USC Viterbi
School of Engineering

Lecture 4: Bayesian Networks

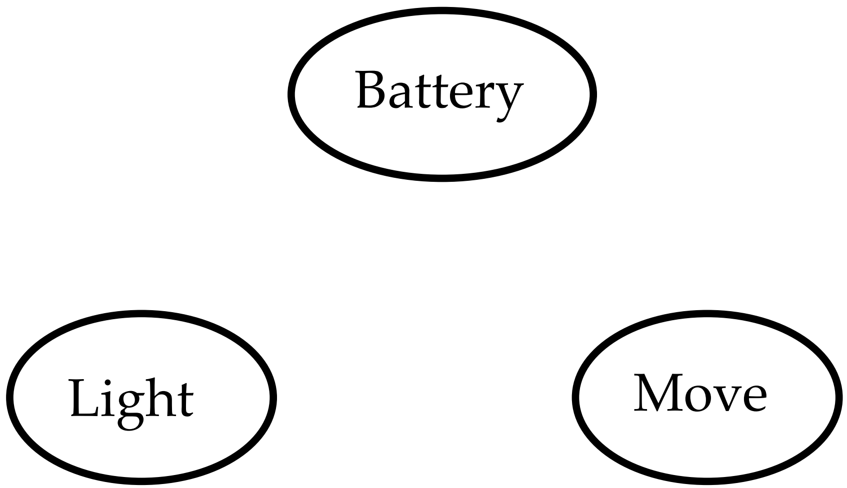
[Resources: 6.825 Techniques in Artificial Intelligence, MIT OCW]

