KRR with Uncertainty

Bayesian Networks

Ira A. Fulton Schools of Engineering
Arizona State University

Objectives

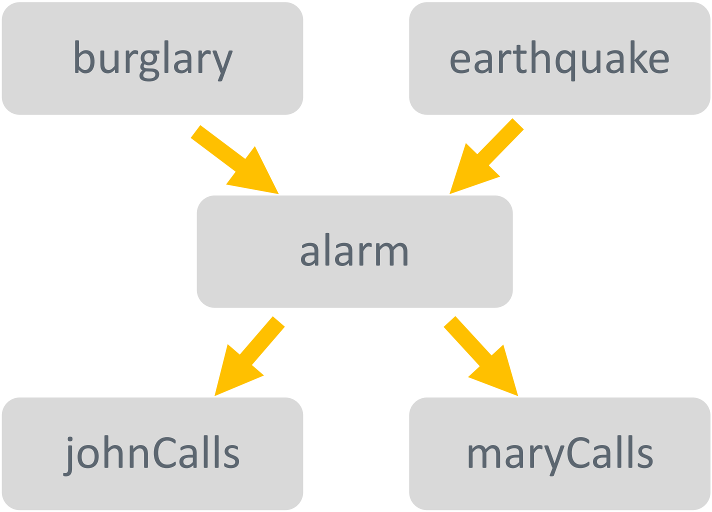
Objective

Explain the concept of Bayesian networks and how Bayesian networks utilize conditional independence

Bayesian Network

- An augmented, directed acyclic graph, where each node corresponds to a random variable X_i and each edge indicates a direct influence among the random variables.

- The influence for each variable X_i is quantified with a conditional probability distribution (CPD) P(X_i | Pa(X_i)), where Pa(X_i) are the parents of X_i in the graph.



Qualitative component: DAG

$\mathbf{P}(burglary)$
$(0.001, 0.999)$

$\mathbf{P}(earthquake)$
$(0.002, 0.998)$

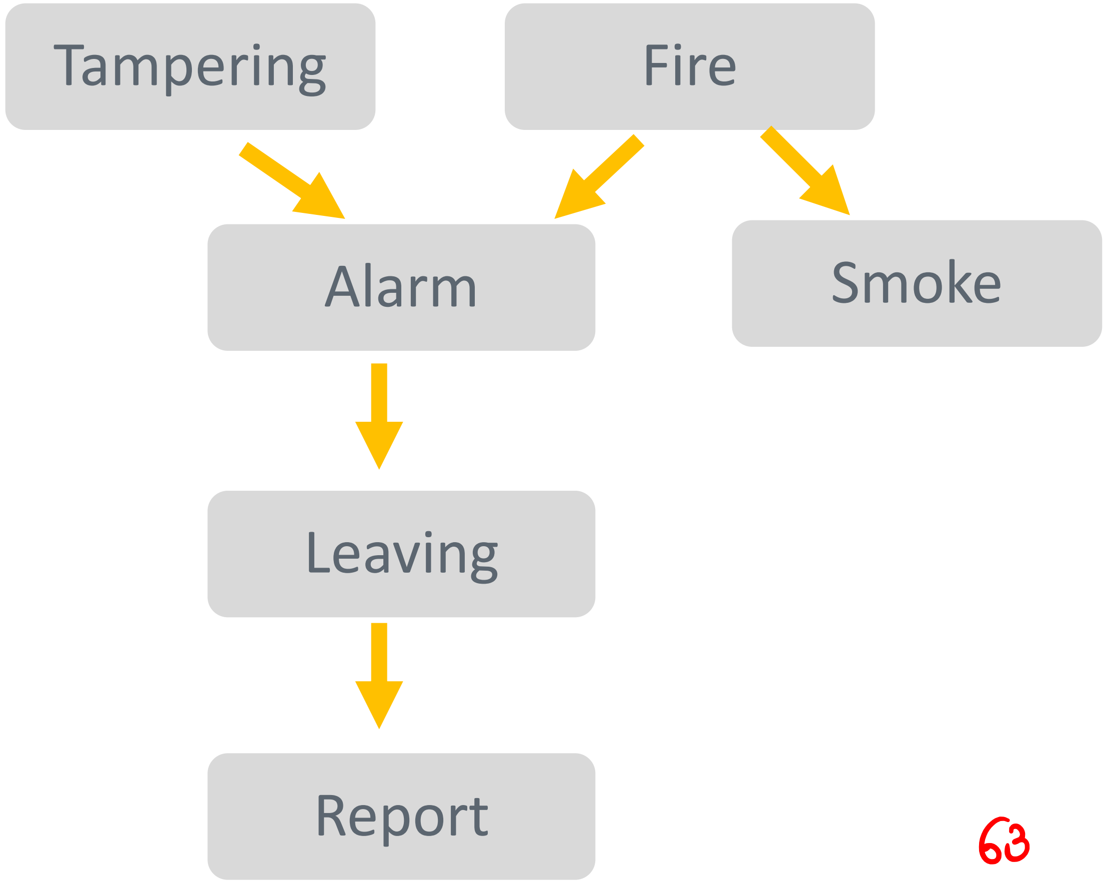
burglary	*earthquake*	$\mathbf{P}(alarm \mid burglary, earthquake)$
true	*true*	$(0.95, 0.05)$
true	*false*	$(0.95, 0.05)$
false	*true*	$(0.29, 0.71)$
false	*false*	$(0.001, 0.999)$

