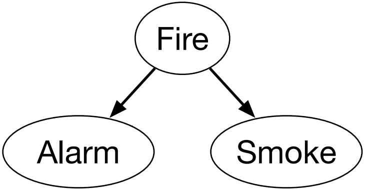
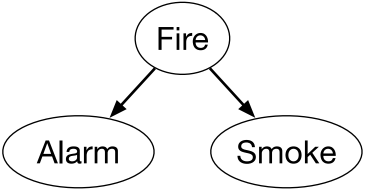
Understanding Independence: Common ancestors



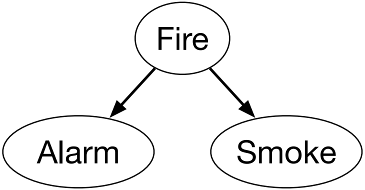
- *Alarm* and *Smoke* are
  given {}

Understanding Independence: Common ancestors



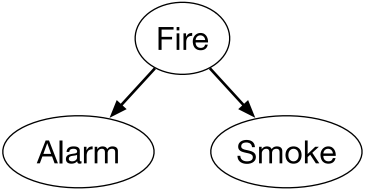
- *Alarm* and *Smoke* are dependent given $\{\}$

Understanding Independence: Common ancestors



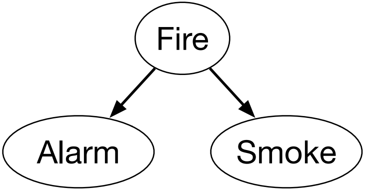
- *Alarm* and *Smoke* are dependent given {}
- *Alarm* and *Smoke* are given *Fire*

Understanding Independence: Common ancestors



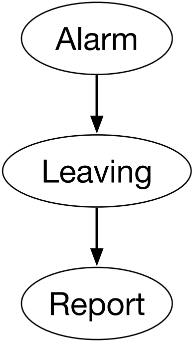
- *Alarm* and *Smoke* are dependent given {}
- *Alarm* and *Smoke* are independent given *Fire*

Understanding Independence: Common ancestors



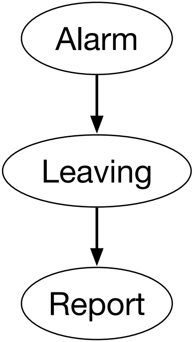
- *Alarm* and *Smoke* are dependent given $\{\}$
- *Alarm* and *Smoke* are independent given *Fire*
- Intuitively, *Fire* can explain *Alarm* and *Smoke*; learning one can affect the other by changing the belief in *Fire*.

Understanding Independence: Chain



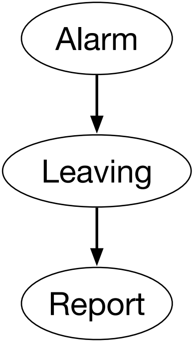
- *Alarm* and *Report* are
  given {}

Understanding Independence: Chain



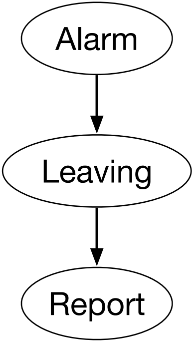
- *Alarm* and *Report* are dependent given $\{\}$

Understanding Independence: Chain



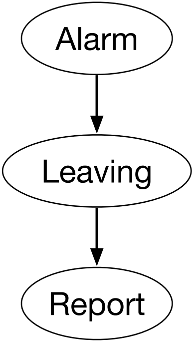
- *Alarm* and *Report* are dependent given $\{\}$
- *Alarm* and *Report* are given *Leaving*

Understanding Independence: Chain



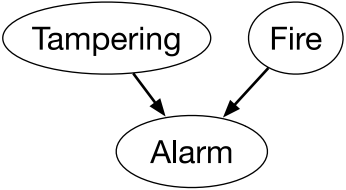
- *Alarm* and *Report* are dependent given $\{\}$
- *Alarm* and *Report* are independent given *Leaving*

Understanding Independence: Chain



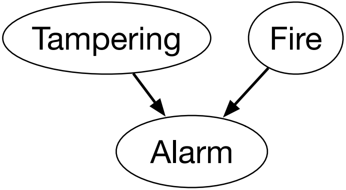
- *Alarm* and *Report* are dependent given {}
- *Alarm* and *Report* are independent given *Leaving*
- The (only) way that the *Alarm* affects *Report* is by affecting *Leaving*.