CSCI 545 Introduction to Robotics
Instructor: Stefanos Nikolaidis

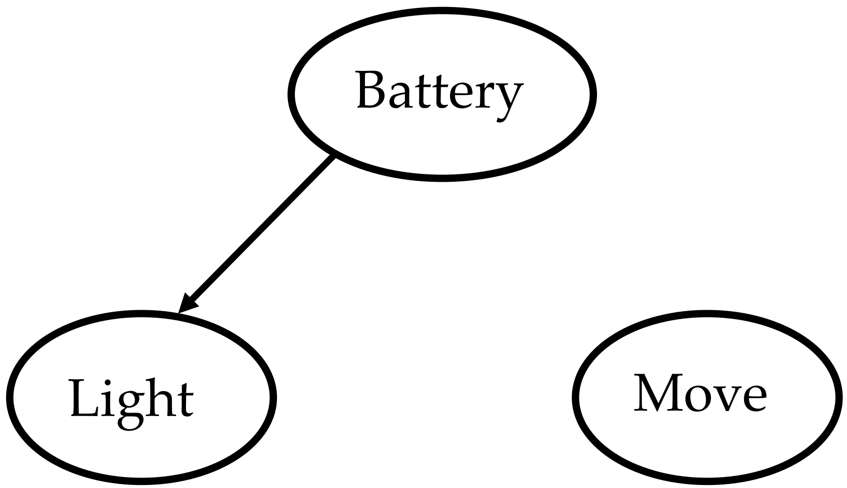
Motivation

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

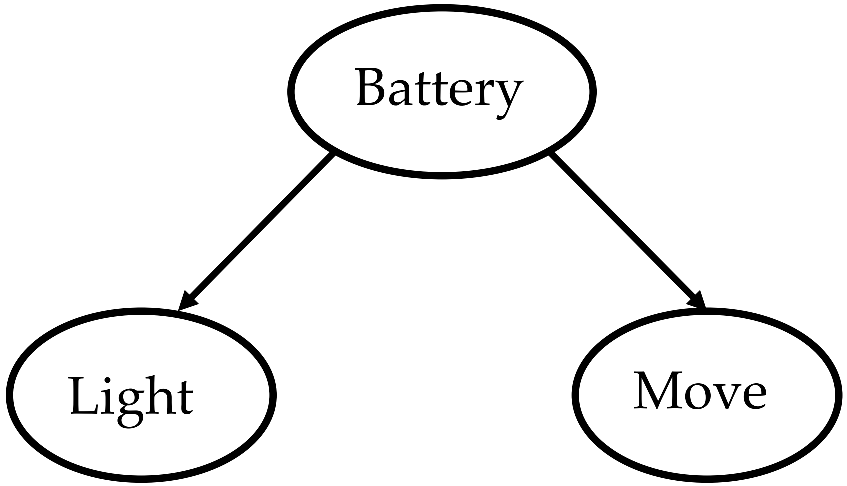
Examples



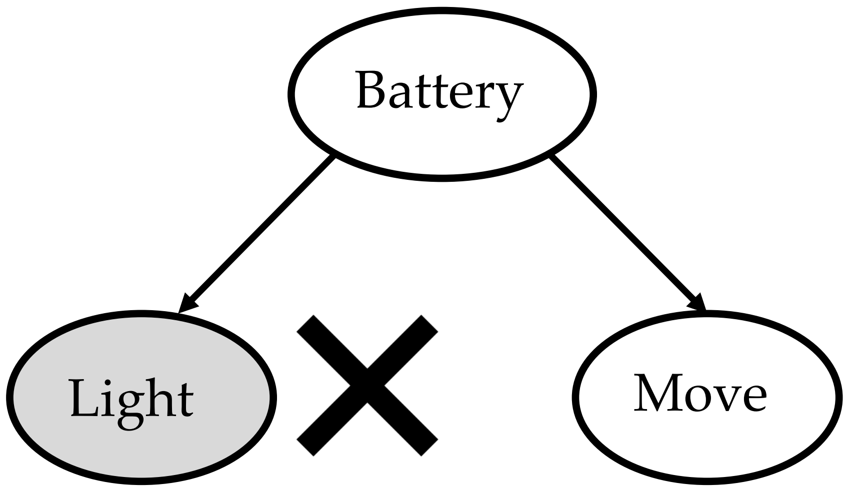
Examples



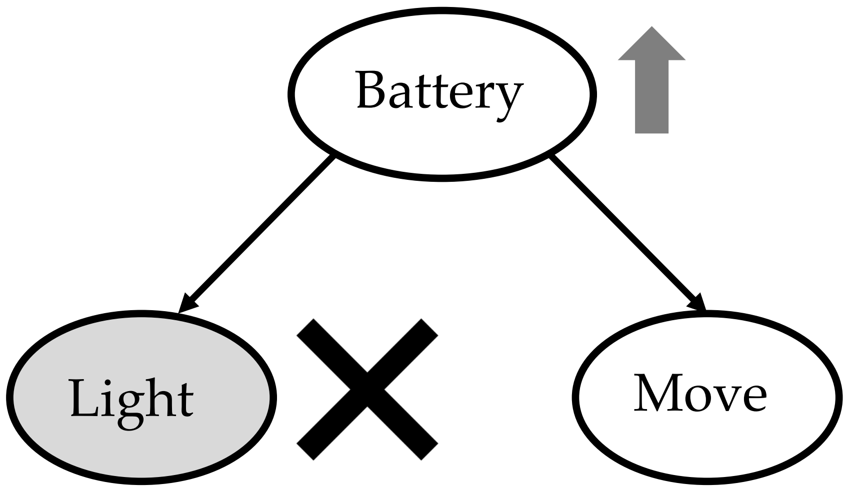
Examples



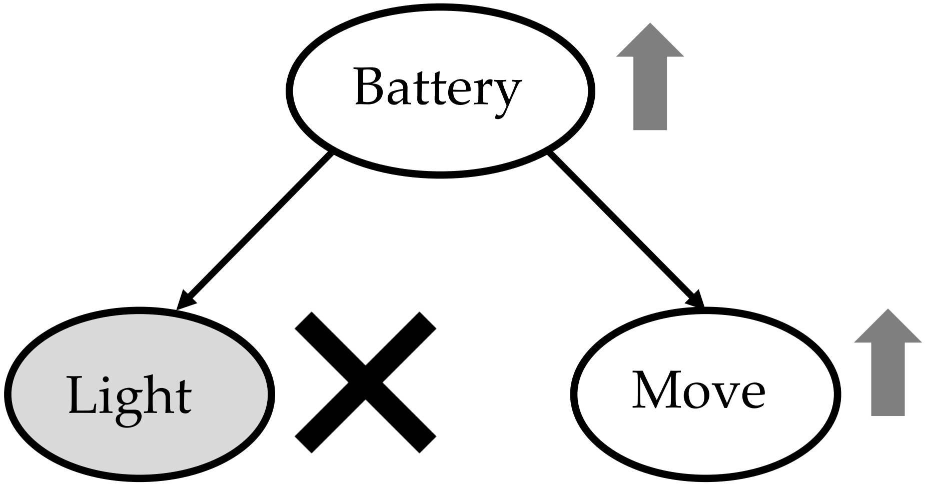
Examples



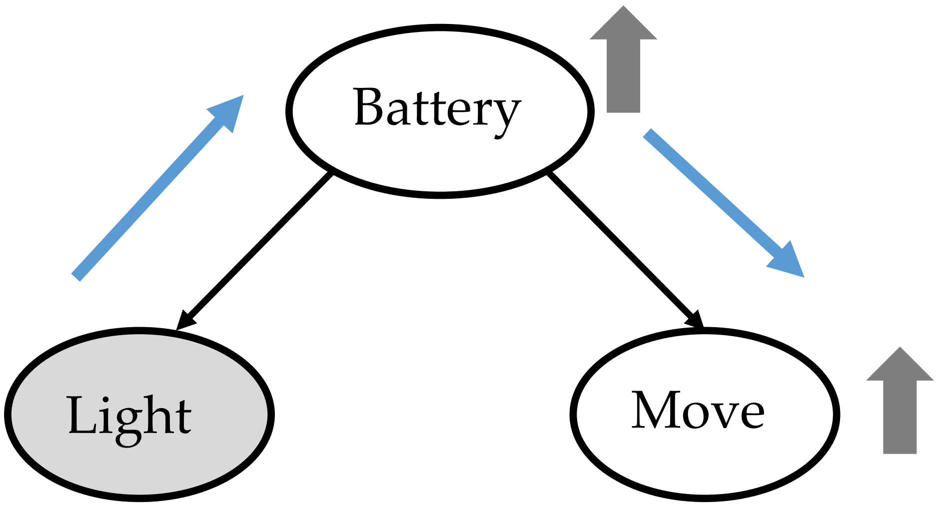
Examples



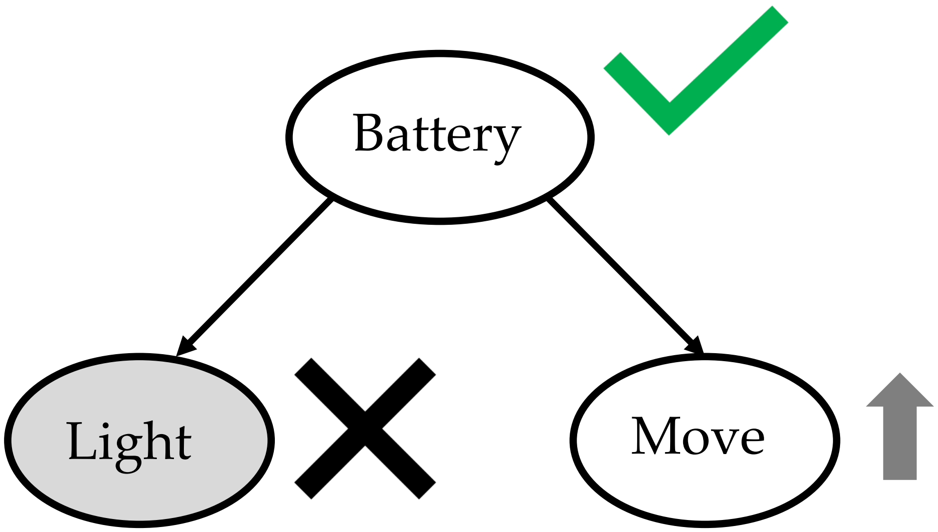
Examples



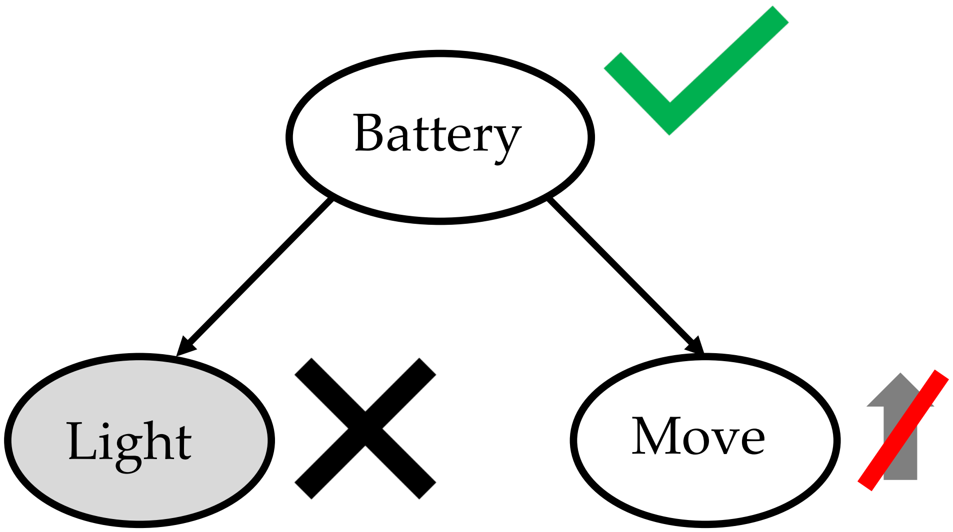
Information Flow



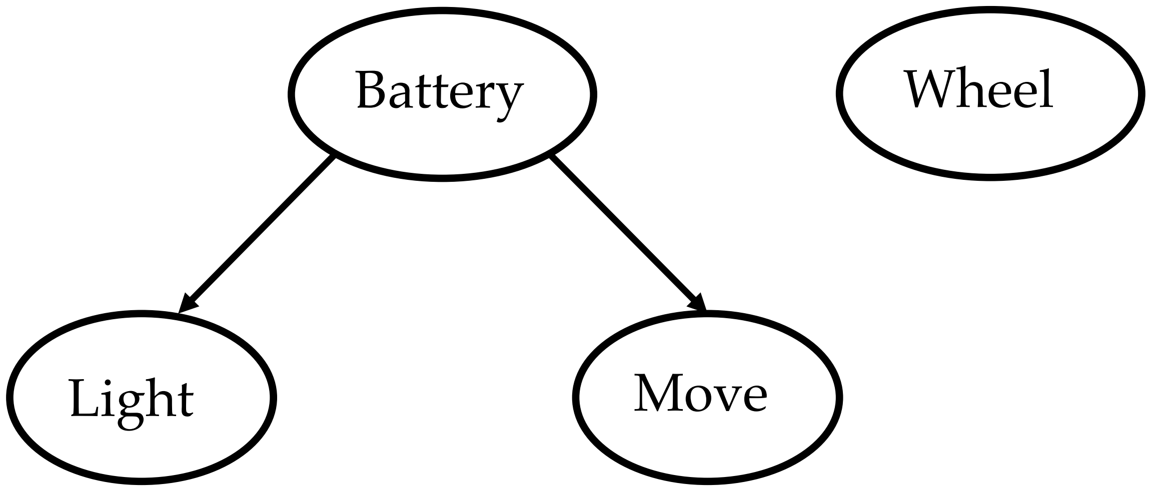
Examples



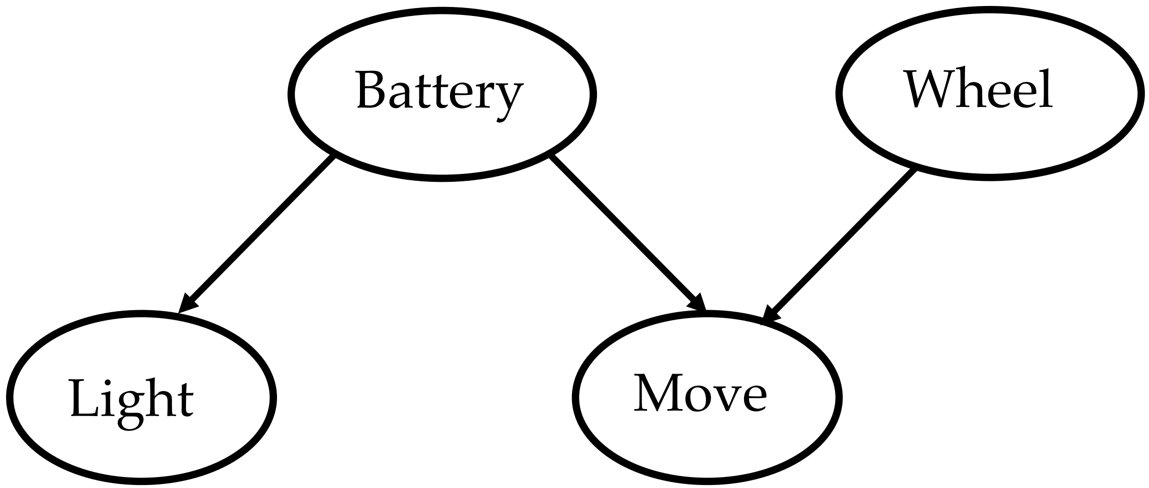
Examples



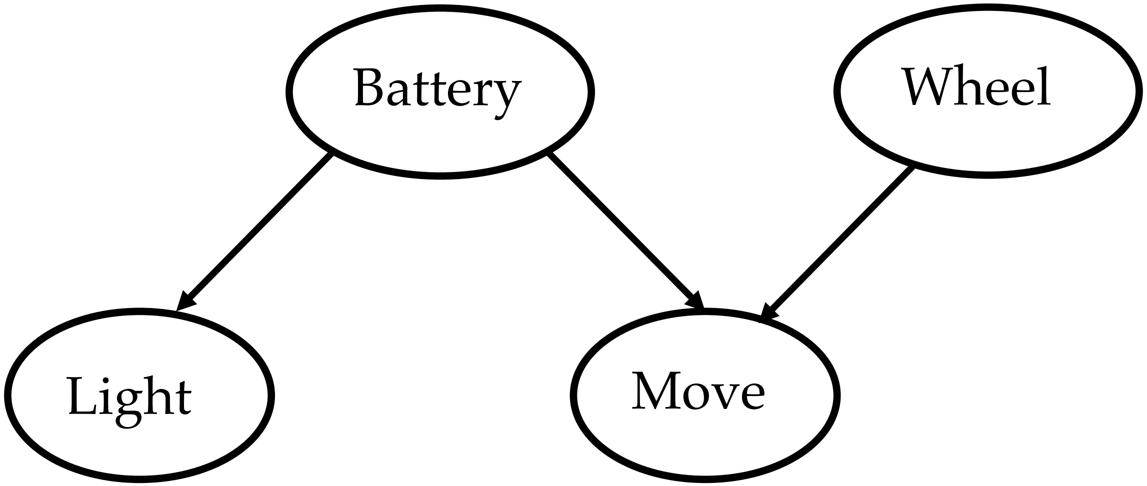
Examples



