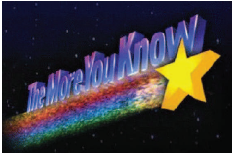
The More you Know - The Information Gain Approach to Path Planning (Part 2)

Liam Paull



Outline

Outline

Conditional Probabilities

We can define probabilities of one RV conditional on another.

e.g:

$$
\begin{aligned}
A &= 1,2,3,4,5,6 \\
B &= \text{even}, \text{odd}
\end{aligned}
$$

$$P(A=2|B=even)=1/3 \quad (1)$$

Outline

Bayes' Theorem

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2}$$

Simple Example: (wikipedia) School with 60% boys and 40% girls. All of the boys wear pants, the girls wear pants or skirts in equal proportion.

Bayes' Theorem

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2}$$

Simple Example: (wikipedia) School with 60% boys and 40% girls. All of the boys wear pants, the girls wear pants or skirts in equal proportion.

Question: You see someone wearing pants, what is that probability that it's a girl?

Bayes' Theorem

Bayes' Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2}$$

Simple Example: (wikipedia) School with 60% boys and 40% girls. All of the boys wear pants, the girls wear pants or skirts in equal proportion.

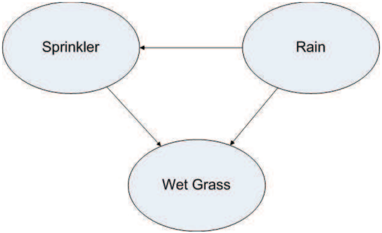
Question: You see someone wearing pants, what is that probability that it's a girl?

A: Person observed is a girl
B: Person ovserved is wearing pants

$$P(A|B) = \frac{0.5 \times 0.4}{0.8} = 0.25 \tag{3}$$

Bayesian Networks - A Simple Example

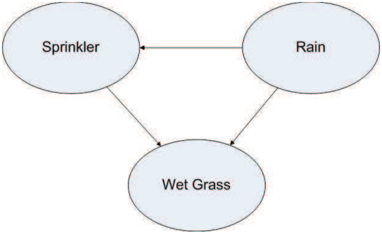
The arrows designate conditionality relationships



$$P(GW, S, R) = P(GW|S, R)P(S|R)P(R) \quad (4)$$

Bayesian Networks - A Simple Example

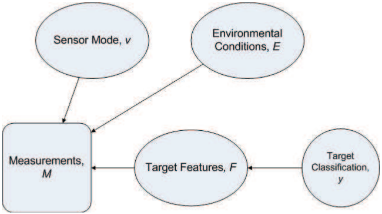
The arrows designate conditionality relationships



$$P(GW, S, R) = P(GW|S, R)P(S|R)P(R) \tag{4}$$

Question: What is the probability that the sprinkler is off given that the grass is wet?

Bayesian Networks - A More Complex Example



$$P(v, E, M, F, y) = P(M|v, E, F)P(F|y)P(y)P(v)P(E) \quad (5)$$

Outline

A Simple Objective Function

Need a way to evaluate all of the potential moves to decide which one is best

e.g.

$$R(t_k) = w_B \cdot B(t_k) - w_J \cdot J(t_k) - w_D \cdot D(t_k) \qquad (6)$$

Where

$R(t_k)$ is the *measurement profit*

$B(t_k)$ is the *information gain*

$J(t_k)$ is the power used

$D(t_k)$ is the distance travelled

t_k is the next timestep

Outline

Entropy

Idea: Entropy is a measure of how much uncertainty there is in the system

Entropy

Idea: Entropy is a measure of how much uncertainty there is in the system

DEFINITION: Entropy of an RV $X = \{x_1, x_2, ...x_n\}$:

$$H(X) = -\sum_{i=1}^{n} P(X = x_i) \log_2 P(X = x_i) \tag{7}$$

Entropy

Idea: Entropy is a measure of how much uncertainty there is in the system

DEFINITION: Entropy of an RV $X = \{x_1, x_2, ...x_n\}$:

$$H(X) = -\sum_{i=1}^{n} P(X = x_i) \log_2 P(X = x_i) \tag{7}$$

E.g. the entropy of a fair die:

$$\begin{aligned} H(X) &= -\sum_{i=1}^{6} P(X = x_i) \log_2 P(X = x_i) \\ &= -6 * (1/6 \log_2 1/6) \\ &= \log_2 6 = 2.585 \end{aligned}$$

