

Lecture 4: Bayesian Networks

[Resources: 6.825 Techniques in Artificial Intelligence, MIT OCW]

CSCI 545 Introduction to Robotics
Instructor: Stefanos Nikolaidis

Motivation

- Why do we need Bayesian Networks?
 - To represent the dependencies among variables.

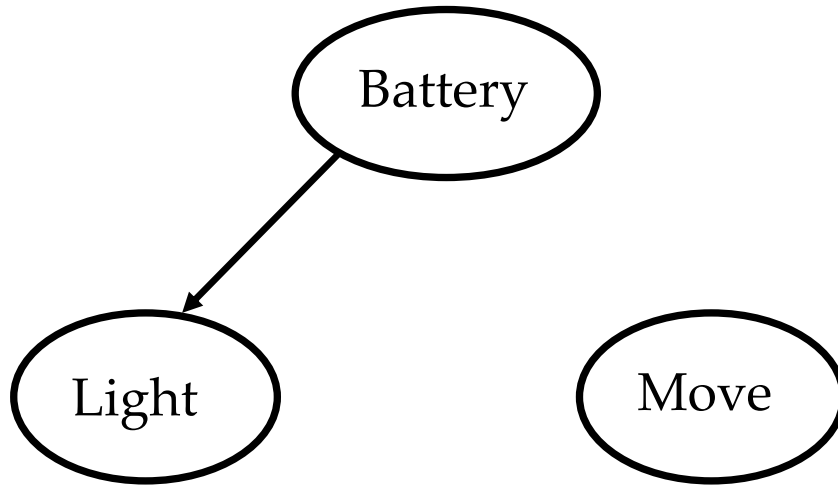
Examples

Battery

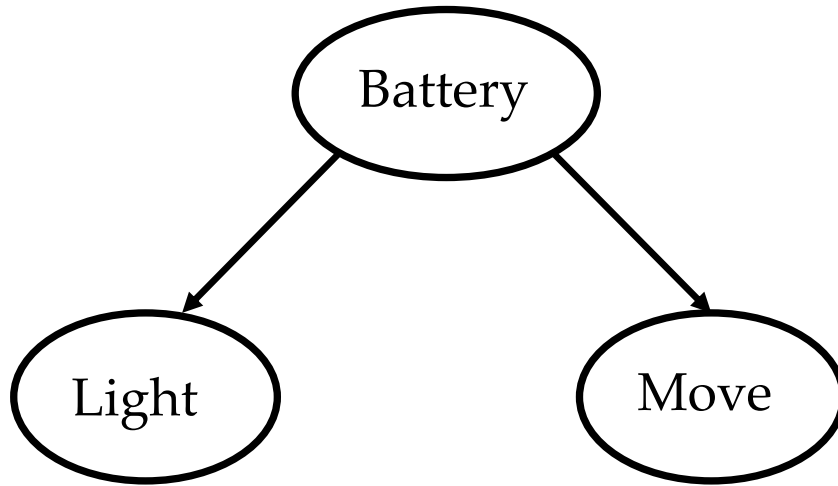
Light

Move

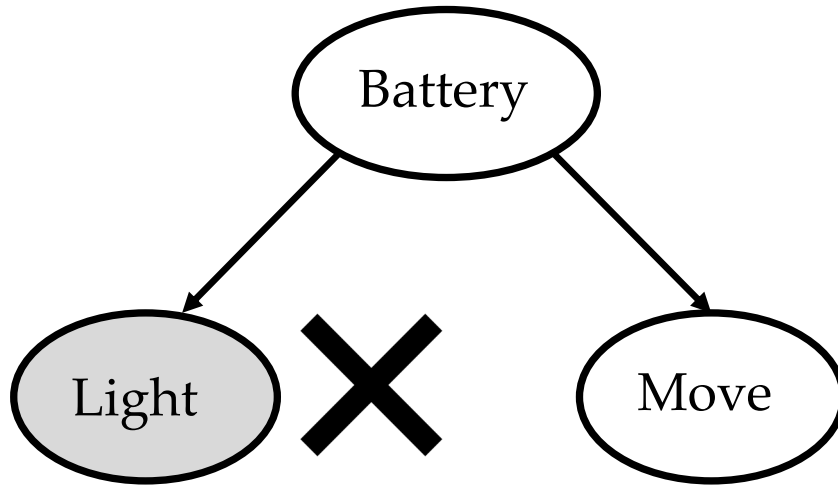
Examples



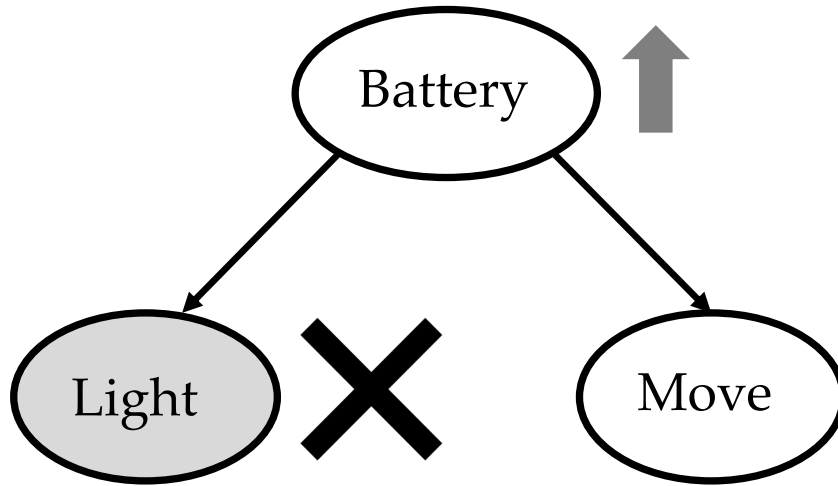