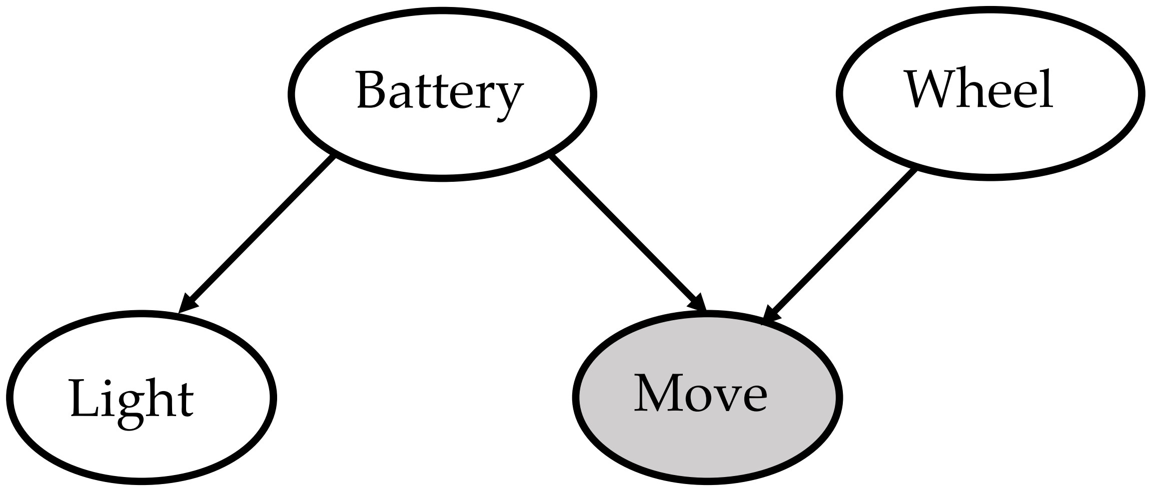
Examples



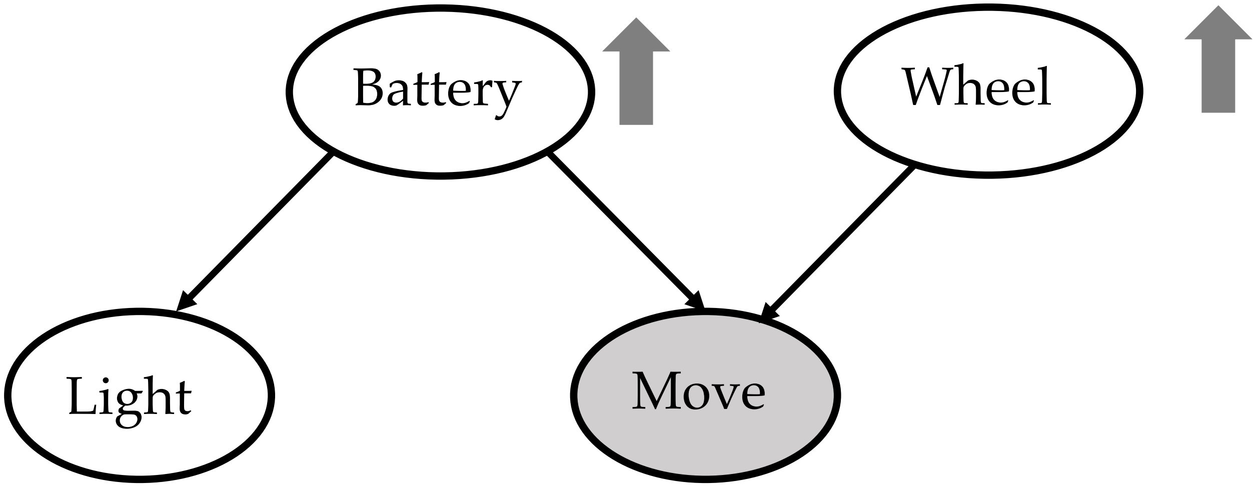
Examples



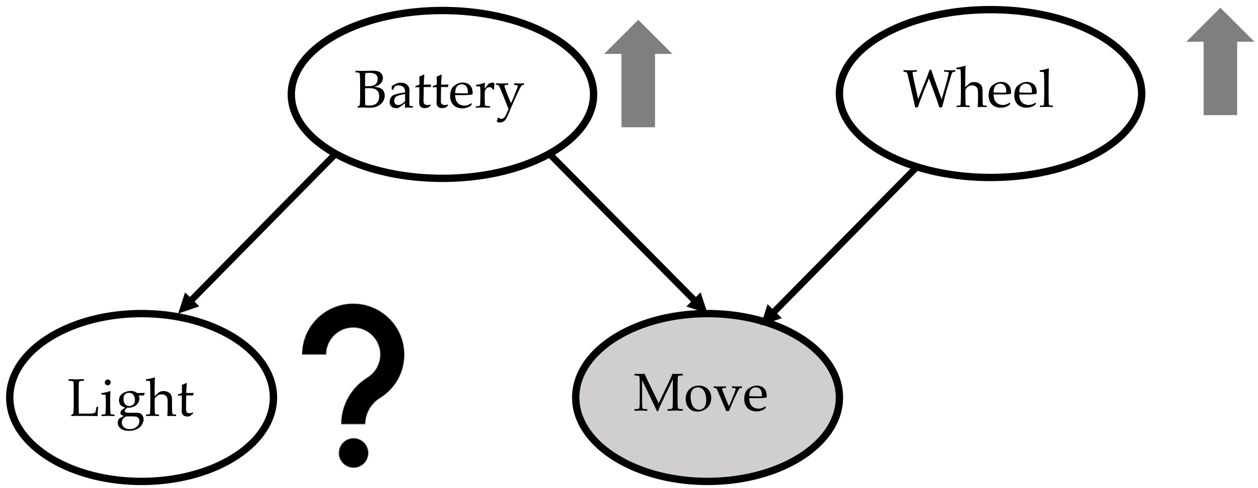
Examples



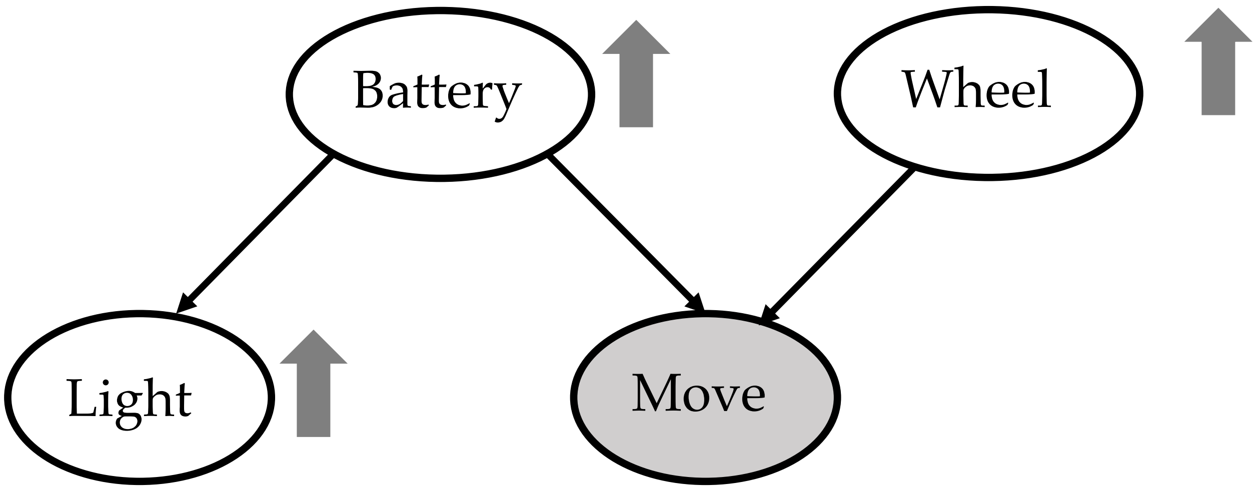
Examples



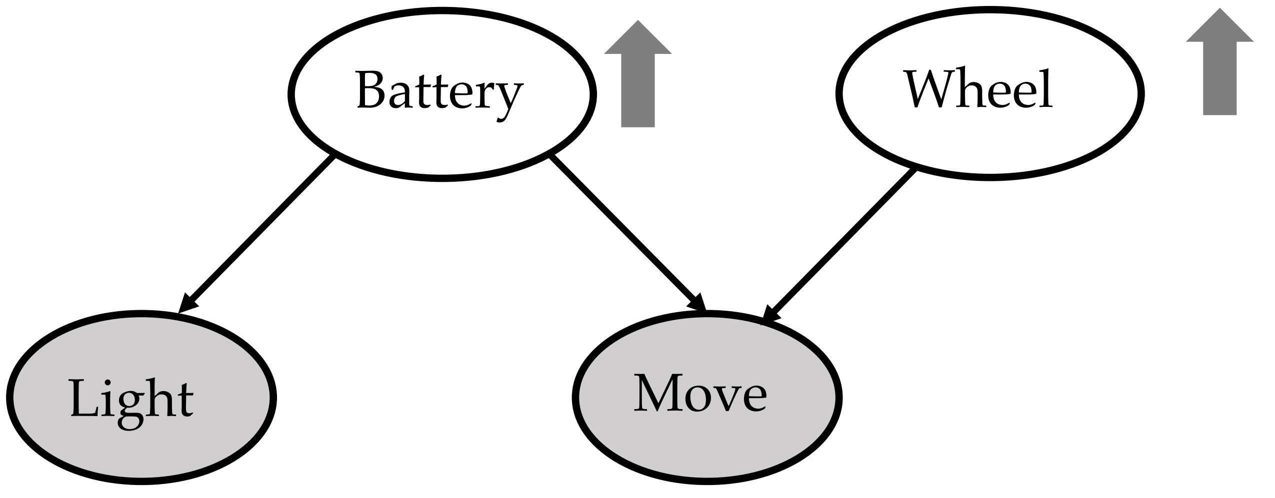
Examples



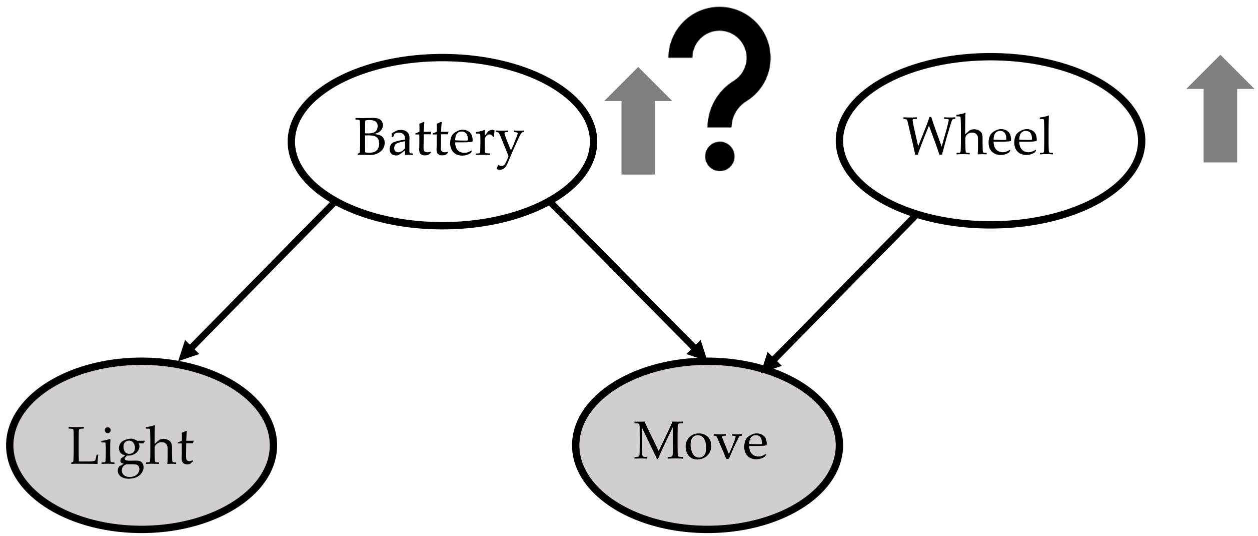
Examples



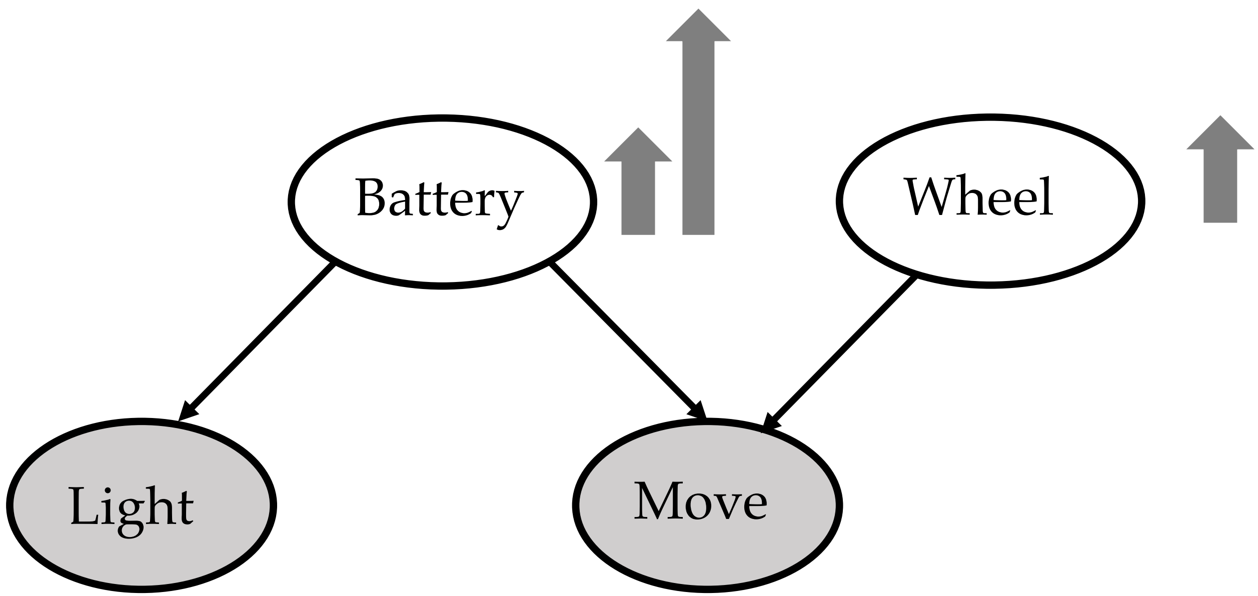
Examples



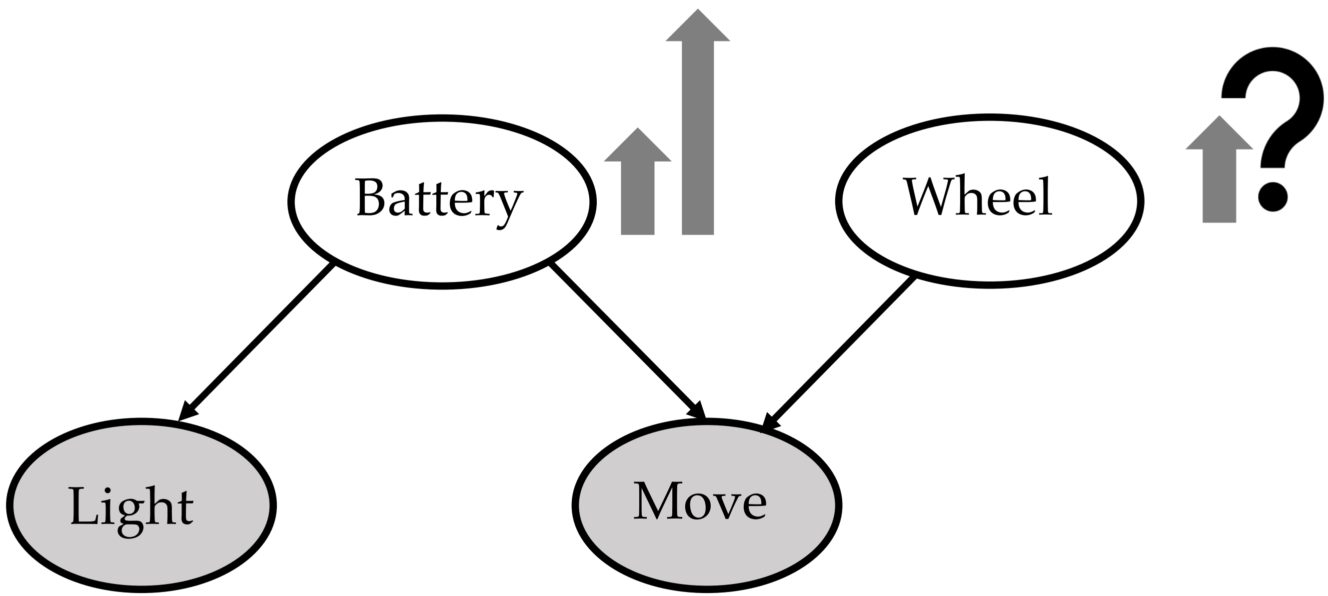
Examples



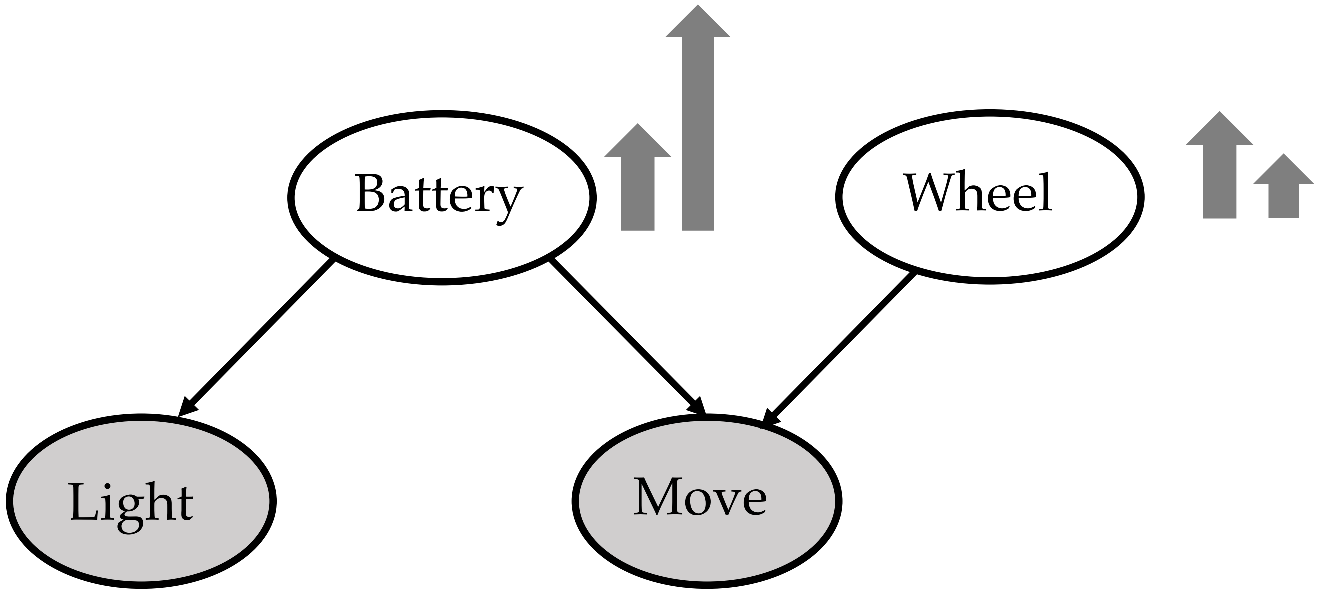
Examples



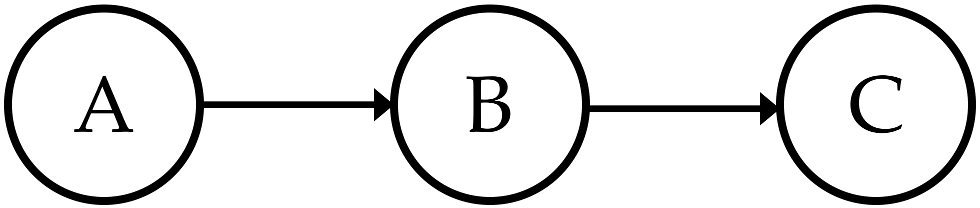
Examples



Examples

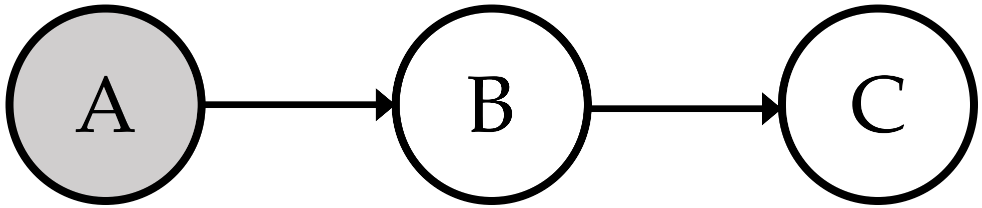


Propagation of Information: Serial Connection



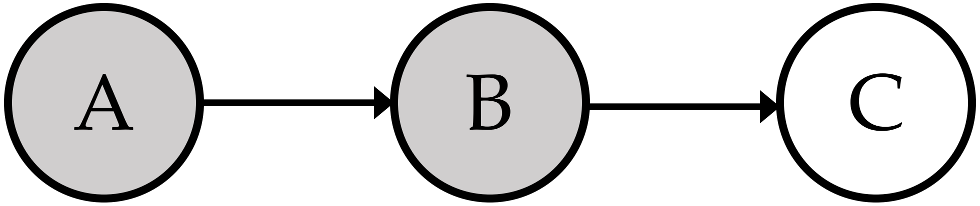
Example:
A: Battery dead
B: Robot won’t start
C: Robot won’t move

