Table of Contents I

Probabilistic Reasoning

- Classical Probabilistic Models
- Basic Probabilistic Reasoning: The Jungle Story
- Multiple Random Selection Rules: Dice
- New Constructs
  - Causal Probability
  - Observations and Interventions
  - Dynamic Range
- Representing Knowledge in P-log

Probabilistic Reasoning: A Finer Gradation of Unknowns

- Defaults allowed us to work with incomplete information.
- Multiple answer sets helped model different possibilities.
- Example 1:

$$p(a) \text{ or } \neg p(a)$$

- Example 2:

$$q(a). \quad q(b). \quad p(b).$$

- In both cases, $p(a)$ is unknown.
- In ASP, propositions could only have three truth values: *true*, *false*, and *unknown*.
- How can we say that "we're pretty sure $p(a)$ is true" without losing our ability to use defaults, nonmonotonicity, recursion, etc. — everything gained by using ASP?

Old Methods, New Reading, New Use

- Probability theory is a well-developed branch of mathematics.
- How do we use it for knowledge representation?
- If we do use it, what do we really mean?
- We will view probabilistic reasoning as *commonsense reasoning about the degree of an agent's beliefs in the likelihood of different events*.
- "There's a fifty-fifty chance." "I'm 99% sure."
- This is known as the **Bayesian view**.

Consequences of the Bayesian View

- Example: the agent's knowledge about whether a particular bird flies will be based on what it knows of the bird, rather than the statistics that apply to the whole population of birds in general.
- A different agent's measure may be different because its knowledge of the bird is different.
- Note that this means that an agent's belief about the probability of an event can change based on the knowledge it has.

Lost in the Jungle

> *Imagine yourself lost in a dense jungle. A group of natives has found you and offered to help you survive, provided you can pass their test. They tell you they have an Urn of Decision from which you must choose a stone at random. (The urn is sufficiently wide for you to easily get access to every stone, but you are blindfolded so you cannot cheat.) You are told that the urn contains nine white stones and one black stone. Now you must choose a color. If the stone you draw matches the color you chose, the tribe will help you; otherwise, you can take your chances alone in the jungle. (The reasoning of the tribe is that they do not wish to help the exceptionally stupid, or the exceptionally unlucky.)*

What is your reasoning about the color you should choose?

Example Train of Thought

> *Suppose I choose white. What would be my chances of getting help? They are the same as the chances of drawing a white stone from the urn. There are nine white stones out of a possible ten. Therefore, my chances of picking a white stone and obtaining help are* $\frac{9}{10}$.

The number $\frac{9}{10}$ can be viewed as the degree of belief that help will be obtained if you select white.

Using a Probabilistic Model I

- **Probabilistic models** consist of a finite set Ω of possible worlds and a probabilistic measure μ.
- **Possible worlds** correspond to possible outcomes of random experiments we attempt to perform (like drawing a stone from the urn).
- The **probabilistic measure** $\mu(W)$ quantifies the agent's degree of belief in the likelihood of the outcomes of random experiments represented by W.

Using a Probabilistic Model II

- The probabilistic measure is a function μ from Ω to 2^{Ω} such that:

$$\text{for all } W \in \Omega,\ \mu(W) \geq 0 \text{ and}$$

$$\sum_{W \in \Omega} \mu(W) = 1.$$

Possible Worlds in Logic-Based Theory

- In logic-based probability theory, possible worlds are often identified with logical interpretations.
- A set E of possible worlds is often represented by a formula F such that $W \in E$ iff W is a model of F.
- In this case the probability function may be defined on propositions

$$P(F) =_{def} P(\{W : W \in \Omega \text{ and } W \text{ is a model of } F\}).$$

Back to the Jungle

- How do we construct a mathematical model of the reasoning behind the stone choice?
- We need to come up with a collection Ω of possible worlds that correspond to possible outcomes of this random experiment.
- Let's enumerate the stone from 1 to 10 starting with the black stone.

Jungle: Possible Worlds

- The possible world describing the effect of the traveler drawing stone number 1 from the urn looks like this:

$$W_1 = \{select\_color = white, draw = 1, \neg help\}.$$

- Drawing the second stone results in possible world

$$W_2 = \{select\_color = white, draw = 2, help\}$$

etc.

- We have 10 possible worlds, 9 of which contain *help*.

The Principle of Indifference

How do we define the probabilistic measure μ on these possible worlds?