Examples



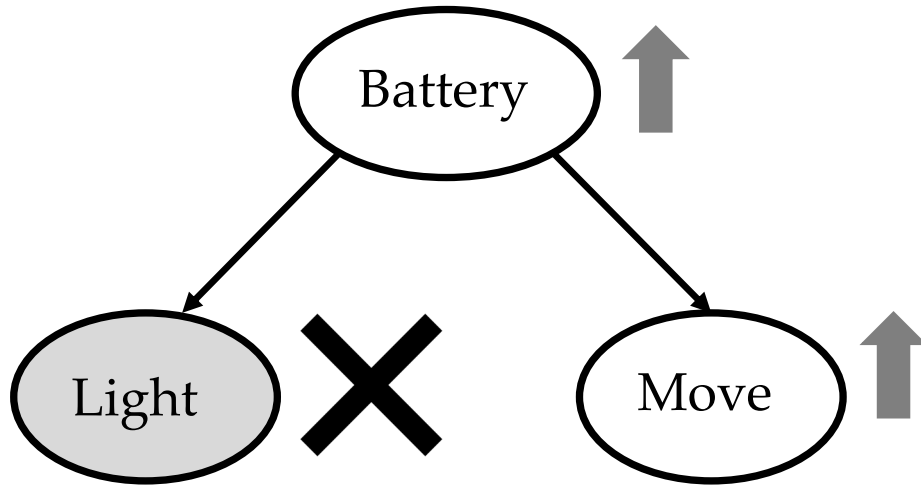
Examples



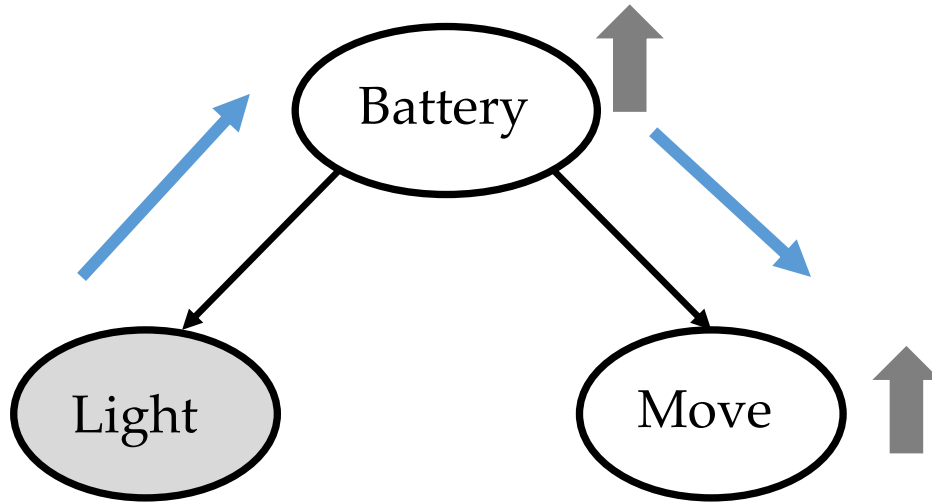
Examples



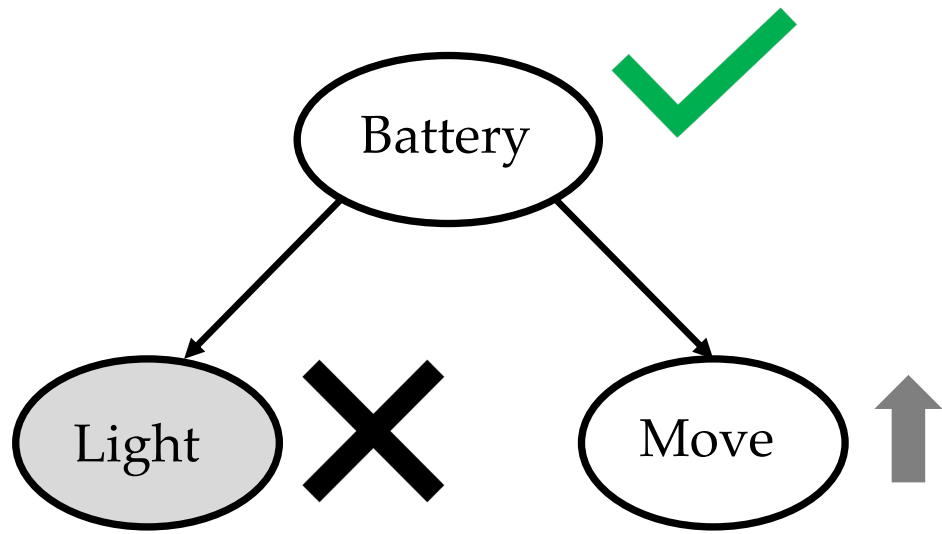
Examples



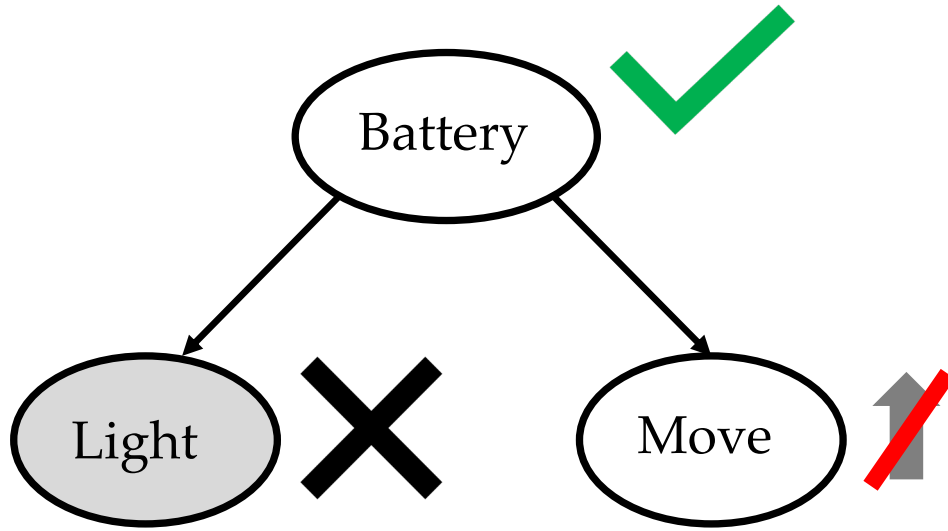
Information Flow



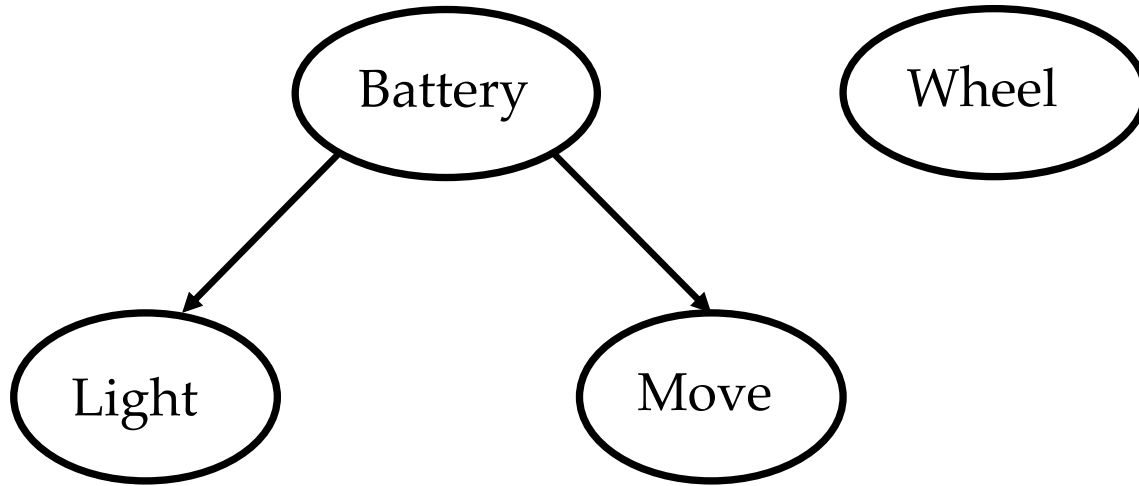
Examples



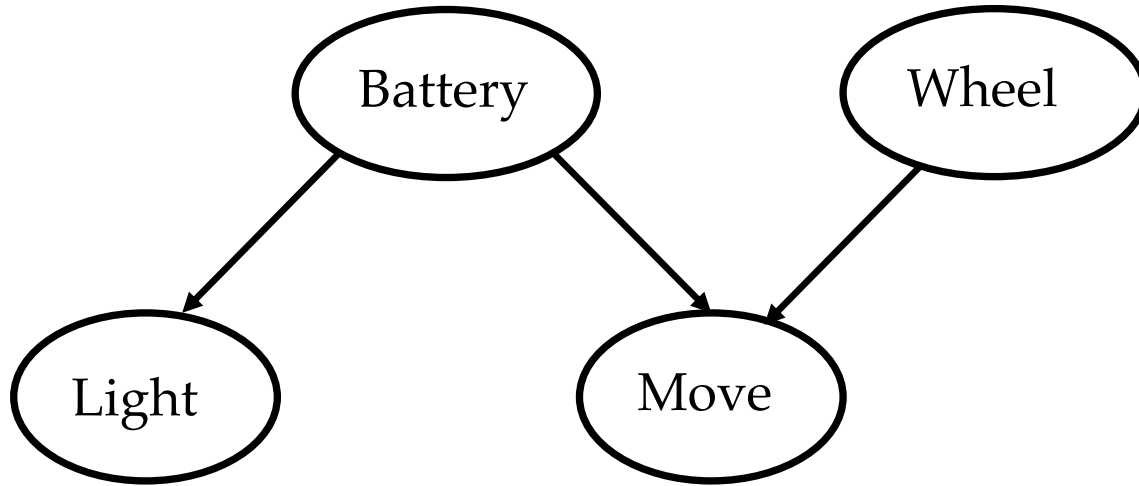
Examples



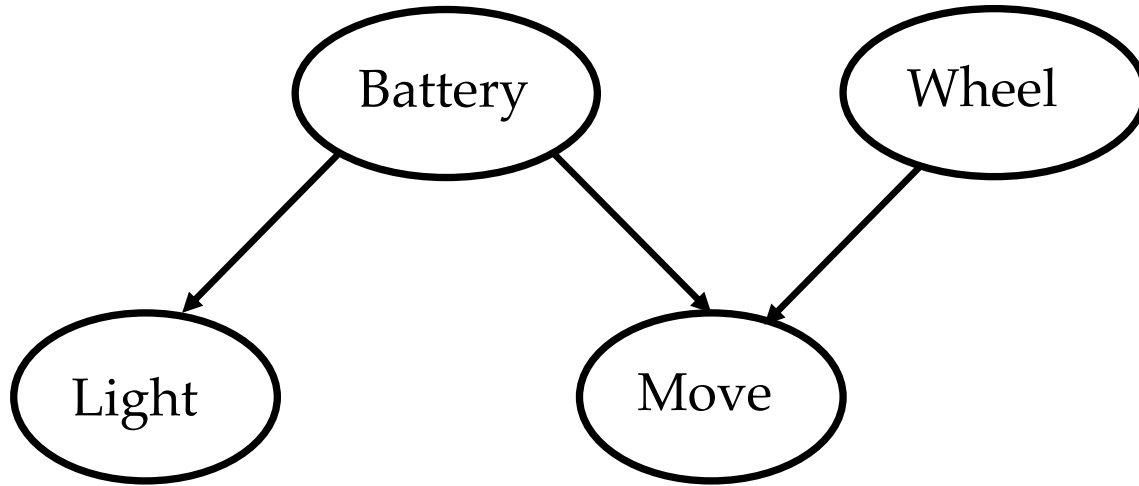
Examples



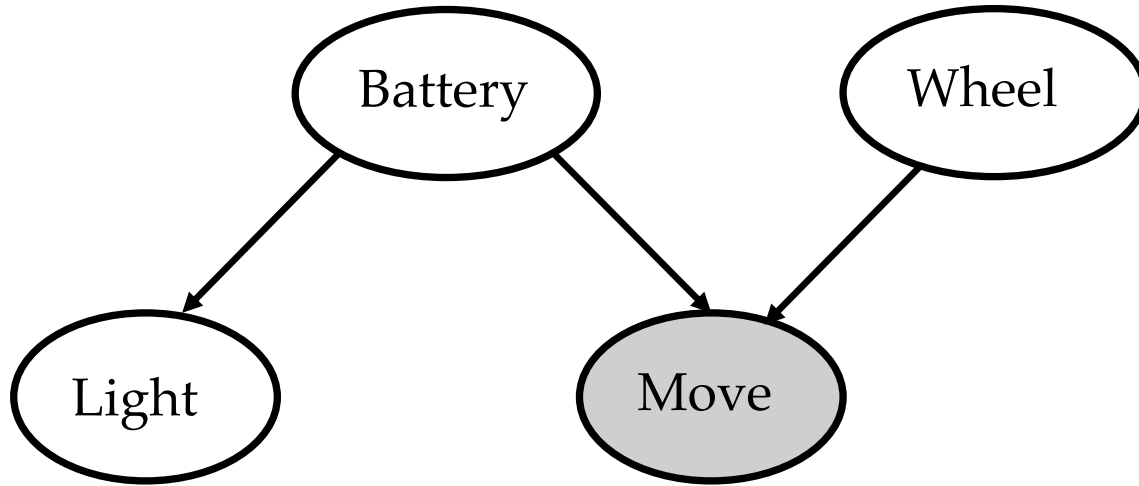