Propagation of Information: Serial Connection



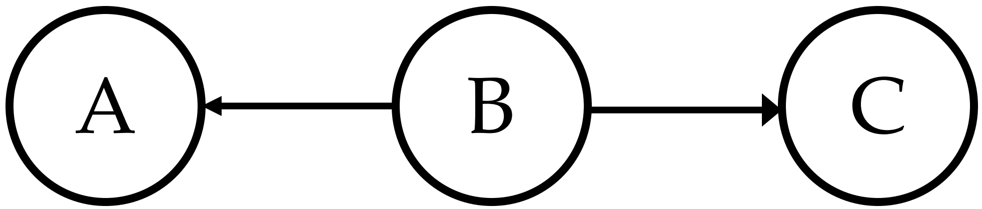
Example:
A: Battery dead
B: Robot won't start
C: Robot won't move

Propagation of Information: Serial Connection



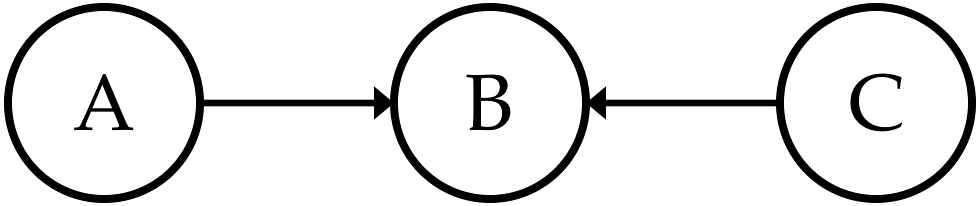
Example:
A: Battery dead
B: Robot won't start
C: Robot won't move

Propagation of Information: Diverging Connection

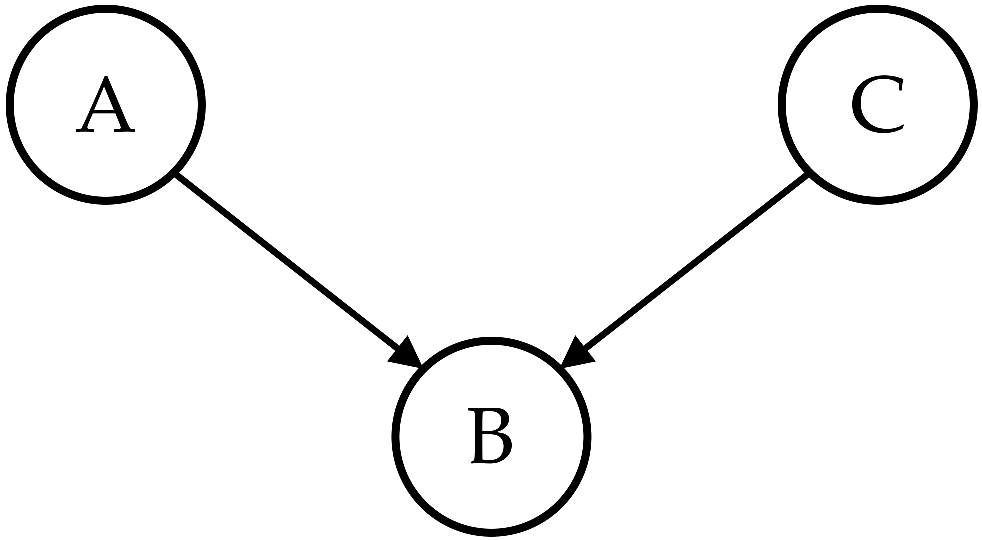


- Knowing about A will tell us something about C. Knowing about C will tell us something about A.
- If we know B propagation is blocked

Propagation of Information: Converging Connection



Propagation of Information: Converging Connection



Example:
A = wheels damaged
B = robot does not move
C= battery off

Bayesian Networks

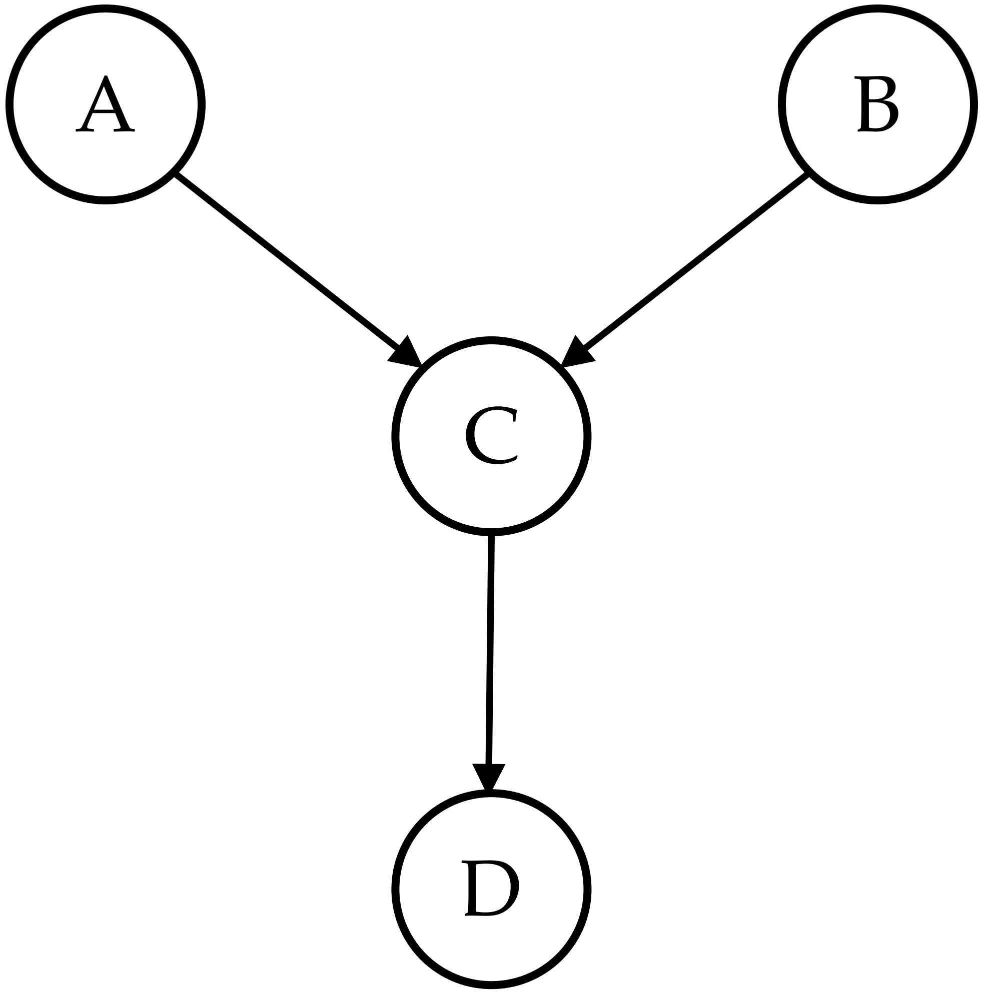
- Set of variables, set of directed arcs forming *acyclic* graphs.
- Every node A with parents $B_1, \ldots, B_n$ has $P(A|B_1, \ldots, B_n)$
- *Each variable is conditionally independent of its non-descendants, given its parents.*

Chain Rule

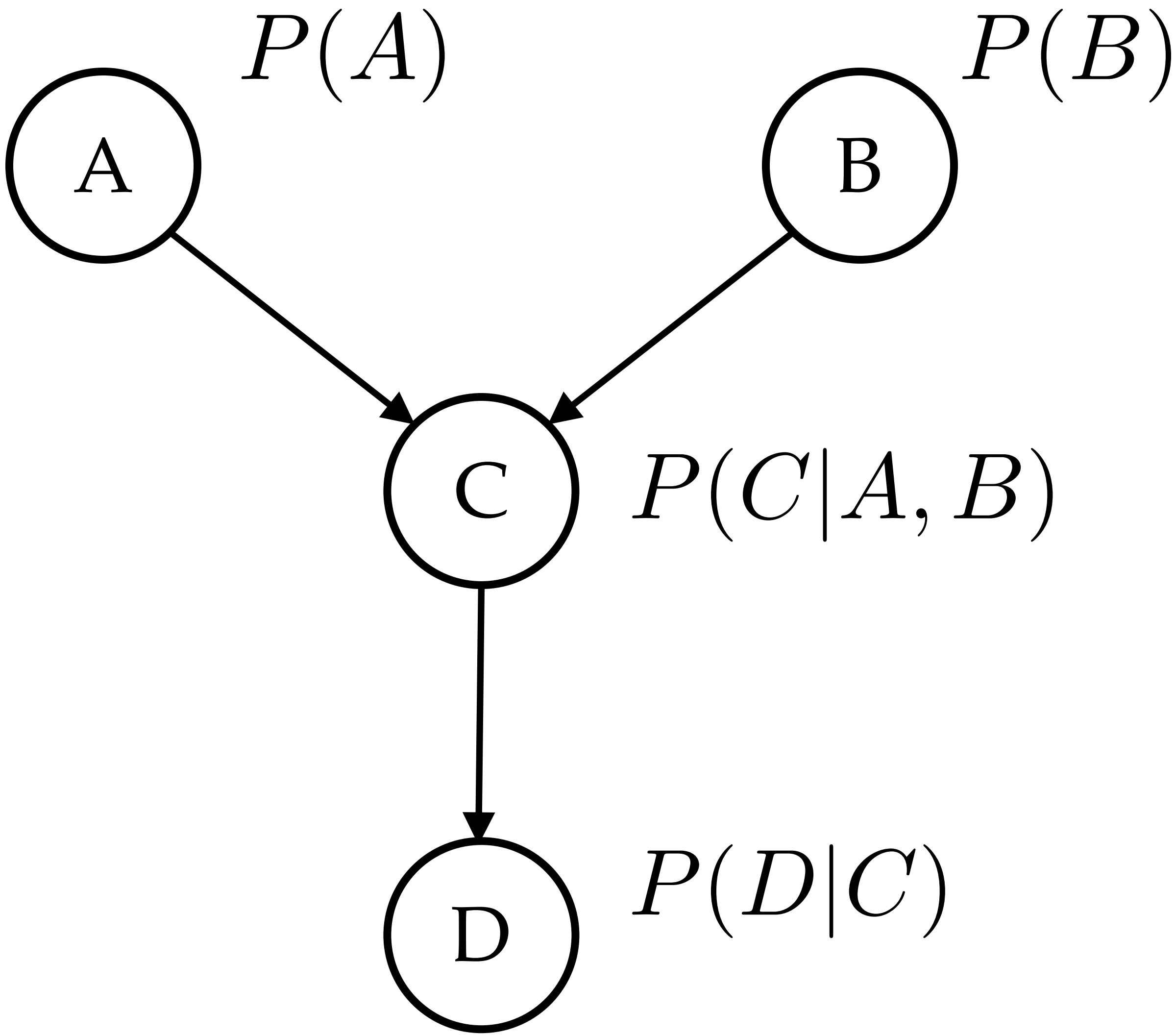
$$P(V_1 = v_1, V_2 = v_2, ..., V_n = v_n) = \Pi_i P(V_i = v_i | \text{parents}(V_i))$$

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.