alarm	$\mathbf{P}(johncalls \mid alarm)$
true	$(0.90, 0.10)$
false	$(0.05, 0.95)$

alarm	$\mathbf{P}(marycalls \mid alarm)$
true	$(0.70, 0.30)$
false	$(0.01, 0.99)$

Quantitative component: CPD

Example BN: Fire Diagnosis

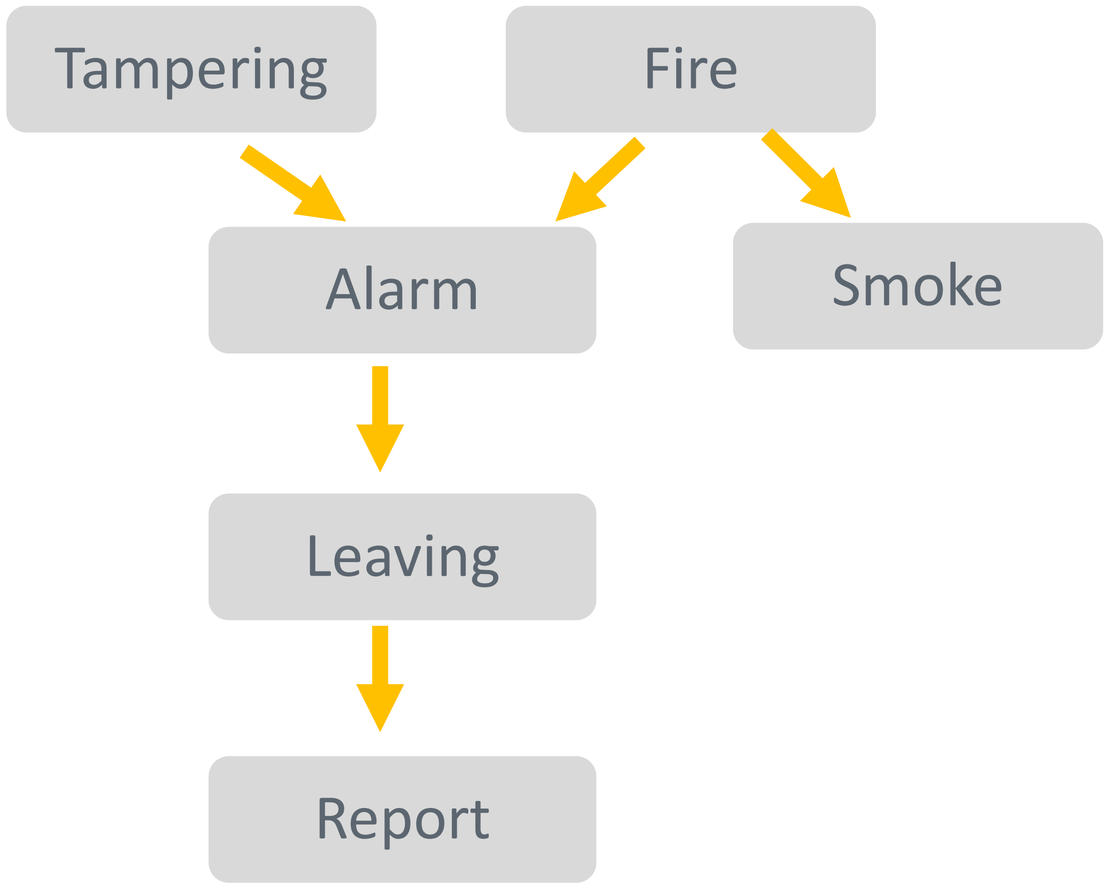


63

- All variables are Boolean.
- **Q:** How many probabilities do we need to specify for this Bayesian network?