- **Principle of Indifference** is a commonsense rule which states that *possible outcomes of a random experiment are assumed to be equally probable if we have no reason to prefer one of them to any other*.
- This rule suggest that $\mu(W) = \frac{1}{10} = 0.1$ for any possible world $W \in \Omega$.
- According to our definition of probability function P, the probability that the outcome of the experiment contains *help* is 0.9.
- A similar argument for the case in which the traveler selects black gives 0.1.
- Thus, we get the expected result.

Creating a Mathematical Model of the Argument

- The hard part of the reasoning is setting up a probabilistic model, especially the selection of possible worlds.
- Key question: *How can possible worlds of a probabilistic model be found and represented?*
- One solution is to use P-log — an extension of ASP and/or CR-Prolog that allows us to combine logical and probabilistic knowledge.
- Answer sets of a P-log program are identified with possible worlds of the domain.

Jungle Story in P-log: Signature

- P-log has a sorted signature.
- Program Π_{jungle} has two sorts: *stones* and *colors*:

$$stones = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}.$$

$$colors = \{black, white\}.$$

Jungle Story in P-log: Mapping Stones to Colors

$$
\begin{aligned}
&color(1) = black. \\
&color(X) = white \quad \leftarrow \quad X \neq 1.
\end{aligned}
$$

Note that the only difference between rules of P-log and ASP is the form of the atoms.

Jungle Story in P-log: Representing the Draw

$$draw : stones.$$
$$random(draw).$$

1. *draw* is a zero-arity function that takes its values from sort *stones*.
2. *random*(*draw*) states that, *normally*, the values for *draw* are selected at random. (**random selection rule**)

Jungle Story in P-log: Tribal Laws

$$select\_color : colors$$

$$help : boolean$$

$$\begin{array}{lcl} help & \leftarrow & draw = X, \\ & & color(X) = C, \\ & & select\_color = C. \\ \neg help & \leftarrow & draw = X, \\ & & color(X) = C, \\ & & select\_color \neq C. \end{array}$$

Here *help* and $\neg$*help* are used as shorthands for $help = true$ and $help = false$.

Jungle Story in P-log: Selecting White

To ask

"Suppose I choose white. What would be my chances of getting help?"

add the following statement to the program:

$$select\_color = white.$$

Jungle Story in P-log: Possible Worlds

- Each possible outcome of random selection for *draw* defines one possible world.
- If the result of our random selection were 1, then the relevant atoms of this world would be

$$W_1 = \{draw = 1, select\_color = white, \neg help\}$$

- Since $color(1) = black$ and $select\_color = white$ are facts of the program, the result follows immediately from the definition of *help*.
- If the result of our random selection were 2, then the world determined by this selection would be

$$W_2 = \{draw = 2, select\_color = white, help\}.$$

- Similarly for stones 3 to 10.

Jungle Story in P-log: Computing the Probability of an Event

- The semantics of P-log uses the Indifference Principle to automatically compute the probabilistic measure of every possible world and hence the probabilities of the corresponding events.
- Since in this case all worlds are equally plausible, the ratio of possible worlds in which arbitrary statement F is true to the number of all possible worlds gives the probability of F.
- Hence the probability of *help* defined by the program $\Pi_{jungle}(white)$ is $\frac{9}{10}$.

P-log: Computing Probabilities

- Collections of atoms from answer sets of $\tau(\Pi)$ are called **possible worlds** of Π.
- The probabilistic measure in P-log is a real number from the interval $[0, 1]$, which represents the degree of a reasoner's belief that a possible world W matches a true state of the world.
- Zero means that the agent believes that the possible world does not correspond to the true state; one corresponds to the certainty that it does.
- The probability of a set of possible worlds is the sum of the probabilistic measures of its elements.
- The **probability of a proposition** is the sum of the probabilistic measures of possible worlds in which this proposition is true.

Dice: The Problem

How do we define a probabilistic measure if there is more than one random selection rule?