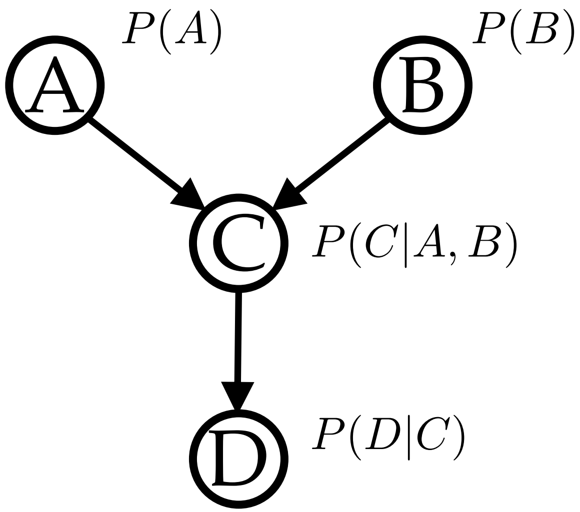
Example



Example

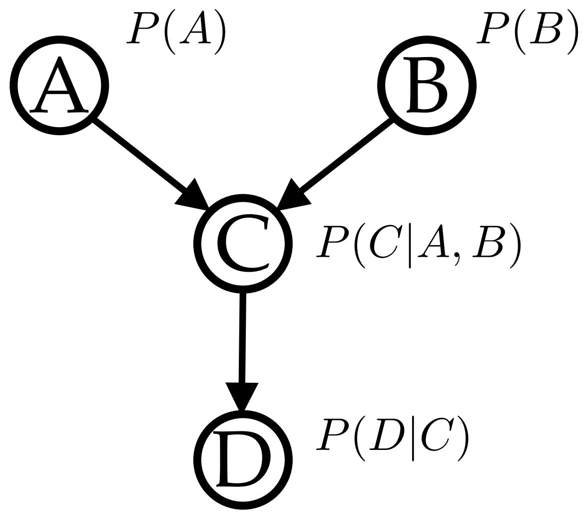


Example



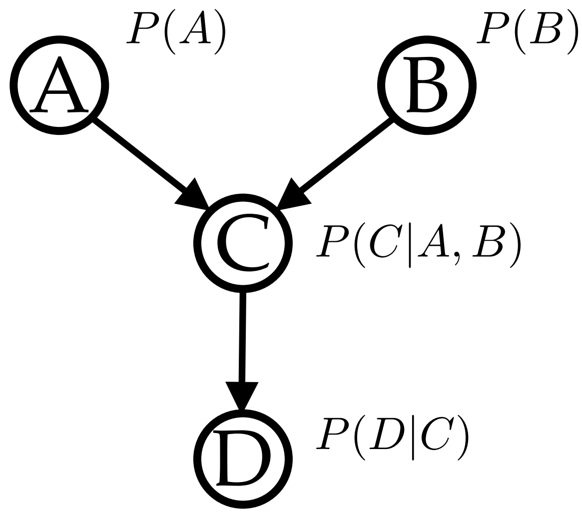
$$P(A,B,C,D) = P(A = True, B = True, C = True, D = True)$$

Example



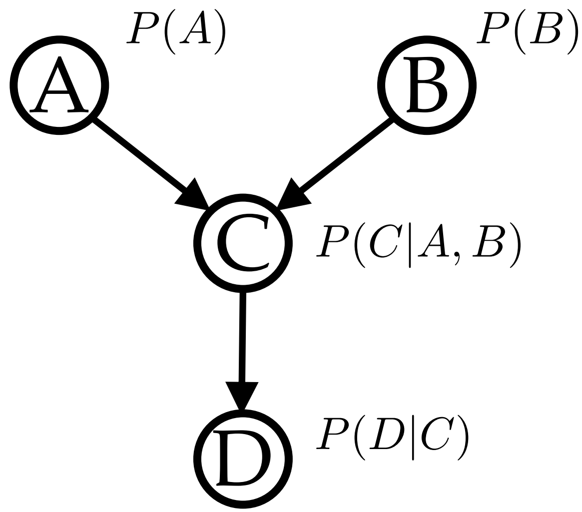
$$
\begin{aligned}
P(A,B,C,D) &= P(A=True, B=True, C=True, D=True) \\
&= P(D|A,B,C)P(A,B,C)
\end{aligned}
$$

Example



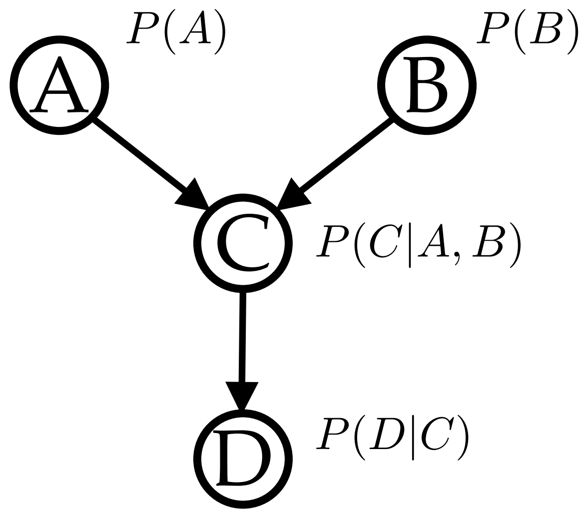
$$
\begin{aligned}
P(A,B,C,D) &= P(A=True, B=True, C=True, D=True) \\
&= P(D|A,B,C)P(A,B,C) \\
&= P(D|C)P(A,B,C)
\end{aligned}
$$

Example



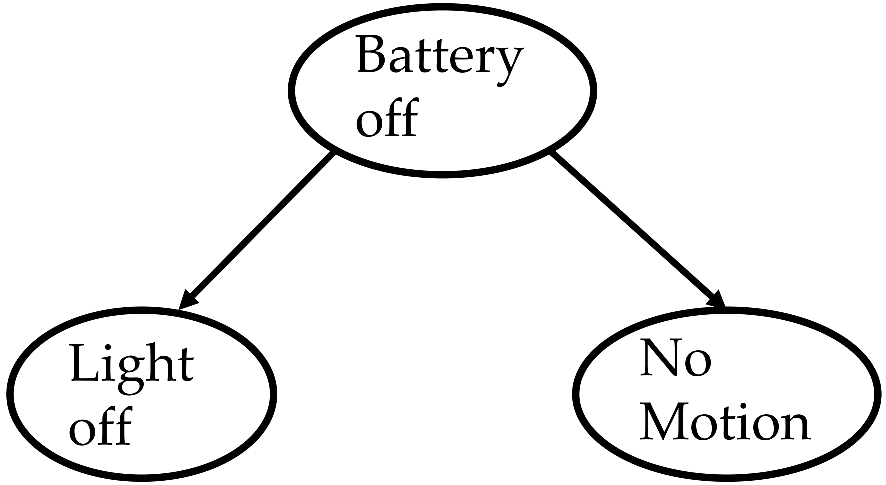
$$
\begin{aligned}
P(A, B, C, D) &= P(A = True, B = True, C = True, D = True) \\
&= P(D|A, B, C)P(A, B, C) \\
&= P(D|C)P(A, B, C) \\
&= P(D|C)P(C|A, B)P(A, B)
\end{aligned}
$$

Example

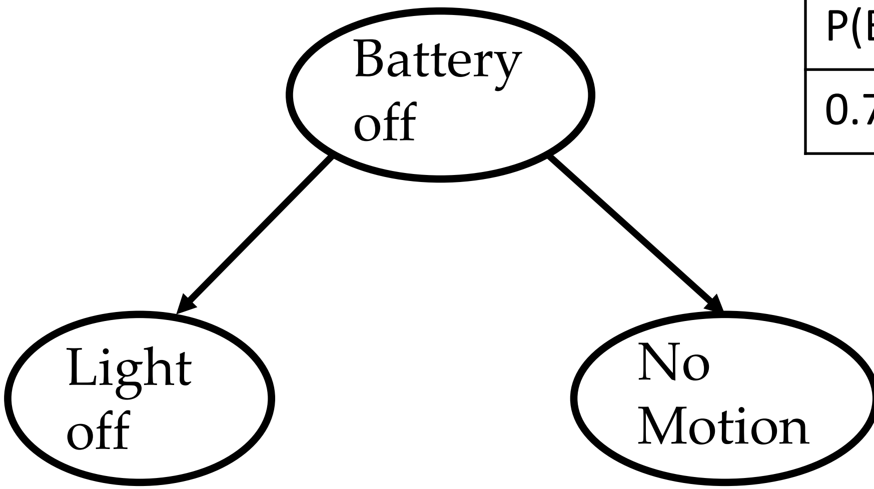


$$
\begin{aligned}
P(A,B,C,D) &= P(A=True, B=True, C=True, D=True) \\
&= P(D|A,B,C)P(A,B,C) \\
&= P(D|C)P(A,B,C) \\
&= P(D|C)P(C|A,B)P(A,B) \\
&= P(D|C)P(C|A,B)P(A)P(B)
\end{aligned}
$$

Numerical Example



Numerical Example

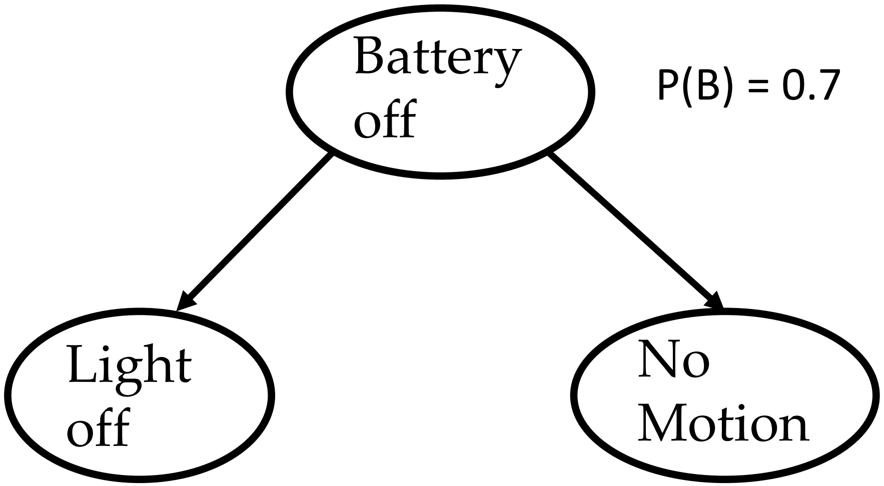


P(B=t)	P(B=f)
0.7	0.3

	P(L=t|B)	P(L=f|B)
B=t	0.8	0.2
B=f	0.1	0.9

	P(M=t|B)	P(M=f|B)
B=t	1.0	0.0
B=f	0.1	0.9

Numerical Example: Shorthand

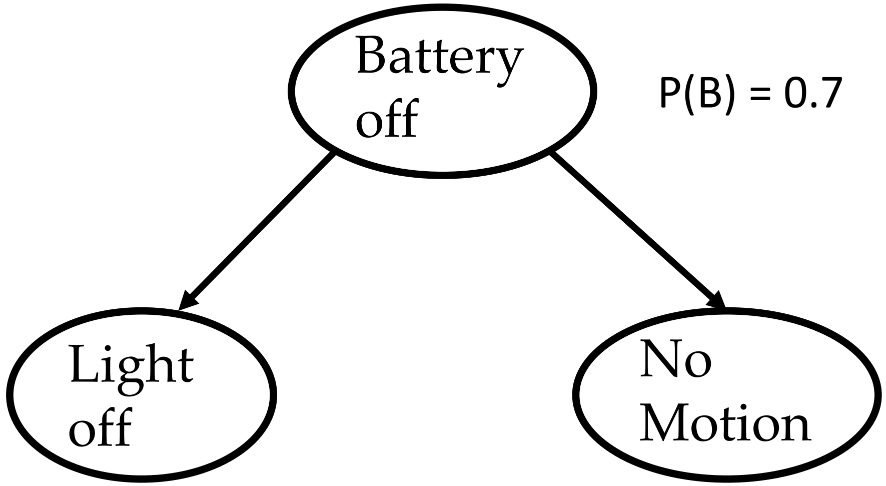


P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

Probability of No Motion

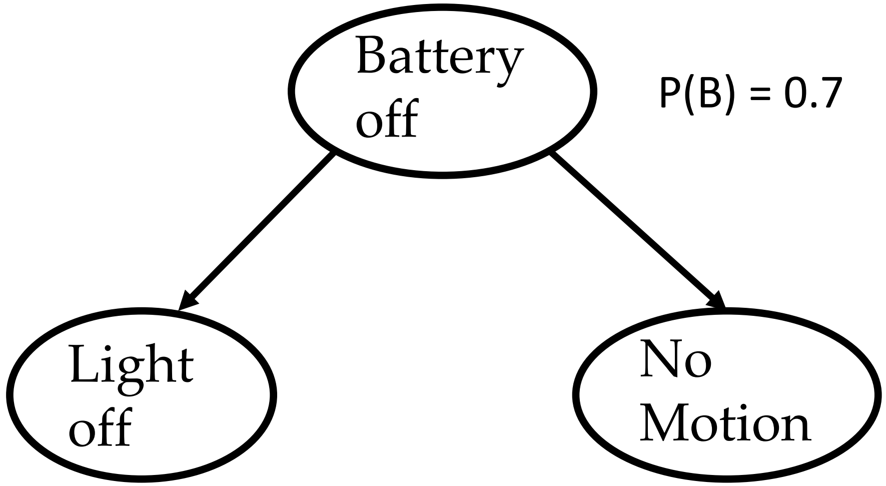


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(M) =$$

Probability of No Motion

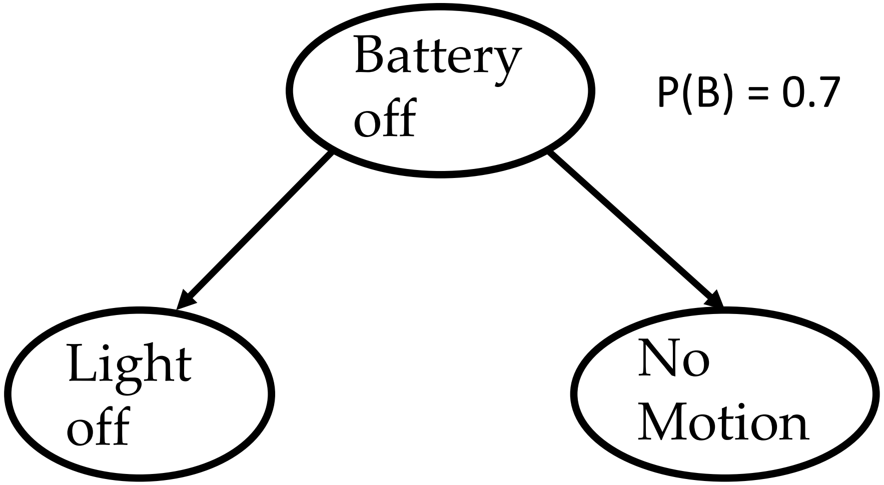


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(M) = P(M|B)P(B) + P(M|\neg B)P(\neg B)$$