  - Take into account that probability tables have to sum to *1*.

Ex. for BN construction: Fire Diagnosis (1 of 3)

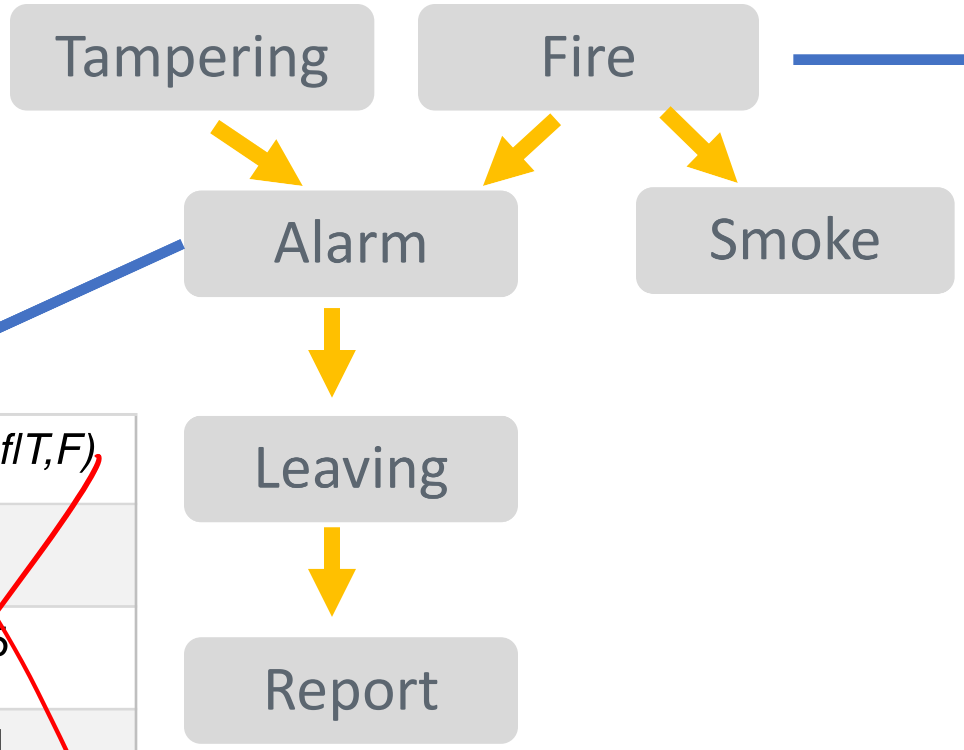
P(Tampering=t)	~~P(Tampering=f)~~
0.02	~~0.98~~



We don’t need to store P(Tampering=f) since probabilities sum to 1

Ex. for BN construction: Fire Diagnosis (2 of 3)