Examples



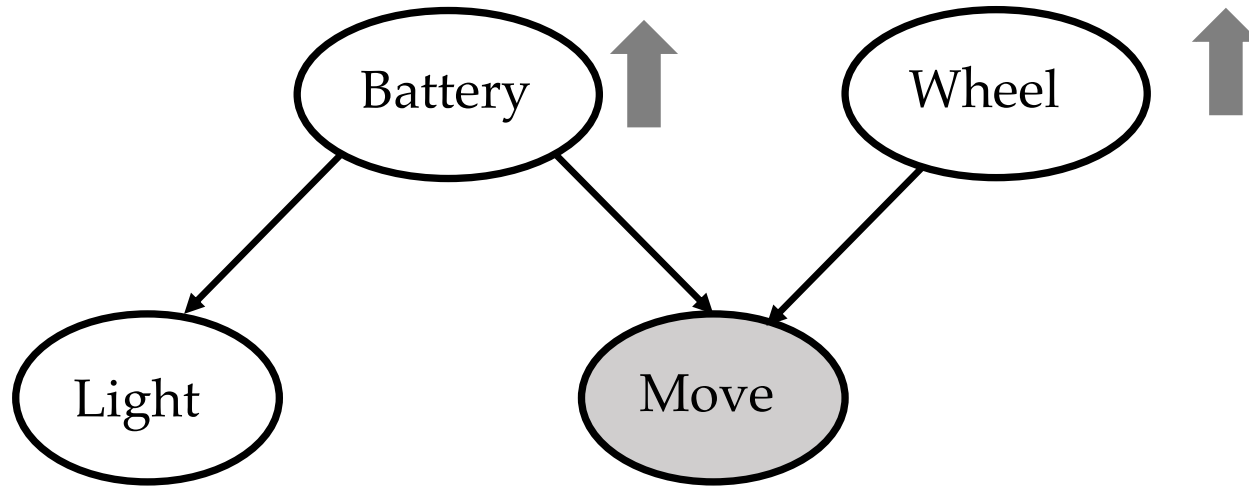
Examples



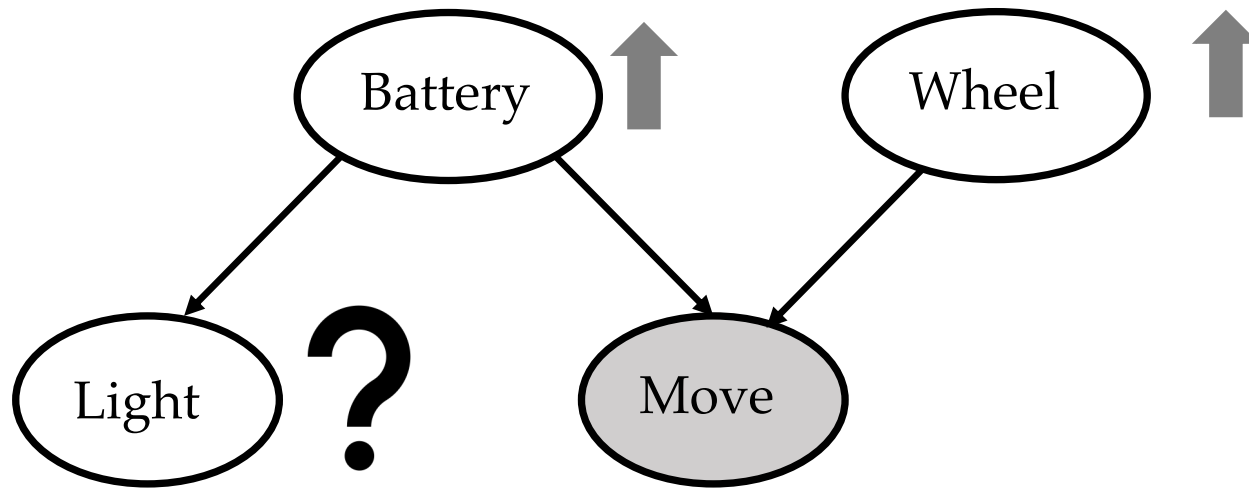
Examples



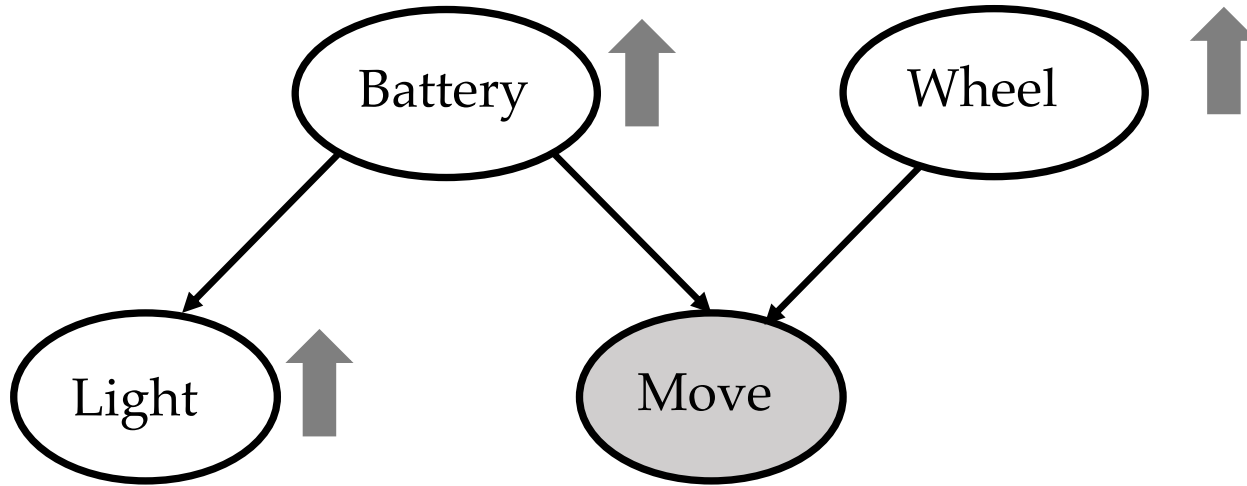
Examples



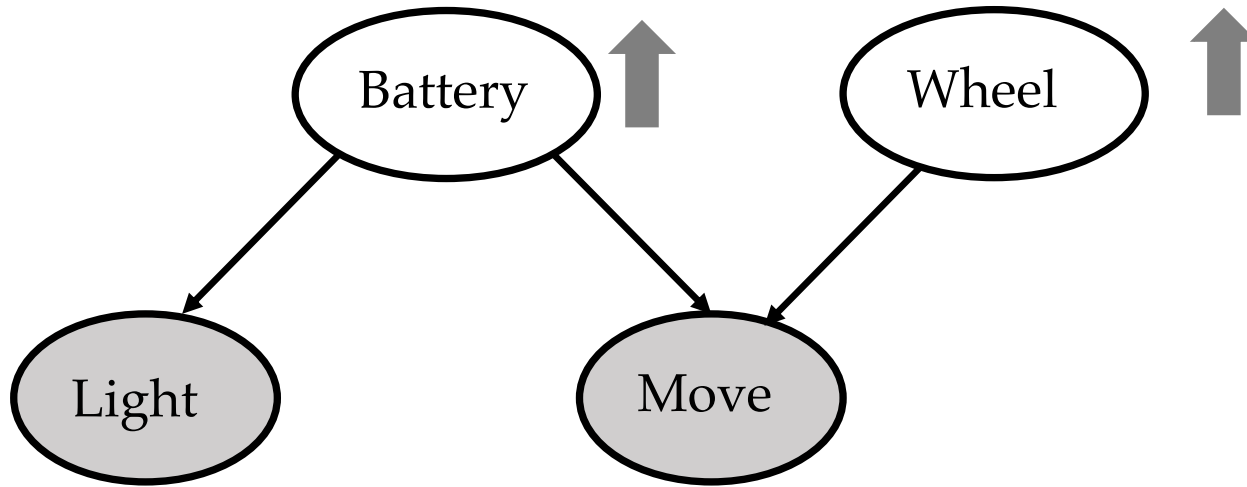
Examples



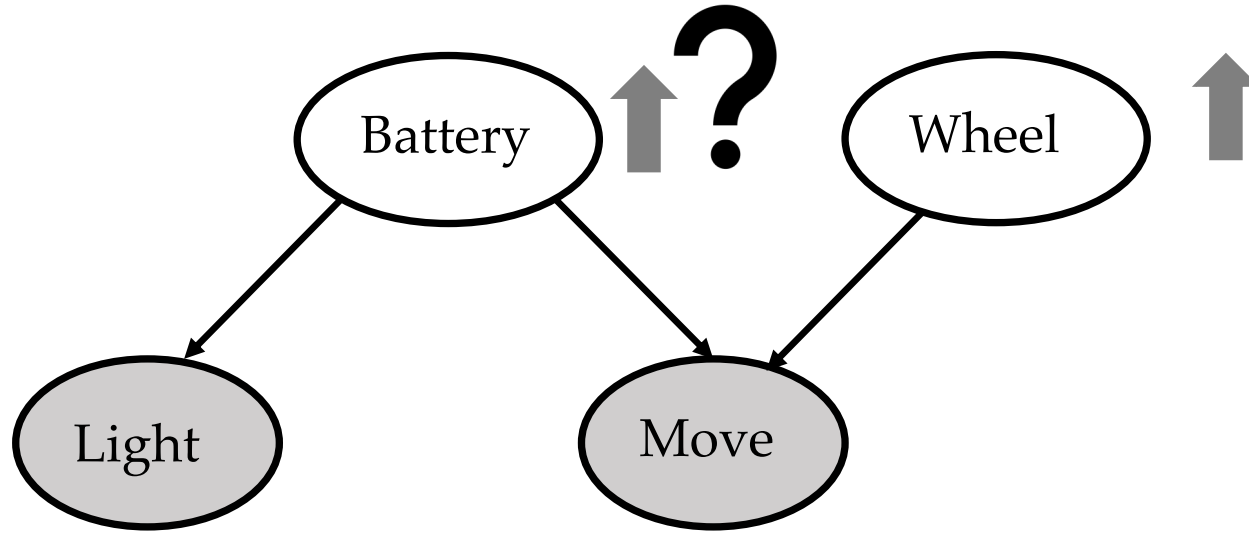
Examples



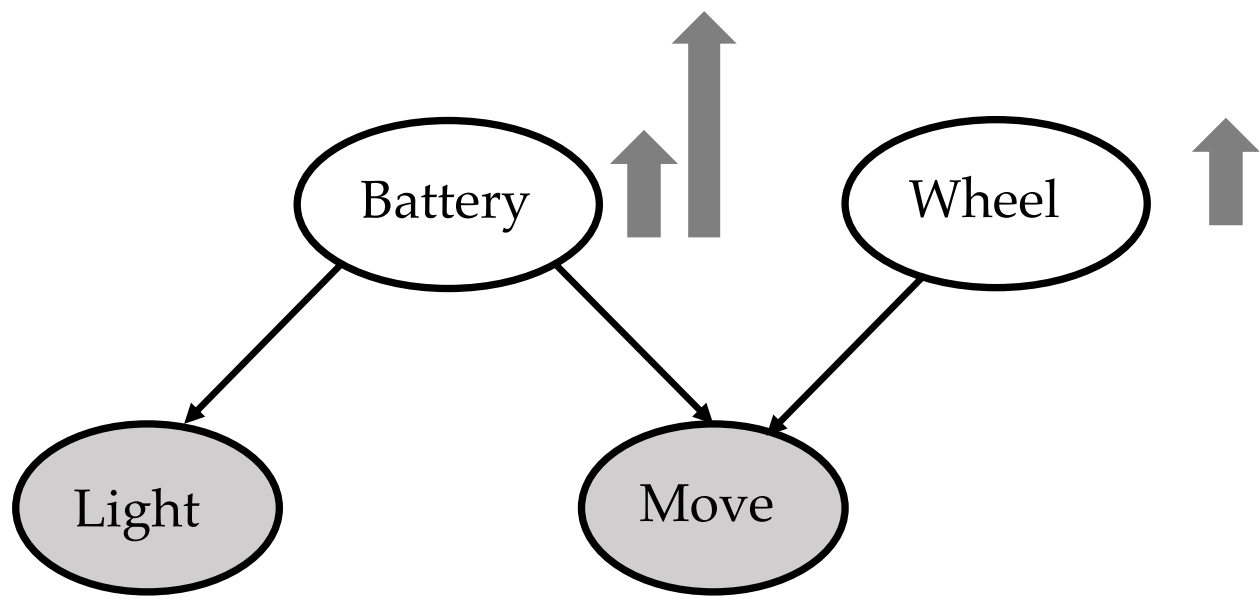
Examples



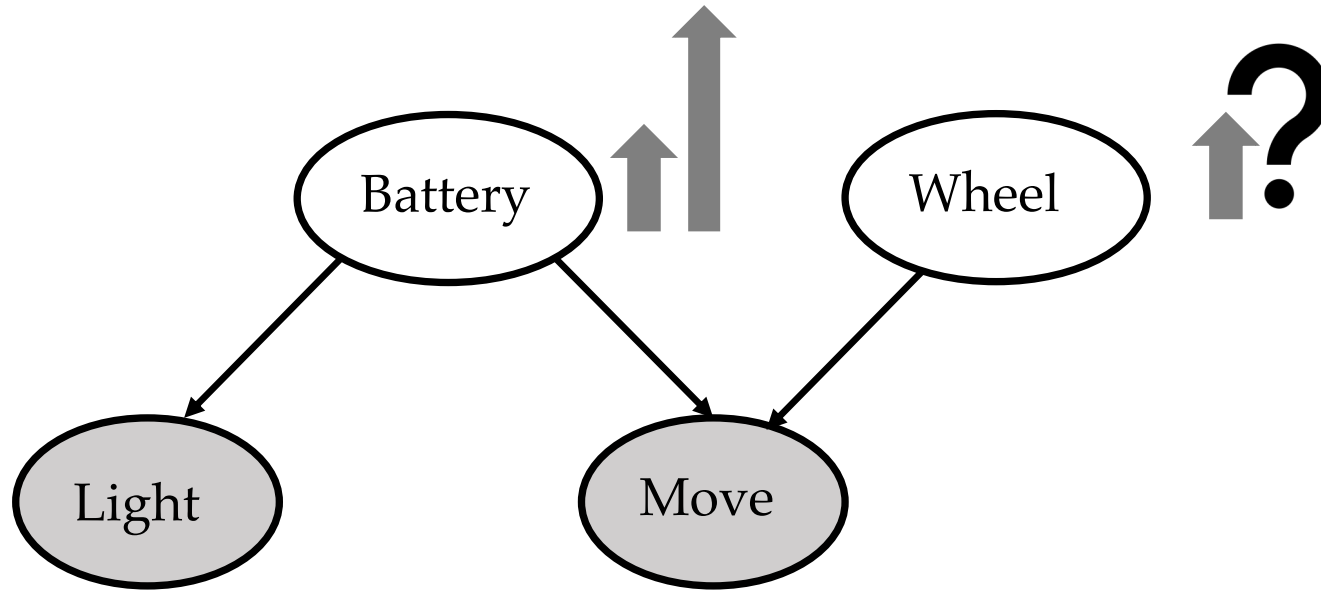
Examples



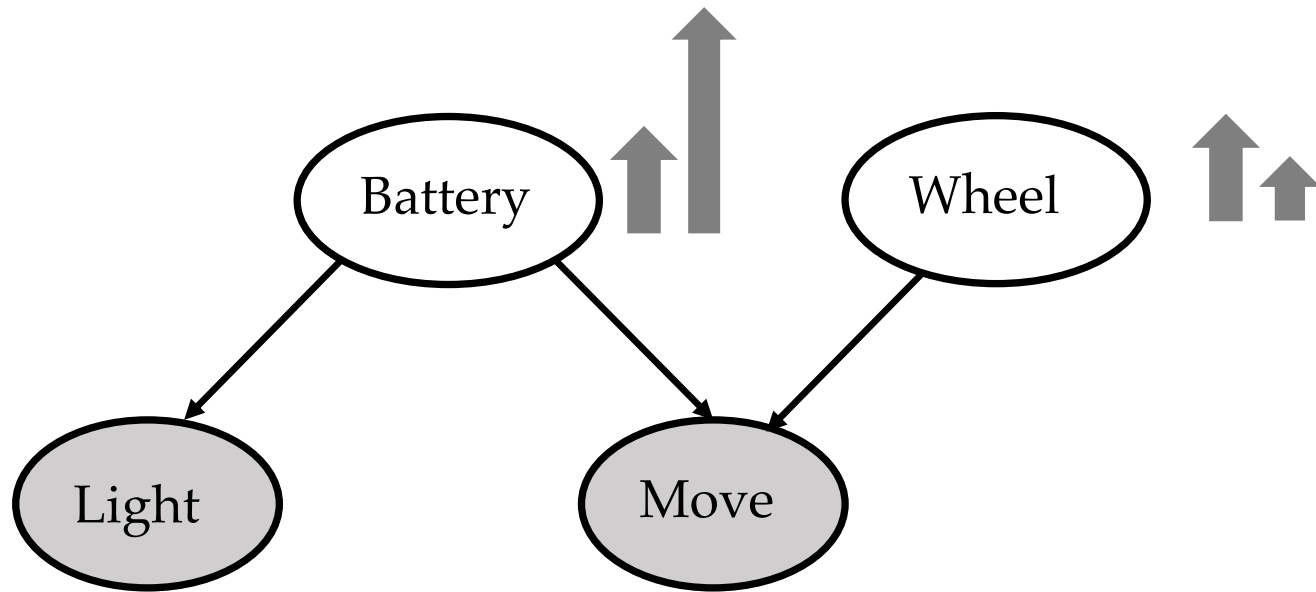
Examples



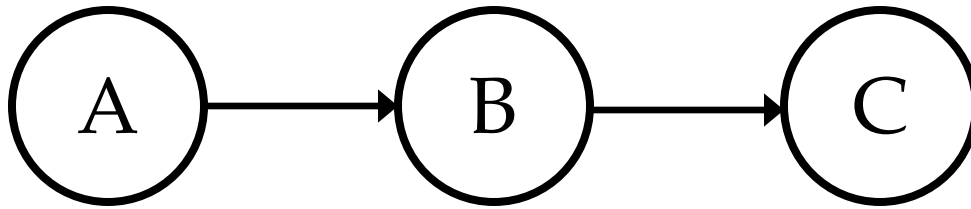