Understanding Independence: Common descendants



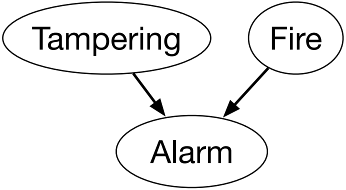
- *Tampering* and *Fire* are
  given {}

Understanding Independence: Common descendants



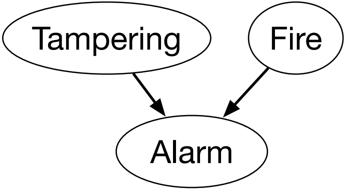
- *Tampering* and *Fire* are independent given {}

Understanding Independence: Common descendants



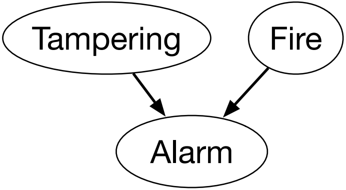
- *Tampering* and *Fire* are independent given {}
- *Tampering* and *Fire* are given *Alarm*

Understanding Independence: Common descendants



- *Tampering* and *Fire* are independent given {}
- *Tampering* and *Fire* are dependent given *Alarm*

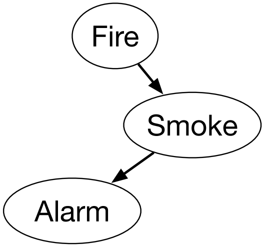
Understanding Independence: Common descendants



- *Tampering* and *Fire* are independent given $\{\}$
- *Tampering* and *Fire* are dependent given *Alarm*
- Intuitively, *Tampering* can explain away *Fire*

Clicker Question

For the belief network, and the ordering *Fire*, *Smoke*, *Alarm*



A *Alarm* is independent of *Smoke* given *Fire*

B *Alarm* is independent of *Fire* given *Smoke*

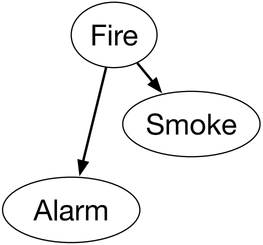
C *Alarm* is independent of *Fire* given $\{\}$

D All of the above independencies hold

E There are no independencies

Clicker Question

For the belief network, and the ordering *Fire*, *Smoke*, *Alarm*



A *Alarm* is independent of *Smoke* given *Fire*

B *Alarm* is independent of *Fire* given *Smoke*

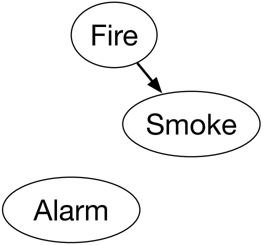
C *Alarm* is independent of *Smoke* given $\{\}$

D All of the above independencies hold

E There are no independencies

Clicker Question

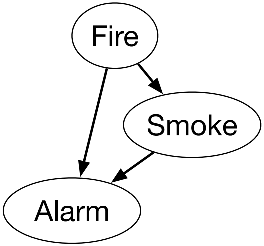
For the belief network, and the ordering *Fire*, *Smoke*, *Alarm*



- A *Alarm* is independent of *Smoke* given *Fire*
- B *Alarm* is independent of *Fire* given *Smoke*
- C *Alarm* is independent of *Fire* given {}
- D All of the above independencies hold
- E There are no independencies

Clicker Question

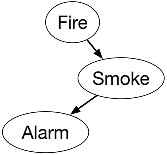
For the belief network, and the ordering *Fire*, *Smoke*, *Alarm*



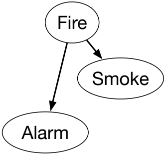
- A *Alarm* is independent of *Smoke* given *Fire*
- B *Alarm* is independent of *Fire* given *Smoke*
- C *Alarm* is independent of *Fire* given {}
- D All of the above independencies hold
- E There are no independencies

Clicker Question

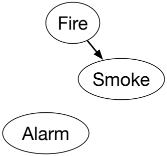
Which network best fits a fire alarm that only detects the heat of the fire?