P(Tampering=t)
0.02



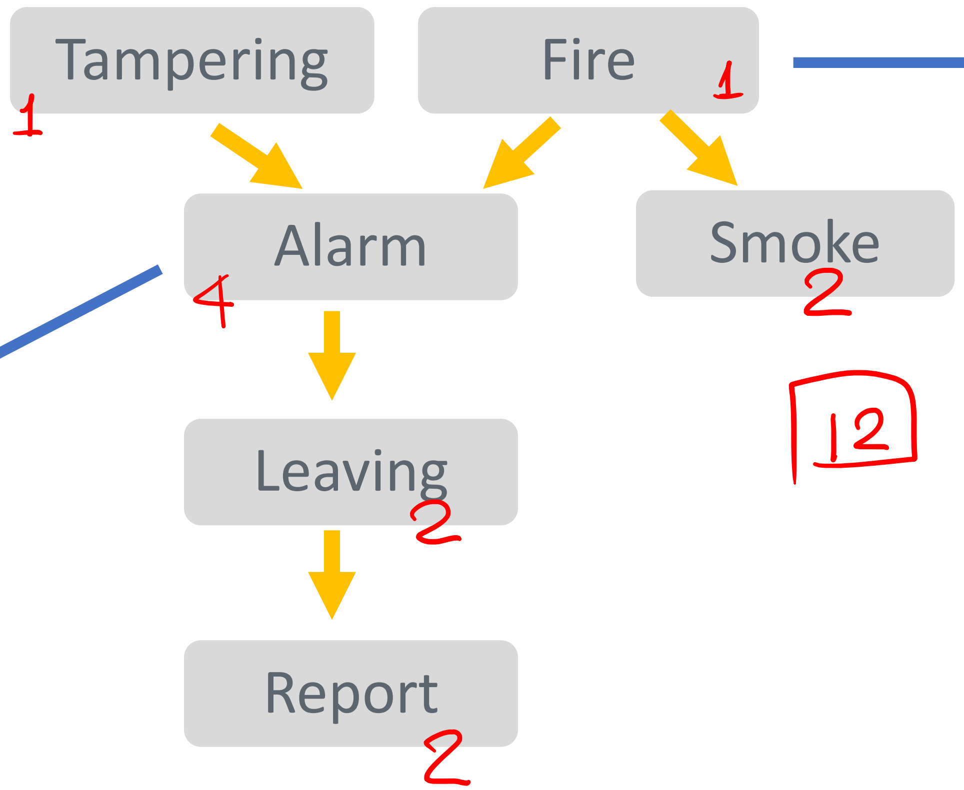
P(Fire=t)
0.01

Tampering T	*Fire F*	*P(Alarm=t|T,F)*	~~*P(Alarm=f|T,F)*~~
t	t	0.5	~~0.5~~
t	f	0.85	~~0.15~~
f	t	0.99	~~0.01~~
f	f	0.0001	~~0.9999~~

Each row of this table is a conditional probability distribution.

Ex. for BN construction: Fire Diagnosis (3 of 3)