Examples



Examples



Propagation of Information: Serial Connection



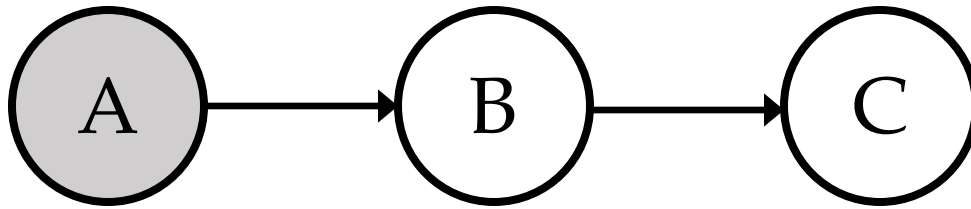
Example:

A: Battery dead

B: Robot won't start

C: Robot won't move

Propagation of Information: Serial Connection



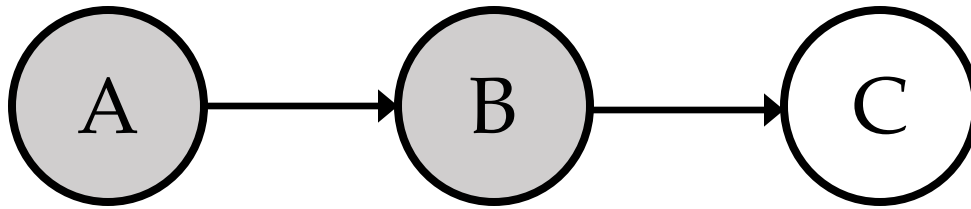
Example:

A: Battery dead

B: Robot won't start

C: Robot won't move

Propagation of Information: Serial Connection



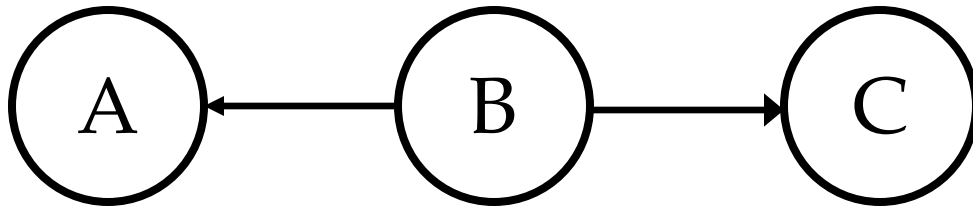