Conditional Entropy

Can also define conditional entropy of an RV X given Y ($Y = \{y_1, y_2, ..., y_m\}$):

$$H(X|Y) = \sum_{j=1}^{m} P(Y = y_j) H(X|Y = y_j) \tag{8}$$

$$= -\sum_{j=1}^{m} \sum_{i=1}^{n} P(X = x_i, Y = y_j) \log_2 P(X = x_i | Y = y_j)$$

Conditional Entropy

Can also define conditional entropy of an RV X given Y ($Y = \{y_1, y_2, ..., y_m\}$):

$$H(X|Y) = \sum_{j=1}^{m} P(Y = y_j) H(X|Y = y_j) \tag{8}$$

$$= -\sum_{j=1}^{m} \sum_{i=1}^{n} P(X = x_i, Y = y_j) \log_2 P(X = x_i | Y = y_j)$$

Question: What is the entropy of the result of rolling a fair die if we know whether the answer is odd or even?

Mutual Information or Entropy Reduction

Now that we have a measure of the uncertainty that is present in the system, we have a way of defining how much the uncertainty is reduced by a measurement or set of measurements:

Mutual Information or Entropy Reduction

Now that we have a measure of the uncertainty that is present in the system, we have a way of defining how much the uncertainty is reduced by a measurement or set of measurements:

DEFINITION: Mutual Information or Entropy Reduction

$$I(X; Y_1 | Y_2) = H(X | Y_1) - H(X | Y_1, Y_2) \tag{9}$$

This equation defines the entropy reduction brought about by Y_2 given what we already knew: Y_1

Mutual Information or Entropy Reduction

Now that we have a measure of the uncertainty that is present in the system, we have a way of defining how much the uncertainty is reduced by a measurement or set of measurements:

DEFINITION: Mutual Information or Entropy Reduction

$$I(X;Y_1|Y_2) = H(X|Y_1) - H(X|Y_1,Y_2) \tag{9}$$

This equation defines the entropy reduction brought about by Y_2 given what we already knew: Y_1

KEY ADVANTAGE: It is additive!

Cumulative Expected Entropy Reduction

Now we can define the expected entropy reduction that will be brought about by a series of measurements (i.e. a move by the robot) and add them all together and put them into our objective function.

Cumulative Expected Entropy Reduction

Now we can define the expected entropy reduction that will be brought about by a series of measurements (i.e. a move by the robot) and add them all together and put them into our objective function.

DEFINITION: Information Gain

Let $Z = \{M_1, M_2, ..., M_k\}$ be the set of measurements
Let $\epsilon_i = \{v_i, E_i, ...\}$ be all the other variables in the BN which can be known or estimated corresponding to measurement M_i

$$B(Z) = \sum_{M_i \in Z} I(y_i; M_i | \epsilon_i) \tag{10}$$

Cumulative Expected Entropy Reduction

Now we can define the expected entropy reduction that will be brought about by a series of measurements (i.e. a move by the robot) and add them all together and put them into our objective function.

DEFINITION: Information Gain

Let $Z = \{M_1, M_2, ..., M_k\}$ be the set of measurements
Let $\epsilon_i = \{v_i, E_i, ...\}$ be all the other variables in the BN which can be known or estimated corresponding to measurement M_i

$$B(Z) = \sum_{M_i \in Z} I(y_i; M_i | \epsilon_i) \tag{10}$$

Now we can evaluate the information gain for each cell and make a decision based on the objective function.

Outline

Probability of Detection

A less computationally intensive alternative to information gain. In the case when the goal is to find targets (e.g. mine hunting), the Probability of Detection defines the probability that a group of measurements will detect a target

Probability of Detection

A less computationally intensive alternative to information gain. In the case when the goal is to find targets (e.g. mine hunting), the Probability of Detection defines the probability that a group of measurements will detect a target

Let each $D_j \in \{d, \bar{d}\}, j = 1..J$ be detection events resulting from sensor measurements

Let $Y = \{y_1, y_2, ..., y_k\}$ be target states (i.e. different targets)

Probability of Detection

A less computationally intensive alternative to information gain. In the case when the goal is to find targets (e.g. mine hunting), the Probability of Detection defines the probability that a group of measurements will detect a target

Let each $D_j \in \{d, \bar{d}\}, j = 1..J$ be detection events resulting from sensor measurements

Let $Y = \{y_1, y_2, ..., y_k\}$ be target states (i.e. different targets)

$$P.O.D. = 1 - \sum_{i=1}^{k} P(Y = y_i) \prod_{j=1}^{J} P(D_j = \bar{d} | Y = y_i) \quad (11)$$