Probability of No Motion

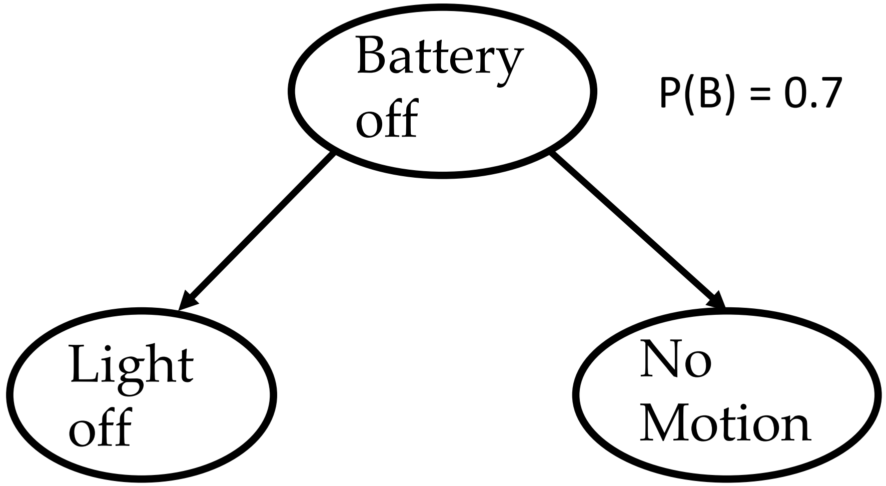


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(M) = P(M|B)P(B) + P(M|\neg B)P(\neg B)$$
$$= 1.0 * 0.7 + 0.1 * 0.3$$

Probability of No Motion

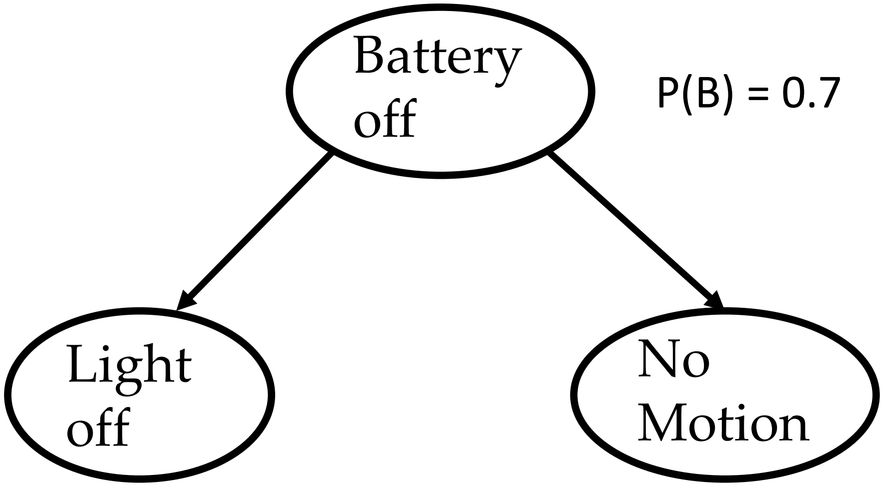


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$
\begin{aligned}
P(M) &= P(M|B)P(B) + P(M|\neg B)P(\neg B) \\
&= 1.0 * 0.7 + 0.1 * 0.3 \\
&= 0.7 + 0.03 \\
&= 0.73
\end{aligned}
$$

Probability of Light Off



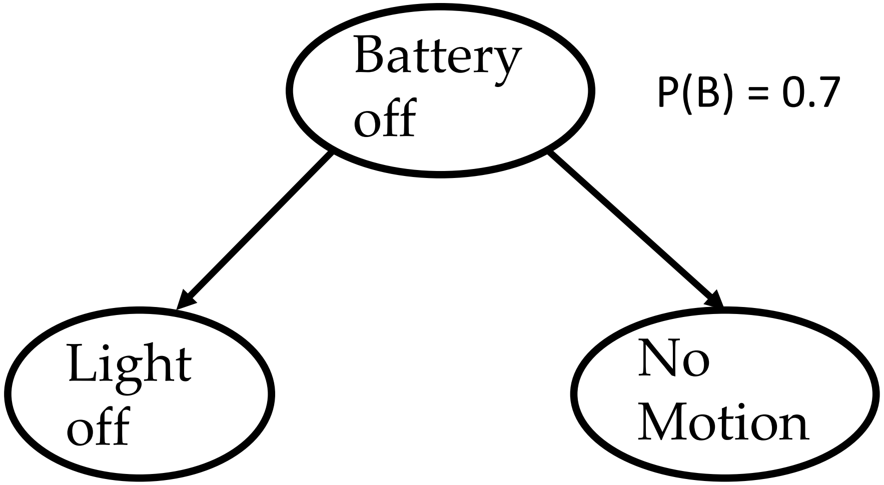
P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(L) =$$

Probability of Light Off



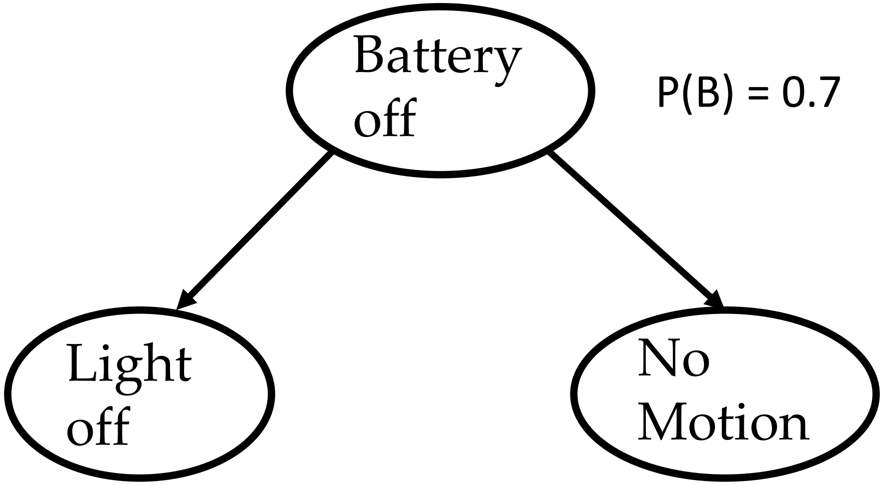
P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(L) = P(L|B)P(B) + P(L|\neg B)P(\neg B)$$

Probability of Light Off



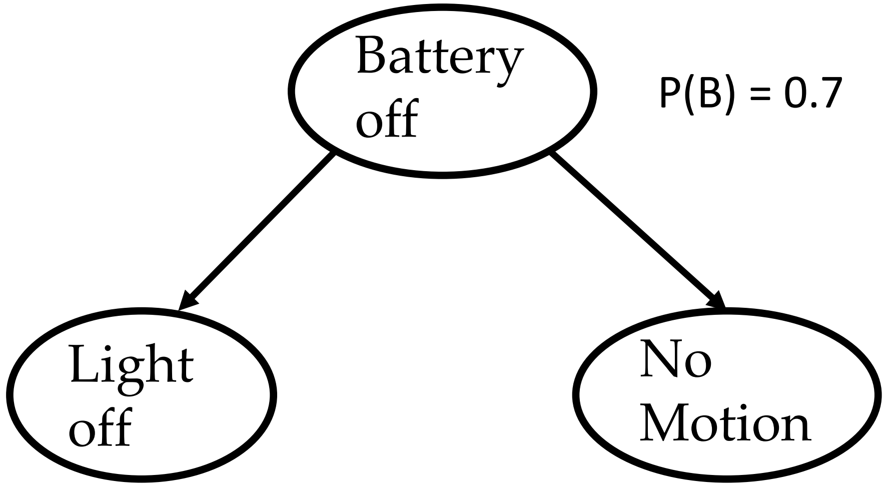
P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(L) = P(L|B)P(B) + P(L|\neg B)P(\neg B)$$
$$= 0.8 * 0.7 + 0.1 * 0.3$$

Probability of Light Off

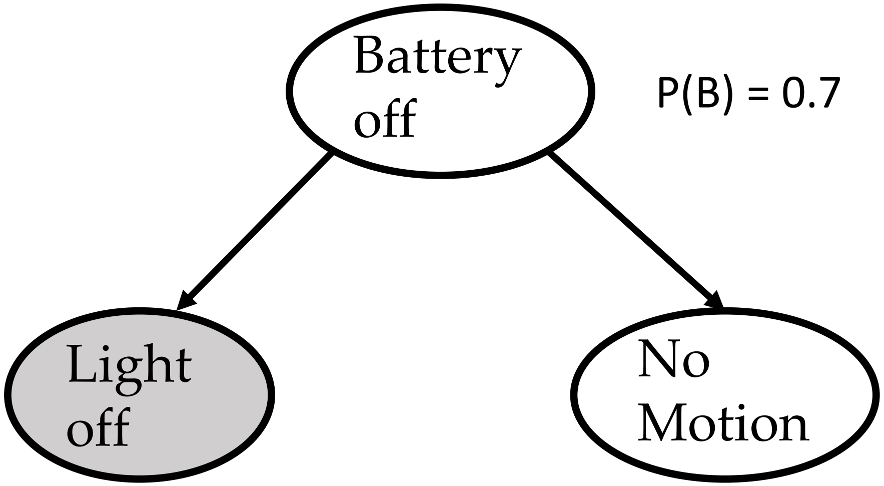


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$
\begin{aligned}
P(L) &= P(L|B)P(B) + P(L|\neg B)P(\neg B) \\
&= 0.8 * 0.7 + 0.1 * 0.3 \\
&= 0.56 + 0.03 \\
&= 0.59
\end{aligned}
$$

Probability of Battery Off

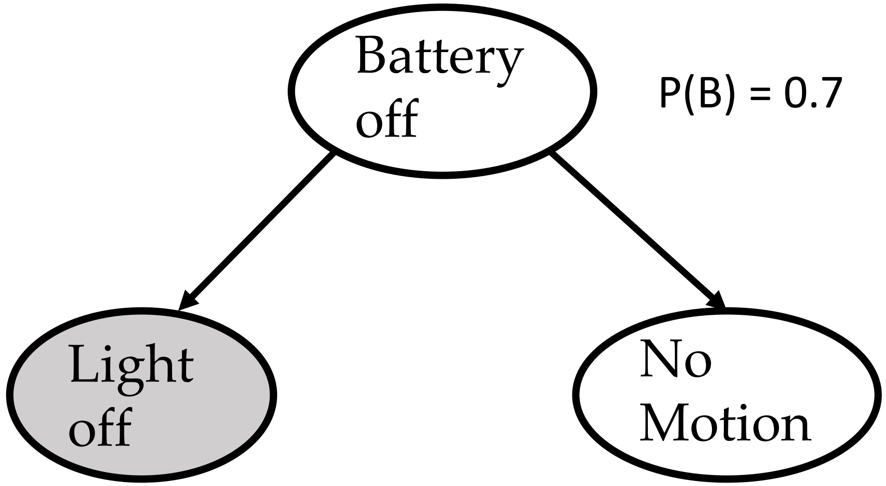


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(B|L) =$$

Probability of Battery Off

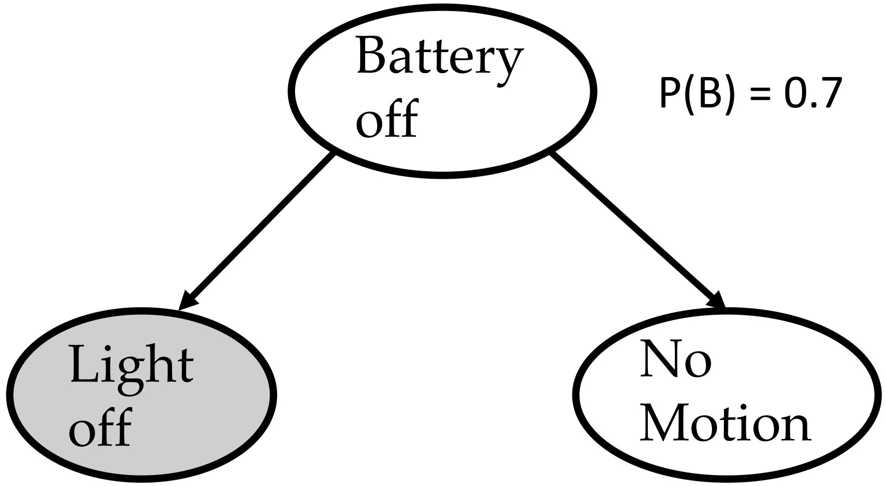


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(B|L) = \frac{P(L|B)P(B)}{P(L)}$$