Example:

A: Battery dead

B: Robot won't start

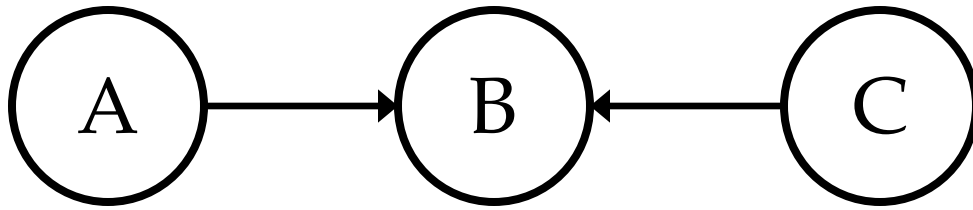
C: Robot won't move

Propagation of Information: Diverging Connection

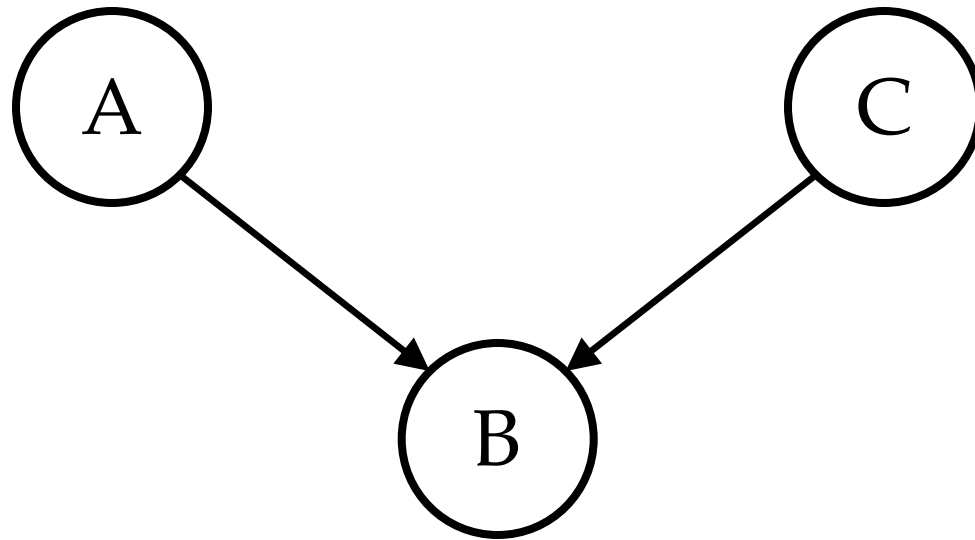


- Knowing about A will tell us something about C.
Knowing about C will tell us something about A.
- If we know B propagation is blocked

Propagation of Information: Converging Connection



Propagation of Information: Converging Connection



Example:

A = wheels damaged

B = robot does not move

C = battery off

Bayesian Networks

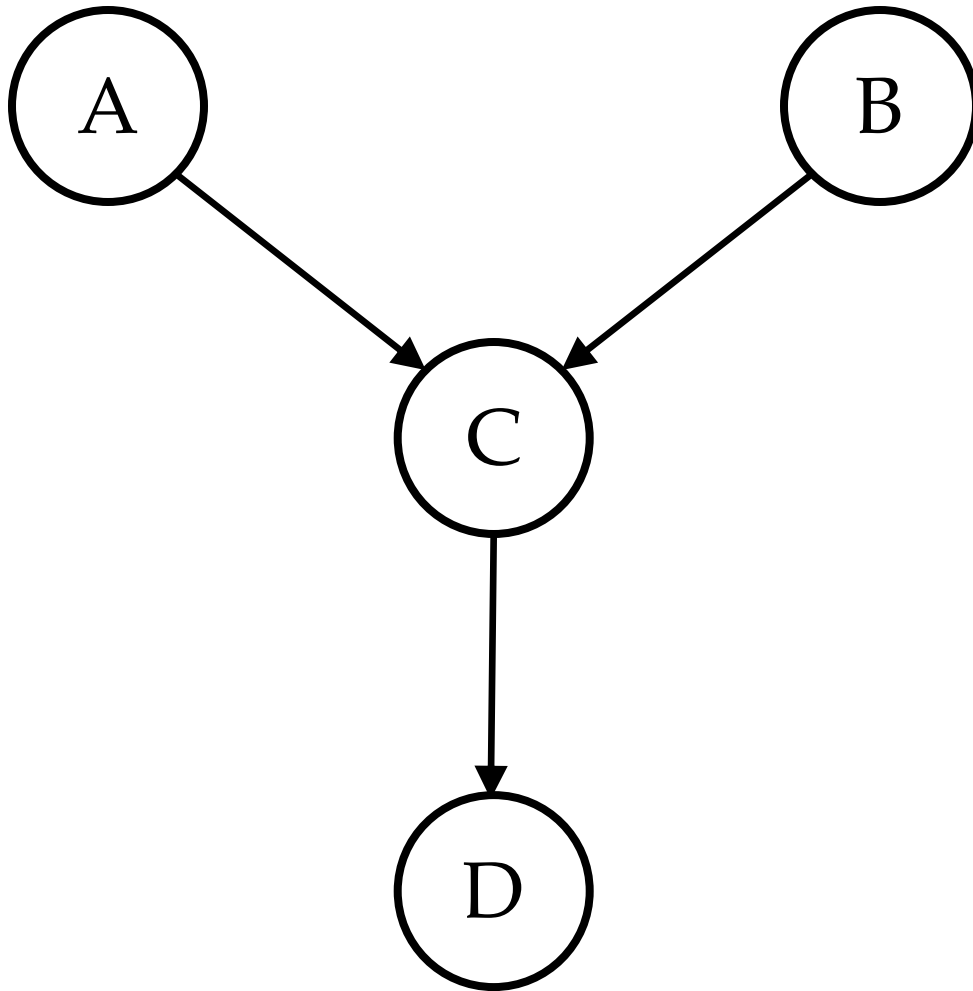
- Set of variables, set of directed arcs forming *acyclic* graphs.
- Every node A with parents B_1, \dots, B_n has $P(A | B_1, \dots, B_n)$
- *Each variable is conditionally independent of its non-descendants, given its parents.*

Chain Rule

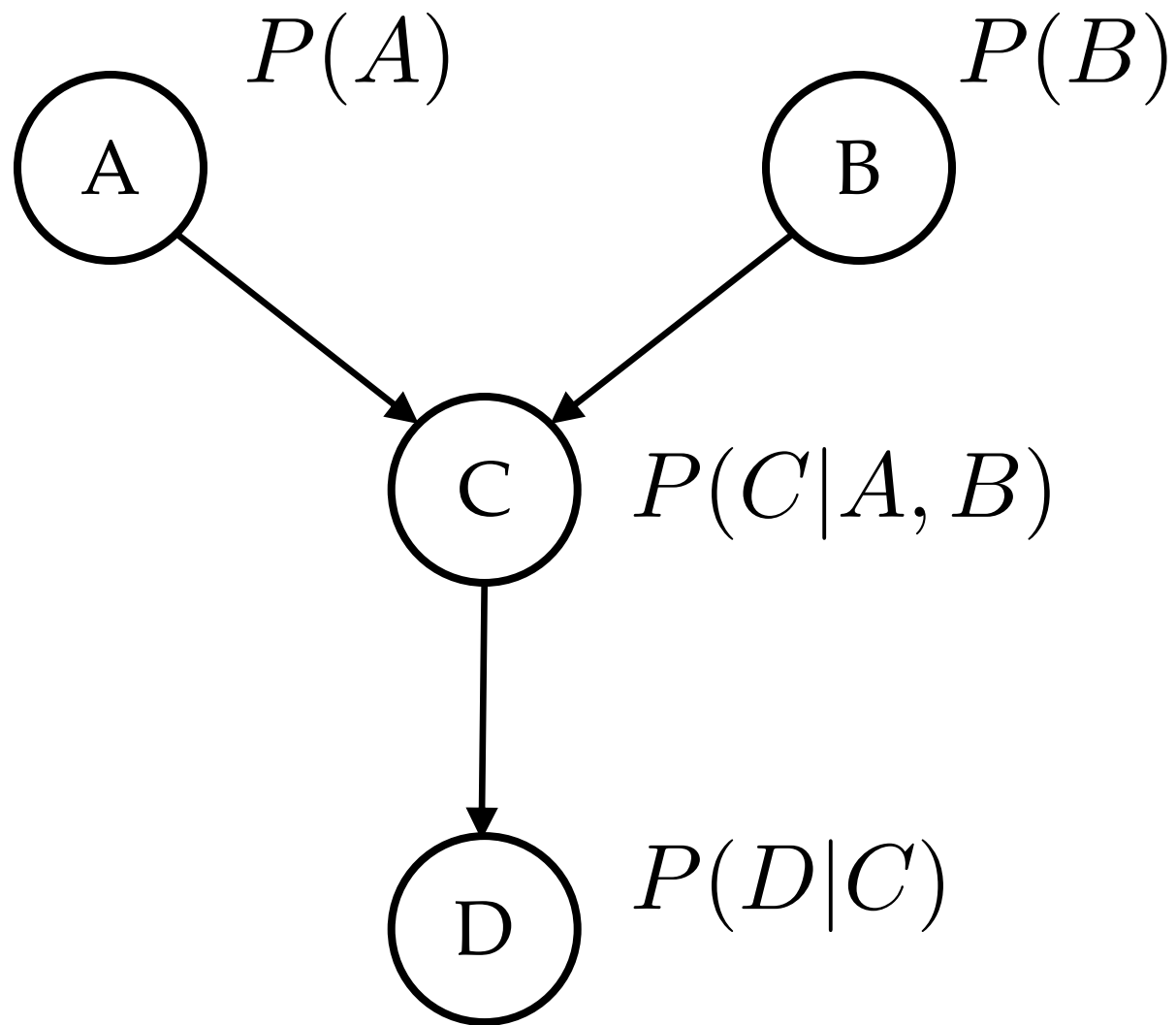
$$P(V_1 = v_1, V_2 = v_2, \dots, V_n = v_n) = \prod_i P(V_i = v_i | \text{parents}(V_i))$$

We say that the joint probability distribution is the product of all the individual probability distributions that are stores in the nodes of the graph.

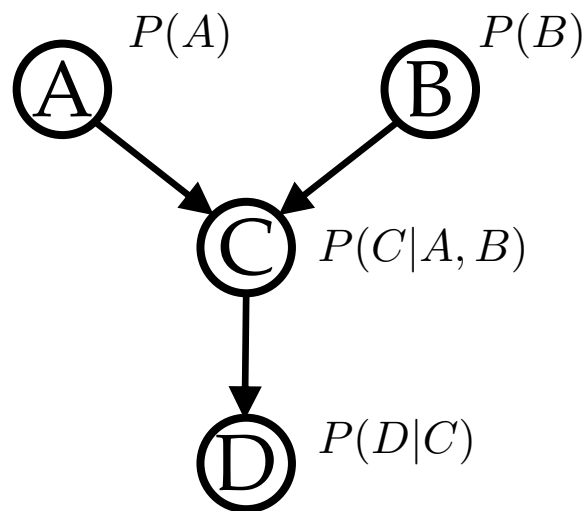
Example



Example

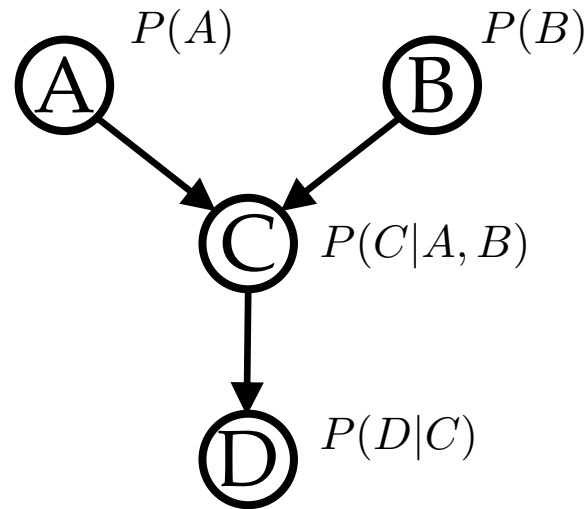


Example



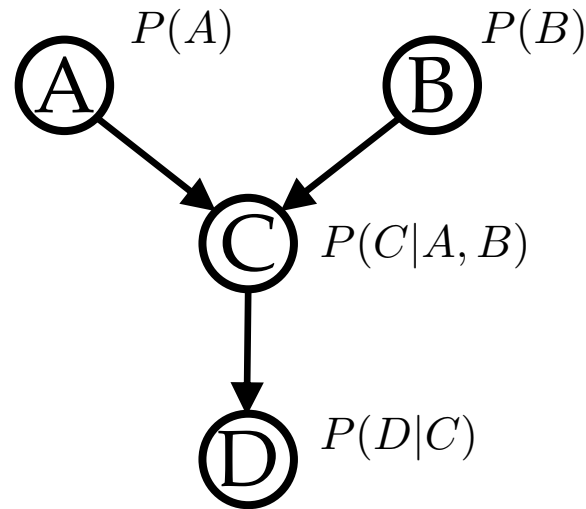
$$P(A, B, C, D) = P(A = \text{True}, B = \text{True}, C = \text{True}, D = \text{True})$$

