = P(X_l|PA_l)$ for $\forall X_l \notin X_i$, whenever PA_l is consistent with $X_i = x$

Causal Bayesian Networks – Definition (cont.)

Mathematical Definition (cont.) :

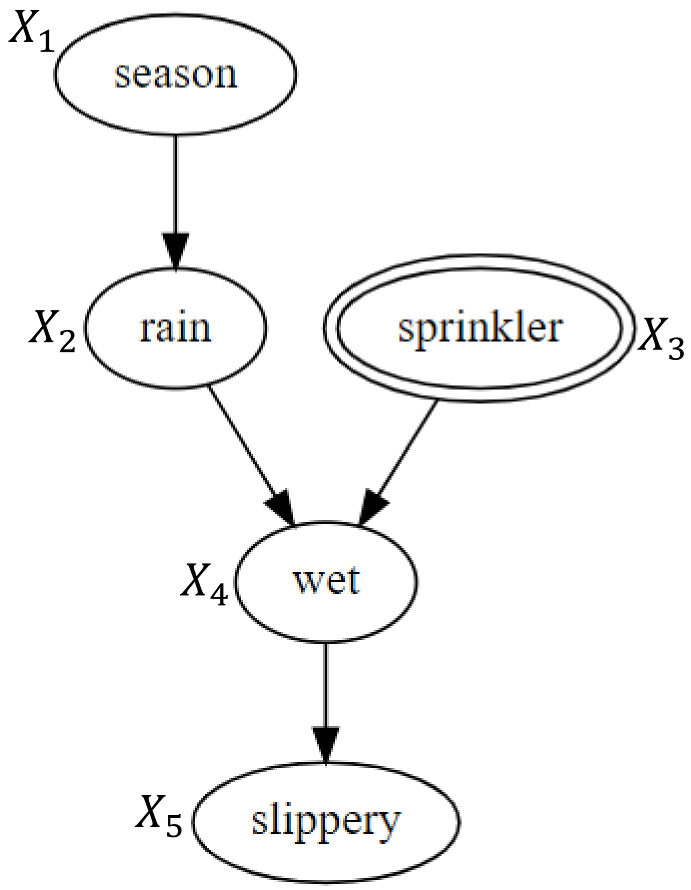
- the mathematical definition of causal Bayesian Networks imposes constraints on P_*
- these constraints enable us to compute the distribution $P_{X_i}(X)$, resulting from any intervention $do(X_i = x)$ as a truncated factorization:

$X = \{X_1, X_2, \dots, X_n\}$ – set of random variables

X_i – variables subject to interventions, $X_i \subseteq X$

$$P_{X_i}(X) = \prod_{l=1}^{n} P(X_l|PA_l) \text{ for } \forall X_l \notin X_i$$

This definition justifies the removal of the parent-child links of the X_i subject to the intervention $do(X_i = x)$ on G.



Causal Bayesian Networks – Properties

Property 1:

$P(X_i|PA_i) = P_{PA_i}(X_i)$, for $\forall X_i$ variables subject to interventions, $PA_i \in X_i \subseteq X$

- renders every parent set PA_i exogenous relative to its child X_i ensuring that the conditional probability $P(X_i|PA_i)$ coincides with the effect of setting PA_i as variables subject to interventions

Property 2:

$P_{PA_{i,S}}(X_i) = P_{PA_i}(X_i)$, for $\forall X_i$ and for every subset S of variables disjoint of $\{X_i, PA_i\}$

- expresses the notion of invariance – if the direct causes PA_i are subject to intervention, other interventions will not affect the probability of X_i

Summary

Causal Bayesian Networks

Causal Networks and Interventions

Causal Bayesian Networks – Definition

Causal Bayesian Networks – Properties

Next Lesson:

- Explainability in the context of AI

Thank You!

Questions?