Probability of Battery Off



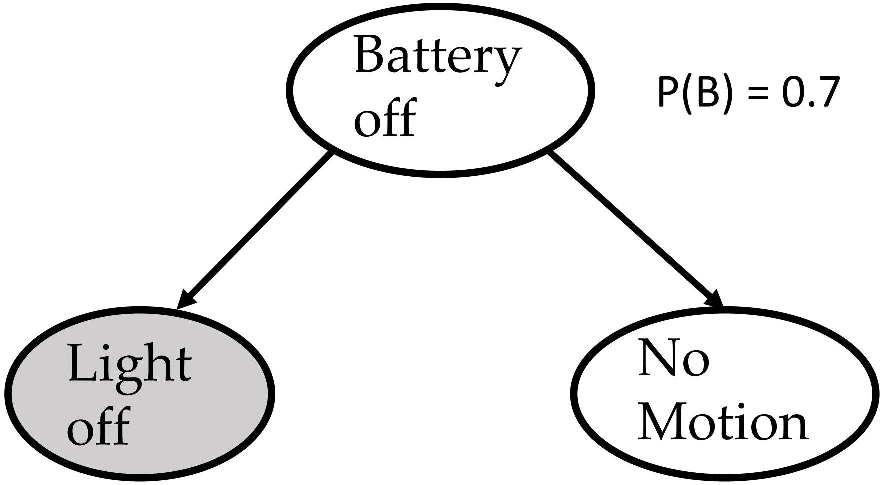
P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(B|L) = \frac{P(L|B)P(B)}{P(L)}$$
$$= \frac{0.8 * 0.7}{0.59}$$

Probability of Battery Off



	P(L|B)
B	0.8
¬B	0.1

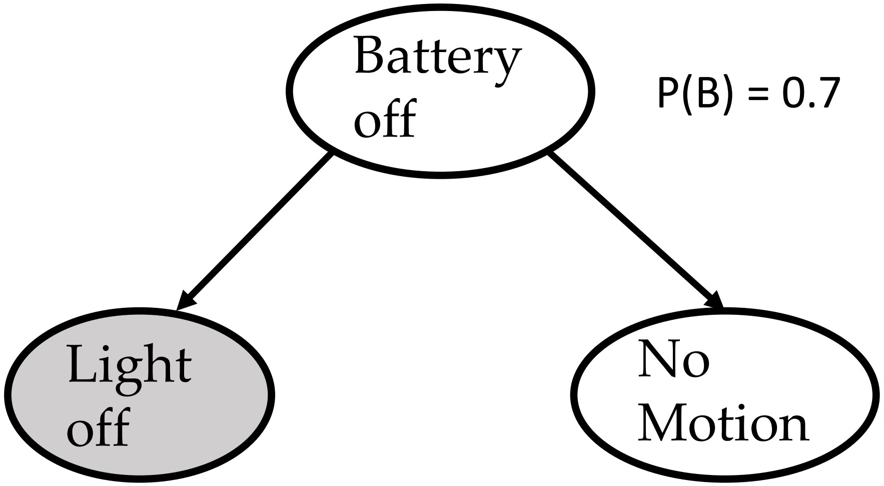
	P(M|B)
B	1.0
¬B	0.1

$$P(B|L) = \frac{P(L|B)P(B)}{P(L)}$$

$$= \frac{0.8 * 0.7}{0.59}$$

$$= 0.95$$

Probability of No Motion Given Light Off

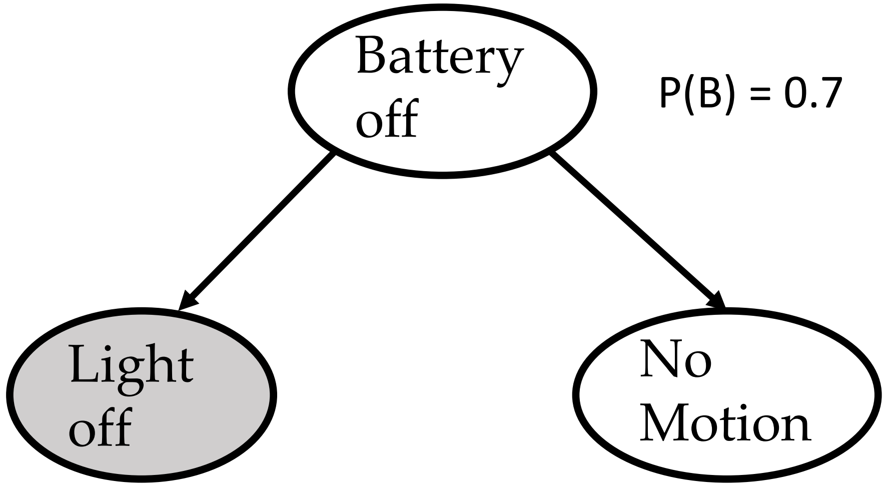


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(M|L) =$$

Probability of No Motion Given Light Off



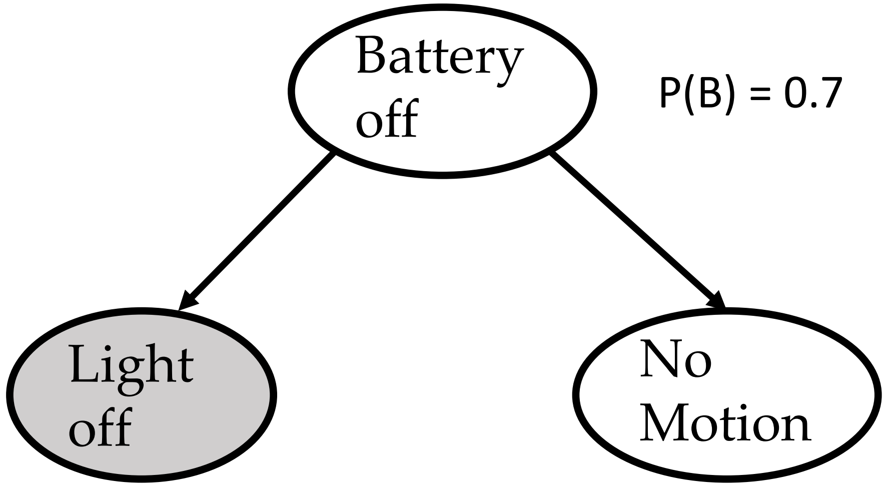
P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$P(M|L) = P(M, B|L) + P(M, \neg B|L)$$

Probability of No Motion Given Light Off



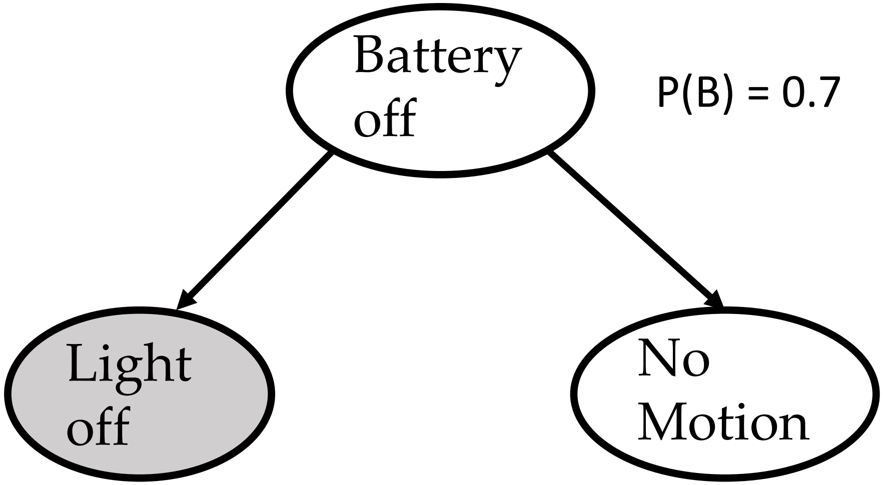
P(B) = 0.7

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$\begin{aligned} P(M|L) &= P(M, B|L) + P(M, \neg B|L) \\ &= P(M|B, L)P(B|L) + P(M|\neg B, L)P(\neg B|L) \end{aligned}$$

Probability of No Motion Given Light Off

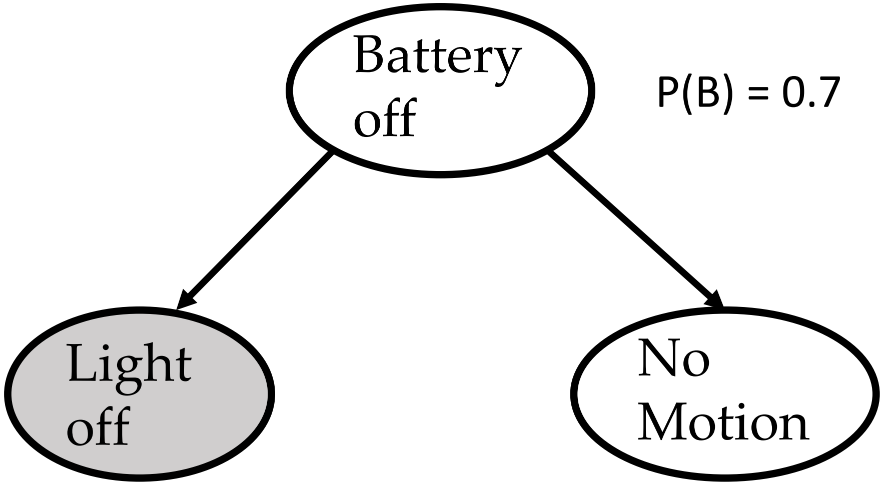


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$
\begin{aligned}
P(M|L) &= P(M, B|L) + P(M, \neg B|L) \\
&= P(M|B, L)P(B|L) + P(M|\neg B, L)P(\neg B|L) \\
&= P(M|B)P(B|L) + P(M|\neg B)P(\neg B|L)
\end{aligned}
$$

Probability of No Motion Given Light Off

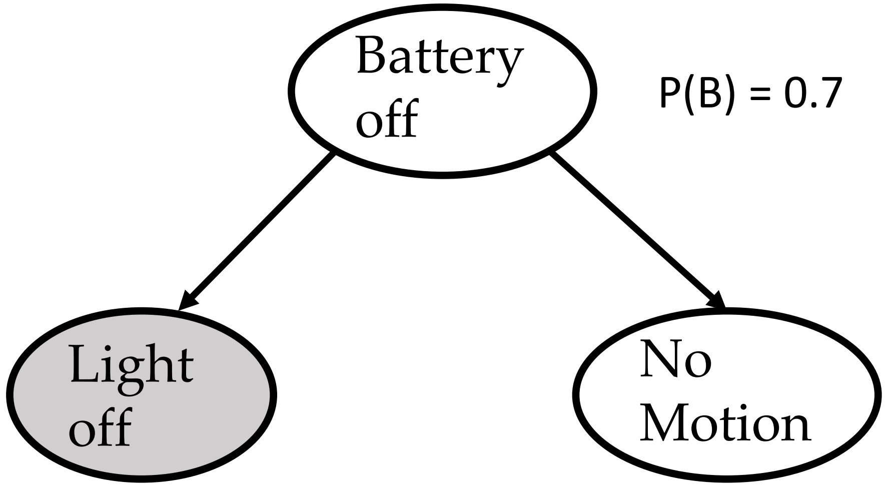


	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

$$\begin{aligned} P(M|L) &= P(M,B|L) + P(M,\neg B|L) \\ &= P(M|B,L)P(B|L) + P(M|\neg B,L)P(\neg B|L) \\ &= P(M|B)P(B|L) + P(M|\neg B)P(\neg B|L) \\ &= 1.0 * 0.95 + 0.1 * 0.05 \\ &= 0.95 + 0.005 \\ &= 0.96 \end{aligned}$$

Probability of No Motion Given Light Off



P(B) = 0.7

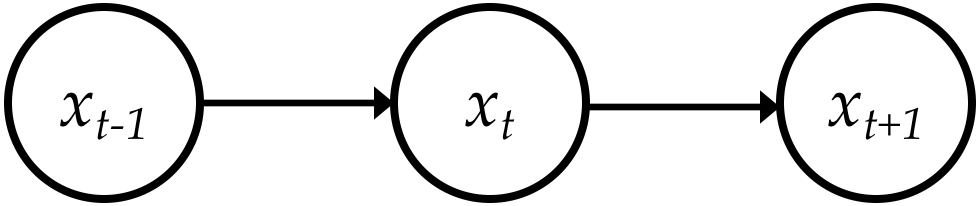
$P(M) = 0.73$

	P(L|B)
B	0.8
¬B	0.1

	P(M|B)
B	1.0
¬B	0.1

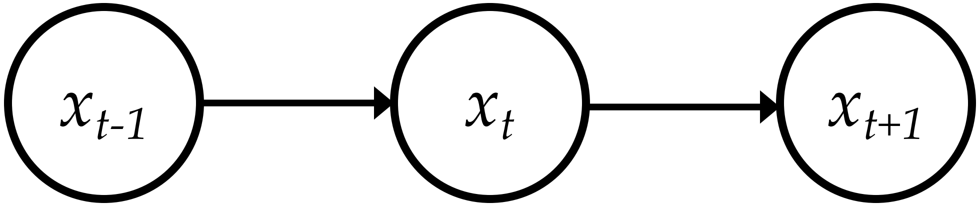
$$
\begin{aligned}
P(M|L) &= P(M, B|L) + P(M, \neg B|L) \\
&= P(M|B, L)P(B|L) + P(M|\neg B, L)P(\neg B|L) \\
&= P(M|B)P(B|L) + P(M|\neg B)P(\neg B|L) \\
&= 1.0 * 0.95 + 0.1 * 0.05 \\
&= 0.95 + 0.005 \\
&= 0.96
\end{aligned}
$$

State Estimation Over Time



What type of connection?