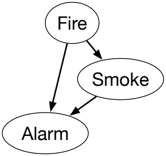
A



B



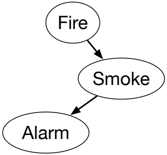
C



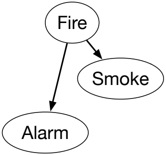
D

Clicker Question

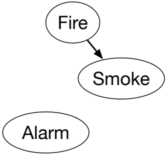
Which network best fits a smoke alarm (that only detects smoke)?



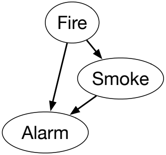
A



B



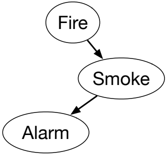
C



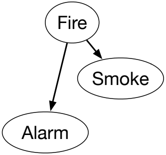
D

Clicker Question

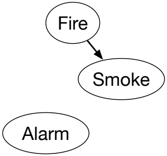
Which network best fits a fire alarm that detects both smoke and the heat of the fire?



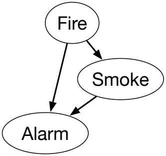
A



B



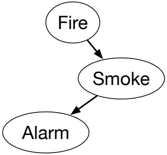
C



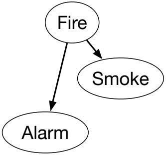
D

Clicker Question

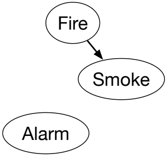
Which network best fits a burglary alarm that doesn't detect heat or smoke?



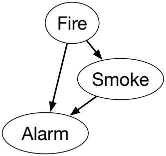
A



B

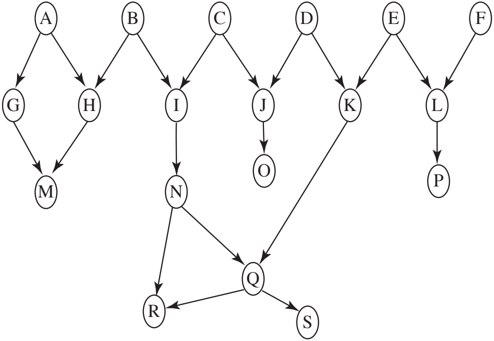


C



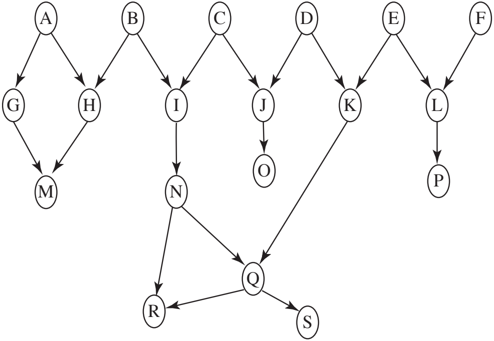
D

Understanding independence: example



(a) On which given probabilities does $P(N)$ depend?

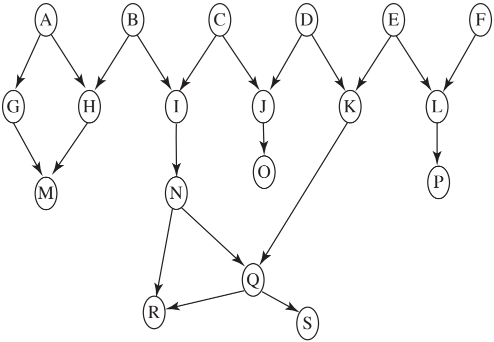
Understanding independence: example



(a) On which given probabilities does $P(N)$ depend?

(b) If you were to observe a value for B, which variables' probabilities will change?