P(Tampering=t)
0.02

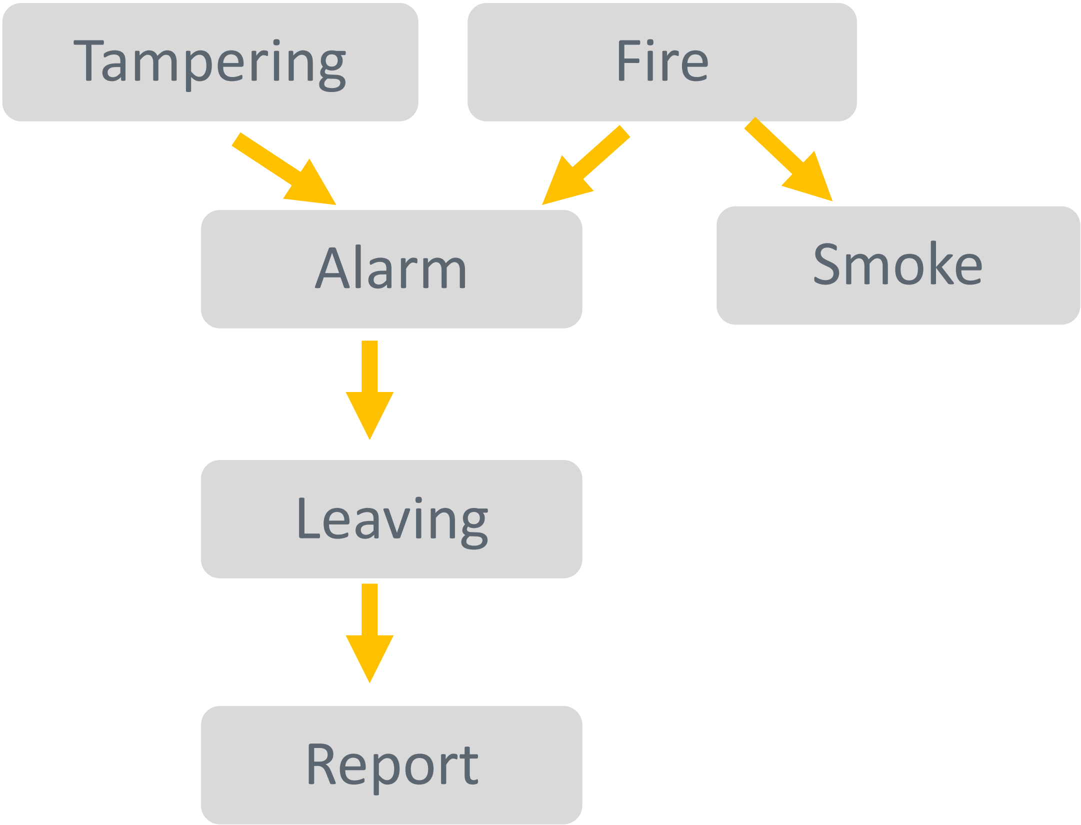


P(TFire=t)
0.01

Tampering T	*Fire F*	*P(Alarm=t|T,F)*
t	t	0.5
t	f	0.85
f	t	0.99
f	f	0.0001

Independence Assumption in Bayesian Networks

- **The Bayesian network represents a set of independence assumptions:**
- **Each variable is independent of its non-descendents given its parents** $(X_i \perp NonDescentdant(X_i) \mid Pa(X_i))$



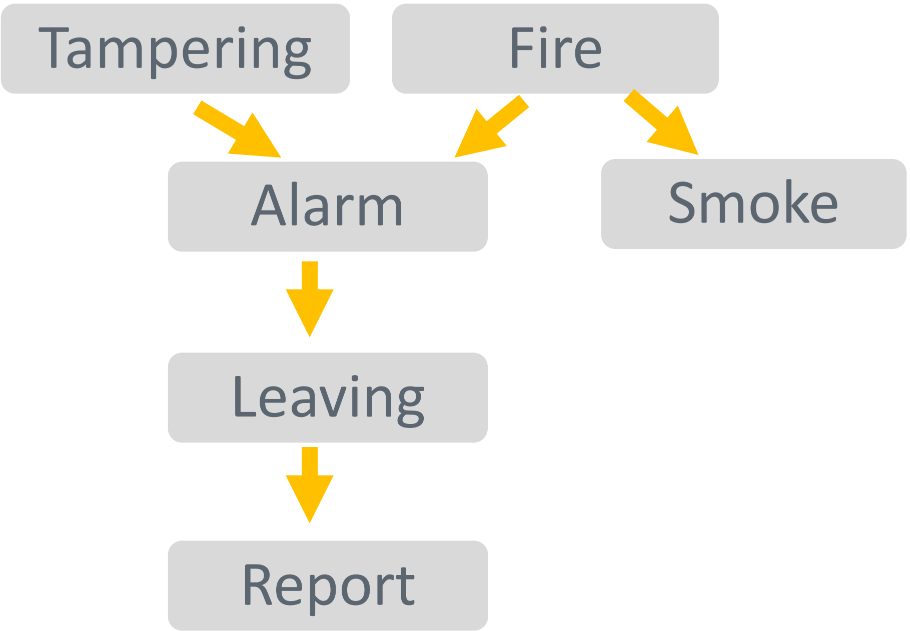
Compact Representation due to Independence Assumption

By the chain rule

$$P(T, F, A, S, L, R) = P(T)\, P(F|\cancel{T})\, P(A|\, T, F)\, P(S \mid \cancel{T}, F, \cancel{A})\, P(L \mid \cancel{T}, \cancel{F}, A, \cancel{S})\, P(R \mid \cancel{T}, \cancel{F}, \cancel{A}, \cancel{S}, L)$$