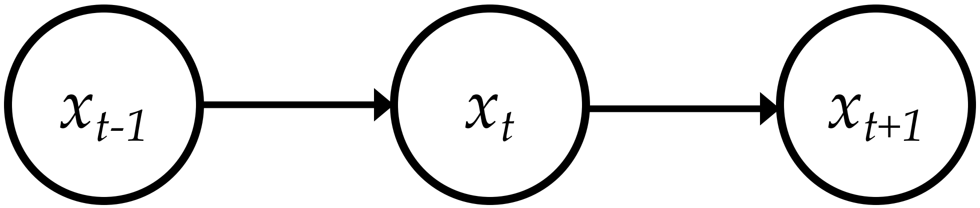
State Estimation Over Time



What type of connection?

- Serial Connection!

State Estimation Over Time



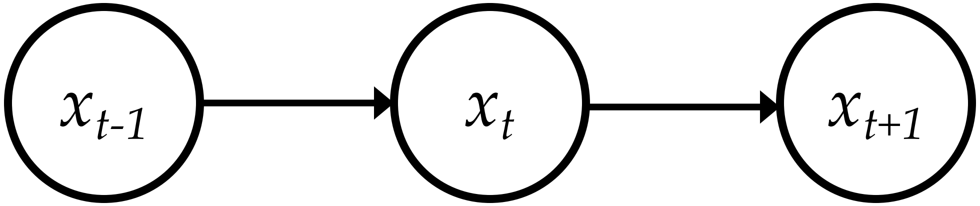
What type of connection?

- Serial Connection!

State Transitions:

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

State Estimation Over Time



What type of connection?

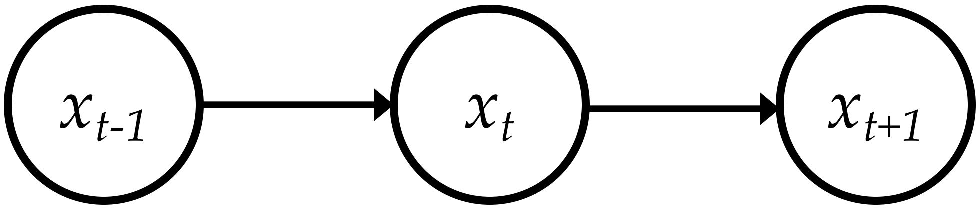
- Serial Connection!

State Transitions:

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

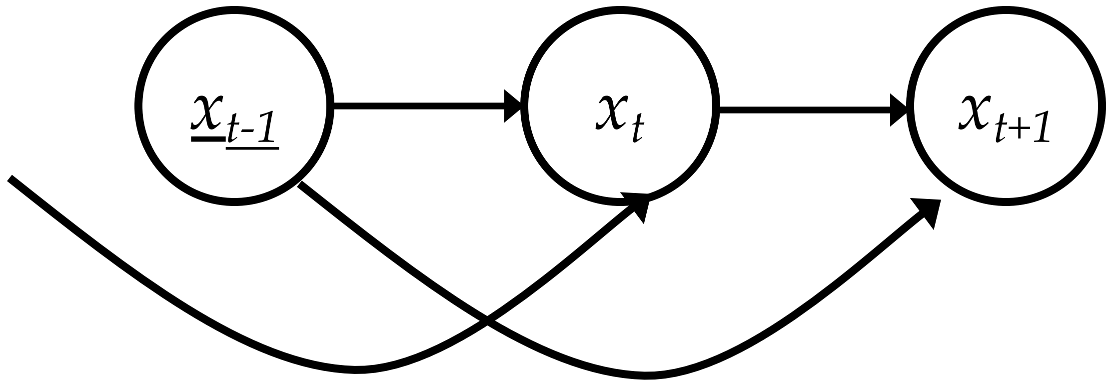
Markov Assumption

Second Order Model



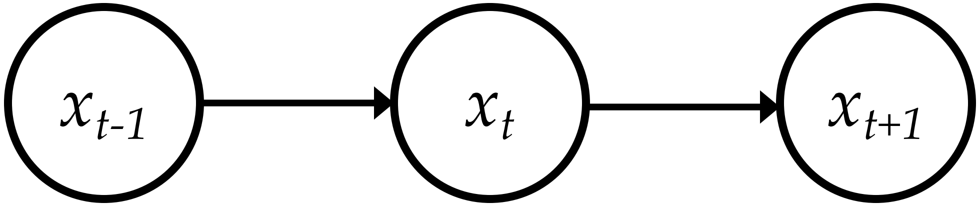
$$P(X_t|X_{0:t-1}) = P(X_t|X_{t-1}, X_{t-2})$$

Second Order Model



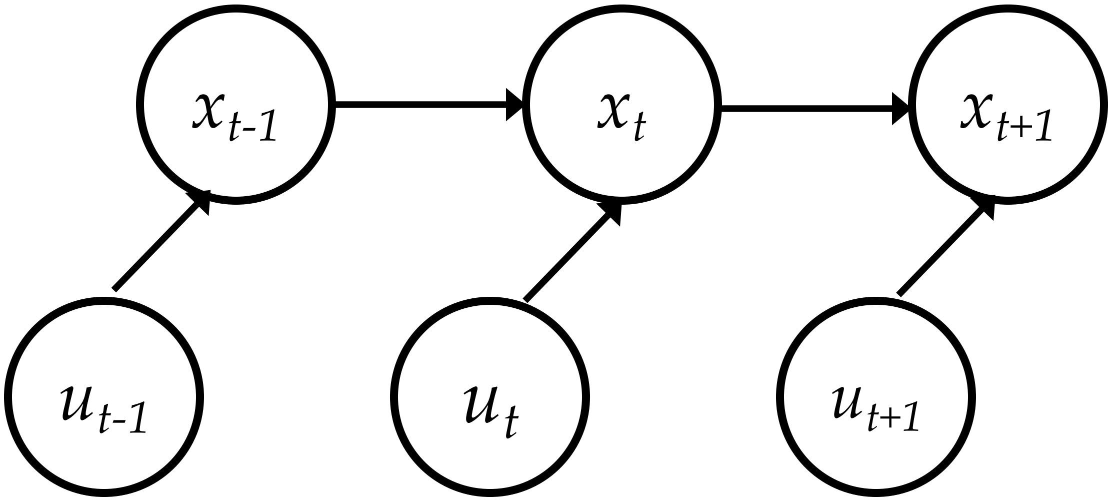
$$P(X_t|X_{0:t-1}) = P(X_t|X_{t-1}, X_{t-2})$$

Assumption of Stationary Dynamics



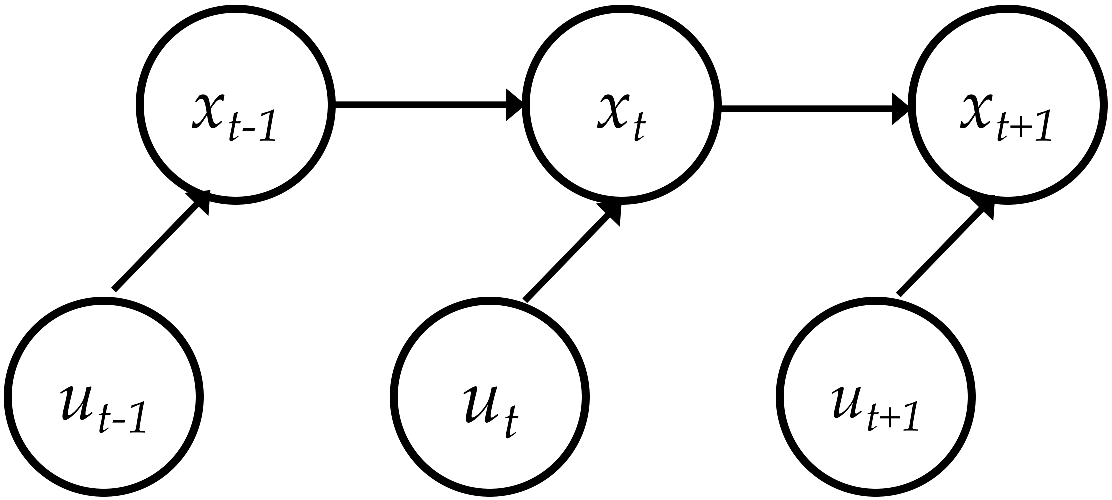
$$P(X_t|X_{t-1}) = P(X_{t-1}|X_{t-2})$$

Control Actions



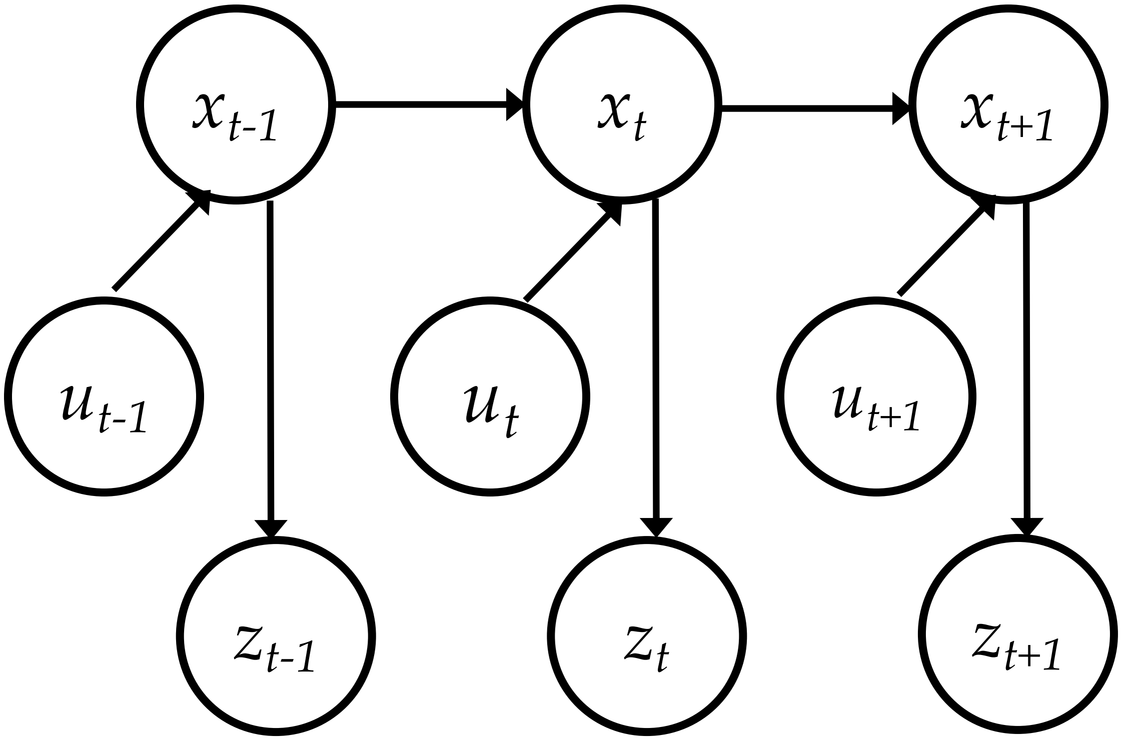
$$P(X_t | X_{0:t-1}, U_{0:t-1}) =$$

Control Actions



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

Sensor Model



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