Understanding independence: example

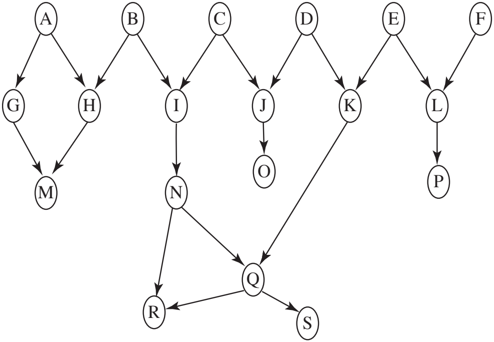


(a) On which given probabilities does $P(N)$ depend?

(b) If you were to observe a value for B, which variables' probabilities will change?

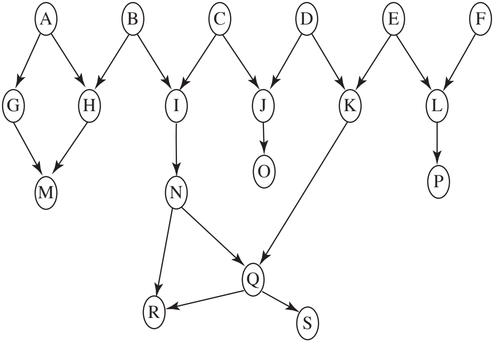
(c) If you were to observe a value for N, which variables' probabilities will change?

Understanding independence: questions



(d) Suppose you had observed a value for M; if you were to then observe a value for N, which variables' probabilities will change?

Understanding independence: questions



(d) Suppose you had observed a value for M; if you were to then observe a value for N, which variables' probabilities will change?

(e) Suppose you had observed B and Q; which variables' probabilities will change when you observe N?

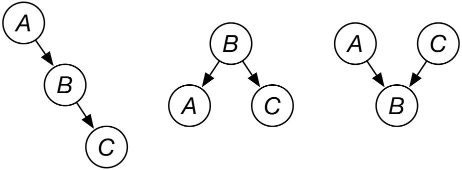
What variables are affected by observing?

- If you observe variable(s) $\overline{Y}$, the variables whose posterior probability is different from their prior are:
  - ancestors of $\overline{Y}$ and
  - their descendants.

What variables are affected by observing?

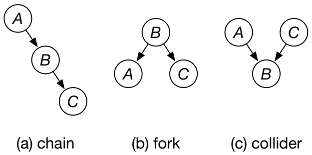
- If you observe variable(s) $\overline{Y}$, the variables whose posterior probability is different from their prior are:
  - ancestors of $\overline{Y}$ and
  - their descendants.
- Intuitively (assuming network ordered so causes are before effects):
  - You do abduction to possible causes and
  - prediction from the causes.

Three types of meetings between arcs



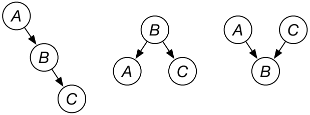
(a) chain (b) fork (c) collider

D-separation



- A path p can follow arrows in either direction.
- Observations Zs block a path p if:
  - (a) p contains a chain $A \rightarrow B \rightarrow C$, and $B \in Zs$
  - (b) p contains a fork $A \leftarrow B \rightarrow C$, and $B \in Zs$
  - (c) p contains a collider $A \rightarrow B \leftarrow C$, and B, and all its descendants, are not in Zs

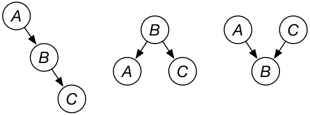
D-separation



(a) chain (b) fork (c) collider

- A path p can follow arrows in either direction.
- Observations Zs block a path p if:
  - (a) p contains a chain $A \rightarrow B \rightarrow C$, and $B \in Zs$
  - (b) p contains a fork $A \leftarrow B \rightarrow C$, and $B \in Zs$
  - (c) p contains a collider $A \rightarrow B \leftarrow C$, and B, and all its descendants, are not in Zs
- X is d-separated from Y given Zs if every path between X and Y is blocked by Zs