BN assumption: **each variable is independent of its non-descendents given its parents**

$$P(T, F, A, S, L, R) = P(T)\, P(F \underline{\quad})\, P(A|\, T, F)\, P(S \mid F)\, P(L \mid A)\, P(R \mid L)$$



$$\mathbf{P}(X_1, \ldots, X_n) = \prod_{i=1}^{n} \mathbf{P}(X_i \mid \mathbf{Pa}(X_i))$$

Example

$P(Tampering{=}t, Fire{=}f, Alarm{=}t, Smoke{=}f, Leaving{=}t, Report{=}t) =$

$$P(Tampering{=}t) \times P(Fire{=}f) \times P(Alarm{=}t \mid Tampering{=}t, Fire{=}f) \times$$

$$\times P(Smoke{=}f \mid Fire{=}f) \times P(Leaving{=}t \mid Alarm{=}t)$$

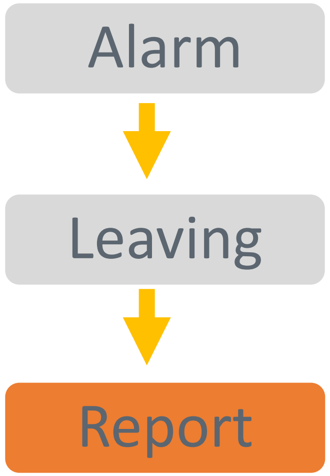
$$\times P(Report{=}t \mid Leaving{=}t) =$$

$$0.2 \text{ x } (1\text{-}0.01) \text{ x } 0.85 \text{ x } (1\text{-}0.01) \text{ x } 0.88 \text{ x } 0.75 = 0.126$$

Bayesian Networks: Types of Inference

Diagnostic

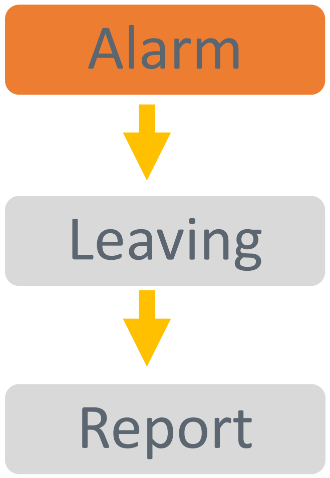
P(A | R=t) ?



$$\frac{P(A=t \wedge R=t)}{P(R=t)}$$

Predictive

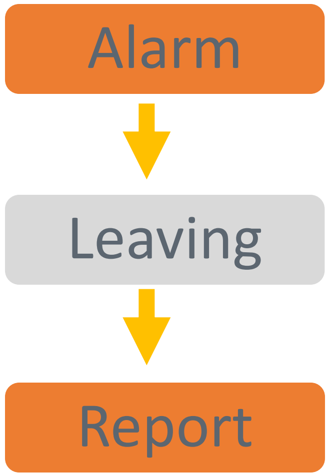
P(R | A=t) ?



$$P(A=t \wedge R=t) = \sum_{L=v} P(A=t \wedge L=v \wedge R=t)$$

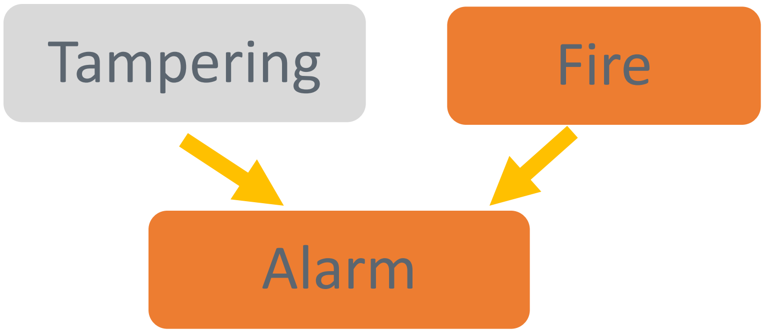
Mixed

P(L | A=f, R=t)?



Intercausal

P(T | A=t, F=t)



Wrap-Up