Example



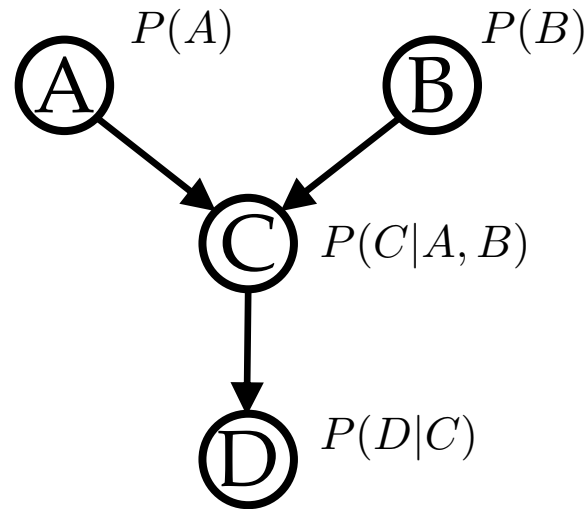
$$\begin{aligned} P(A, B, C, D) &= P(A = \text{True}, B = \text{True}, C = \text{True}, D = \text{True}) \\ &= P(D|A, B, C)P(A, B, C) \end{aligned}$$

Example



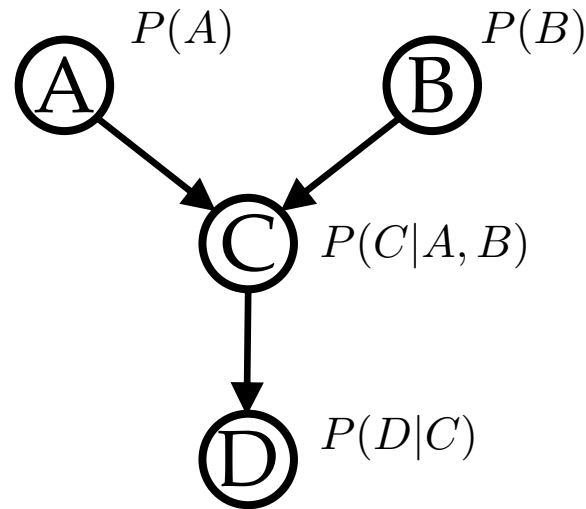
$$\begin{aligned} P(A, B, C, D) &= P(A = \text{True}, B = \text{True}, C = \text{True}, D = \text{True}) \\ &= P(D|A, B, C)P(A, B, C) \\ &= P(D|C)P(A, B, C) \end{aligned}$$

Example



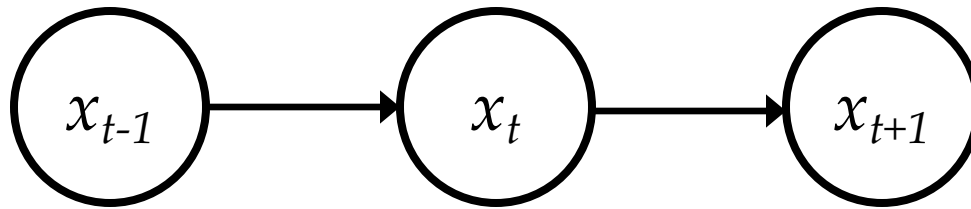
$$\begin{aligned} P(A, B, C, D) &= P(A = \text{True}, B = \text{True}, C = \text{True}, D = \text{True}) \\ &= P(D|A, B, C)P(A, B, C) \\ &= P(D|C)P(A, B, C) \\ &= P(D|C)P(C|A, B)P(A, B) \end{aligned}$$

Example



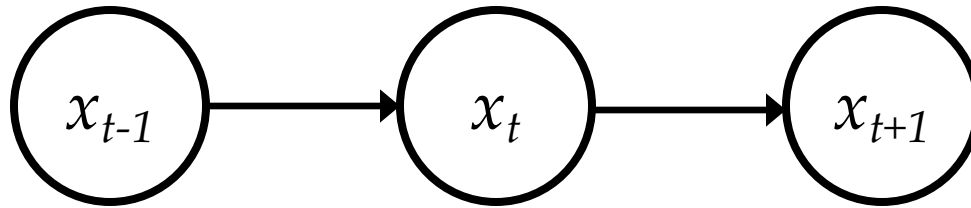
$$\begin{aligned} P(A, B, C, D) &= P(A = \text{True}, B = \text{True}, C = \text{True}, D = \text{True}) \\ &= P(D|A, B, C)P(A, B, C) \\ &= P(D|C)P(A, B, C) \\ &= P(D|C)P(C|A, B)P(A, B) \\ &= P(D|C)P(C|A, B)P(A)P(B) \end{aligned}$$

State Estimation Over Time



What type of connection?

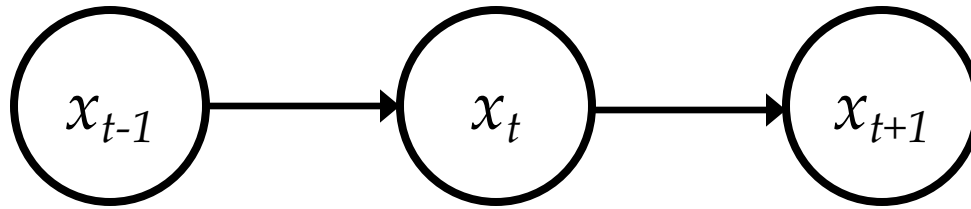
State Estimation Over Time



What type of connection?

- **Serial Connection!**

State Estimation Over Time



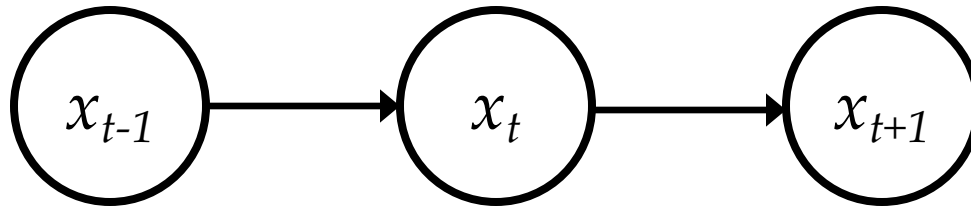
What type of connection?

- **Serial Connection!**

State Transitions:

$$P(X_t | X_{0:t-1}) =$$

State Estimation Over Time



What type of connection?