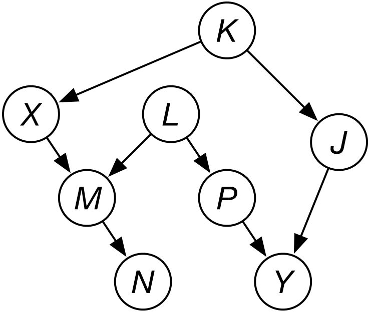
D-separation



(a) chain (b) fork (c) collider

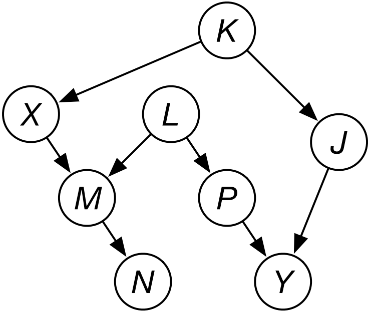
- A path p can follow arrows in either direction.
- Observations Zs block a path p if:
  (a) p contains a chain $A \rightarrow B \rightarrow C$, and $B \in Zs$
  (b) p contains a fork $A \leftarrow B \rightarrow C$, and $B \in Zs$
  (c) p contains a collider $A \rightarrow B \leftarrow C$, and B, and all its descendants, are not in Zs
- X is d-separated from Y given Zs if every path between X and Y is blocked by Zs
- X is independent Y given Zs for all conditional probabilities iff X is d-separated from Y given Zs

Example



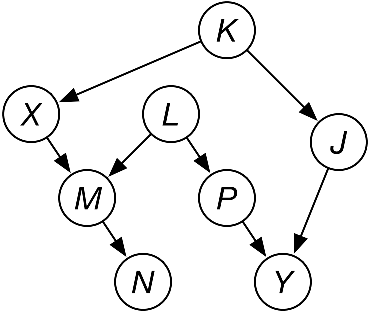
- Are X and Y d-separated by $\{\}$?

Example



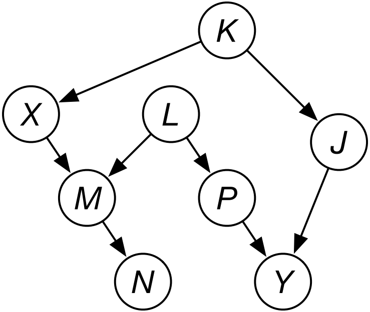
- Are X and Y d-separated by $\{\}$?
- Are X and Y d-separated by $\{K\}$?

Example



- Are X and Y d-separated by $\{\}$?
- Are X and Y d-separated by $\{K\}$?
- Are X and Y d-separated by $\{K, N\}$?

Example



- Are X and Y d-separated by $\{\}$?
- Are X and Y d-separated by $\{K\}$?
- Are X and Y d-separated by $\{K, N\}$?
- Are X and Y d-separated by $\{K, N, P\}$?

Markov Random Field

A Markov random field is composed of

- of a set of random variables: $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ and
- a set of factors $\{f_1, \ldots, f_m\}$, where a factor is a non-negative function of a subset of the variables.

Markov Random Field

A Markov random field is composed of

- of a set of random variables: $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ and
- a set of factors $\{f_1, \ldots, f_m\}$, where a factor is a non-negative function of a subset of the variables.

and defines a joint probability distribution:

$$P(\mathbf{X} = \mathbf{x}) \propto \prod_k f_k(\mathbf{X}_k = \mathbf{x}_k) .$$

Markov Random Field

A Markov random field is composed of

- of a set of random variables: $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ and
- a set of factors $\{f_1, \ldots, f_m\}$, where a factor is a non-negative function of a subset of the variables.

and defines a joint probability distribution:

$$P(\mathbf{X} = \mathbf{x}) \propto \prod_k f_k(\mathbf{X}_k = \mathbf{x}_k) .$$

$$P(\mathbf{X} = \mathbf{x}) = \frac{1}{Z} \prod_k f_k(\mathbf{X}_k = \mathbf{x}_k) .$$

$$Z = \sum_{\mathbf{x}} \prod_k f_k(\mathbf{X}_k = \mathbf{x}_k)$$

where $f_k(\mathbf{X}_k)$ is a factor on $\mathbf{X}_k \subseteq \mathbf{X}$, and
$\mathbf{x}_k$ is $\mathbf{x}$ projected onto $\mathbf{X}_k$.
Z is a normalization constant known as the partition function.