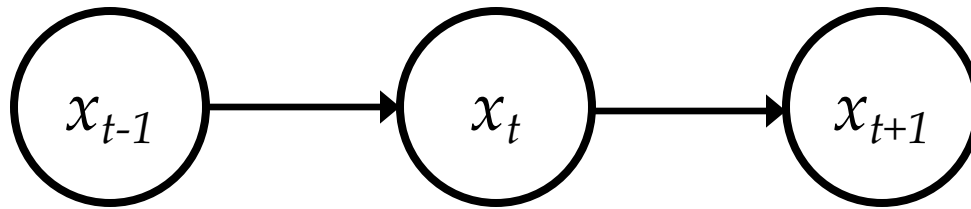
- **Serial Connection!**

State Transitions:

$$P(X_t | X_{0:t-1}) = P(X_t | X_{t-1})$$

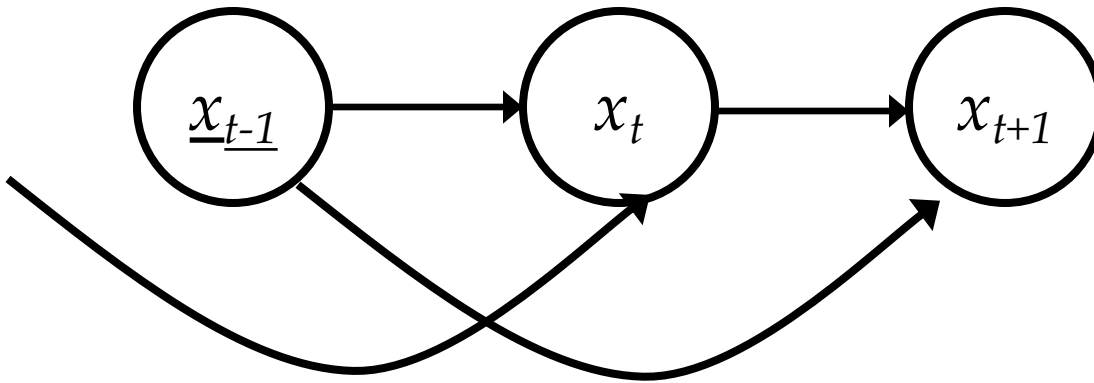
Markov Assumption

Second Order Model



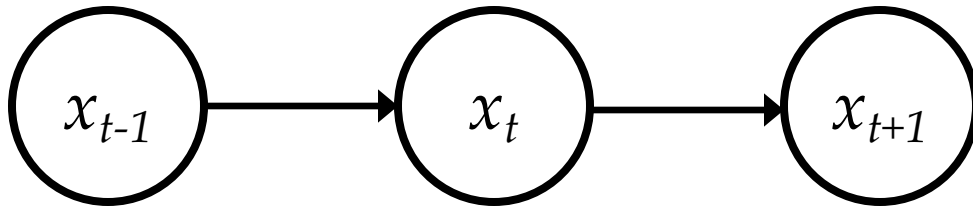
$$P(X_t | X_{0:t-1}) = P(X_t | X_{t-1}, X_{t-2})$$

Second Order Model



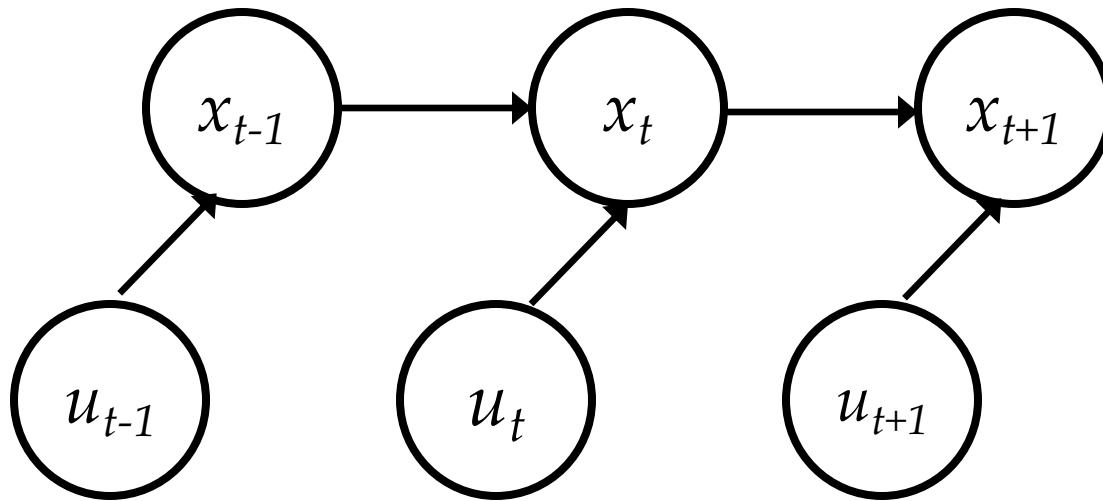
$$P(X_t | X_{0:t-1}) = P(X_t | X_{t-1}, X_{t-2})$$

Assumption of Stationary Dynamics



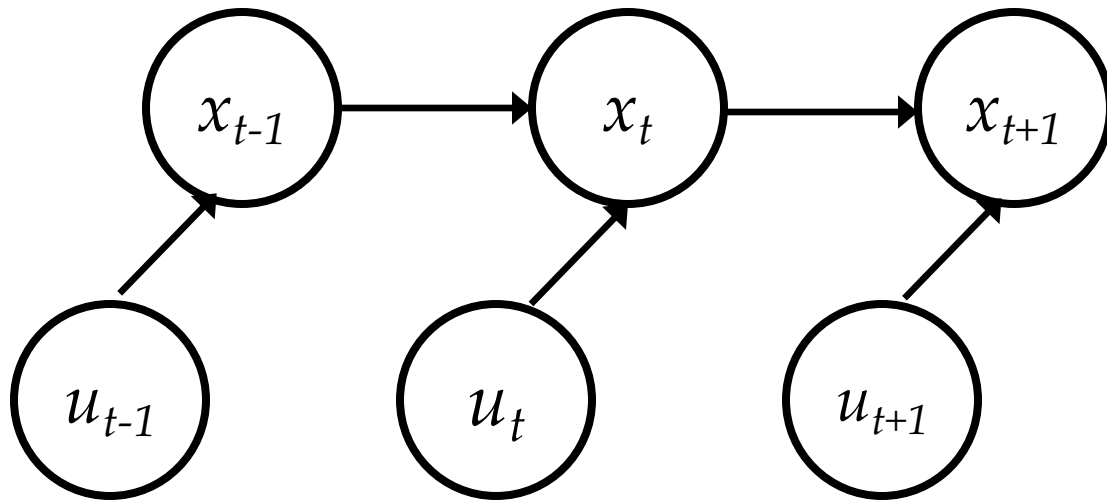
$$P(X_t | X_{t-1}) = P(X_{t-1} | X_{t-2})$$

Control Actions



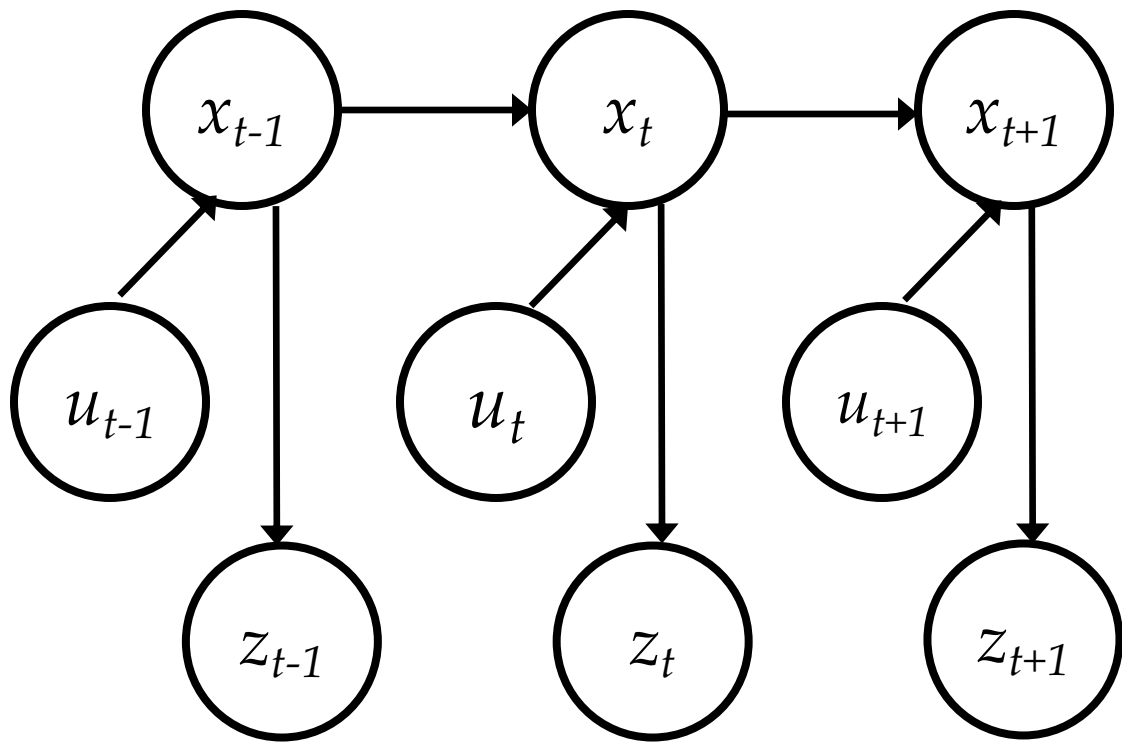
$$P(X_t | X_{0:t-1}, U_{0:t-1}) =$$

Control Actions



$$P(X_t | X_{0:t-1}, U_{0:t}) = P(X_t | X_{t-1}, U_t)$$

Sensor Model



$$P(Z_t | X_{0:t}, U_{0:t}) = P(Z_t | X_t)$$