Markov Networks and Factor graphs

- A factor graph is a bipartite graph, which contains a variable node for each random variable and a factor node for each factor. There is an edge between a variable node and a factor node if the variable appears in the factor.

Markov Networks and Factor graphs

- A factor graph is a bipartite graph, which contains a variable node for each random variable and a factor node for each factor. There is an edge between a variable node and a factor node if the variable appears in the factor.
- A Markov network is a graphical representation of a Markov random field where the nodes are the random variables and there is an arc between any two variables that are in a factor together.

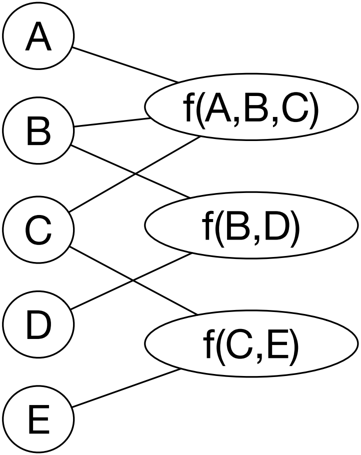
Markov Networks and Factor graphs

- A factor graph is a bipartite graph, which contains a variable node for each random variable and a factor node for each factor. There is an edge between a variable node and a factor node if the variable appears in the factor.
- A Markov network is a graphical representation of a Markov random field where the nodes are the random variables and there is an arc between any two variables that are in a factor together.
- A belief network is a type of Markov random field where

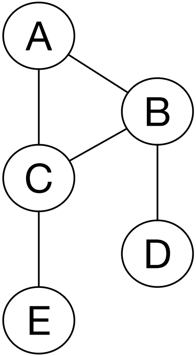
Markov Networks and Factor graphs

- A factor graph is a bipartite graph, which contains a variable node for each random variable and a factor node for each factor. There is an edge between a variable node and a factor node if the variable appears in the factor.
- A Markov network is a graphical representation of a Markov random field where the nodes are the random variables and there is an arc between any two variables that are in a factor together.
- A belief network is a type of Markov random field where the factors represent conditional probabilities, there is a factor for each variable, and directed graph is acyclic.

Factor graph and Markov network example



Factor Graph



Markov Network

Independence in a Markov Network

- The Markov blanket of a variable X is the set of variables that are in factors with X.
- A variable is independent of the other variables given its Markov blanket.

Independence in a Markov Network

- The Markov blanket of a variable X is the set of variables that are in factors with X.
- A variable is independent of the other variables given its Markov blanket.
- X is connected to Y given $\overline{Z}$ if there is a path from X to Y in the Markov network, which does not contain an element of Z.
- X is separated from Y given $\overline{Z}$ if it is not connected.

Independence in a Markov Network

- The Markov blanket of a variable X is the set of variables that are in factors with X.
- A variable is independent of the other variables given its Markov blanket.
- X is connected to Y given $\overline{Z}$ if there is a path from X to Y in the Markov network, which does not contain an element of Z.
- X is separated from Y given $\overline{Z}$ if it is not connected.
- A positive factor is one that does not contain zero values.
- $\overline{X}$ is independent $\overline{Y}$ given $\overline{Z}$ for all positive factors iff $\overline{X}$ is separated from $\overline{Y}$ given $\overline{Z}$

Canonical Representations

- The parameters of a graphical model are the numbers that define the model.
- A belief network is a canonical representation: given the structure and the distribution, the parameters are uniquely determined.
- A Markov random field is not a canonical representation. Many different parameterizations result in the same distribution.