> *Mike and John each own a die. Each die is rolled once. We would like to estimate the chance that the sum of the rolls is high, i.e. greater than 6.*

- Let's construct program Π_{dice}.
- What are our objects? dice, score, people.
- What are our relations? roll a die, get a random score, owner of a die, high (boolean)

Dice: Sort Declarations

The corresponding declarations look like this:

$$die = \{d_1, d_2\}.$$
$$score = \{1, 2, 3, 4, 5, 6\}.$$
$$person = \{mike, john\}.$$

$$roll : die \rightarrow score.$$
$$random(roll(D)).$$

$$owner : die \rightarrow person.$$
$$high : boolean.$$

Dice: Rules

The regular part of the program consists of the following rules:

$$
\begin{aligned}
&owner(d_1) = mike.\\
&owner(d_2) = john.
\end{aligned}
$$

$$
\begin{aligned}
high \leftarrow\ & roll(d_1) = Y_1,\\
& roll(d_2) = Y_2,\\
& (Y_1 + Y_2) > 6.\\
\neg high \leftarrow\ & roll(d_1) = Y_1,\\
& roll(d_2) = Y_2,\\
& (Y_1 + Y_2) \leq 6.
\end{aligned}
$$

Dice: Translation $\tau(\Pi_{dice})$

$$
\begin{array}{l}
die(d_1). \\
die(d_2). \\
score(1..6). \\
person(mike). \\
person(john). \\
roll(D,1) \text{ or } \ldots \text{ or } roll(D,6) \leftarrow \ not\ intervene(roll(D)). \\
\neg roll(D, Y_2) \leftarrow roll(D, Y_1),\ Y_1 \neq Y_2. \\
owner(d_1, mike). \\
owner(d_2, john). \\
\neg owner(D, P_2) \leftarrow owner(D, P_1),\ P_1 \neq P_2. \\
high \leftarrow roll(d_1, Y_1),\ roll(d_2, Y_2),\ (Y_1 + Y_2) > 6. \\
\neg high \leftarrow roll(d_1, Y_1),\ roll(d_2, Y_2),\ (Y_1 + Y_2) \leq 6.
\end{array}
$$

Dice: Possible Worlds from Answer Sets

By computing answer sets of $\tau(\Pi_{dice})$ we obtain 36 possible worlds — each world corresponding to a possible selection of values for random attributes $roll(d_1)$ and $roll(d_2)$; i.e.,

$$\begin{array}{c} W_1 = \{roll(d_1) = 1, roll(d_2) = 1, high = false, \dots\}, \\ W_2 = \{roll(d_1) = 1, roll(d_2) = 2, high = false, \dots\}, \\ \vdots \\ W_{35} = \{roll(d_1) = 6, roll(d_2) = 5, high = true, \dots\}, \\ W_{36} = \{roll(d_1) = 6, roll(d_2) = 6, high = true, \dots\}. \end{array}$$

(Atoms that are the same for all possible worlds are not shown.)

A Review of Independence

- In probability theory two events A and B are called independent if the occurrence of one does not affect the probability of another.
- Mathematically, events A and B are independent (with respect to probability function P) if $P(A \wedge B) = P(A) \times P(B)$. This implies: $P(A) = P(A|B)$ and $P(B) = P(B|A)$.
- For example,
  - the event d_1 *shows a* 5 is independent of d_2 *shows a* 5,
  - the event *the sum of the scores on both dice shows a* 5 is dependent on the event d_1 *shows a* 5.

Dice: Using Independence to Compute the Probabilistic Measure

- The selection for d_1 has six possible outcomes which, by the principle of indifference, are equally likely. Similarly for d_2.
- *The mechanisms controlling the way the agent selects the values of $roll(d_1)$ and $roll(d_2)$ during the construction of its beliefs are independent from each other.*
- This independence justifies the definition of the probabilistic measure of a possible world containing $roll(d_1) = i$ and $roll(d_2) = j$ as the product of the agent's degrees of belief in $roll(d_1) = i$ and $roll(d_2) = j$.
- Hence the measure of a possible world containing $roll(d_1) = i$ and $roll(d_2) = j$ for every possible i and j is $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$.

Dice: Bet on *high*

- The probability $P_{\Pi_{dice}}(high)$ is the sum of the measures of the possible worlds which satisfy *high*.
- Since *high* holds in 21 worlds, the probability $P_{\Pi_{dice}}(high)$ of *high* being true is $\frac{7}{12}$.
- Thus, if the reasoner associated with Π_{dice} had to bet on the outcome of the game, betting on *high* would be better.
- (Note that the jungle example did not require the use of the product rule because it contained only one random selection rule.)

Modeling Bias

Suppose now that we learned from a reliable source that while the die owned by John is fair, the die owned by Mike is biased. On average, Mike's die rolls a 6 *in* 1 *out of* 4 *rolls.*

We need a new construct to encode such knowledge.

Causal Probability Statements

$$pr_r(a(\bar{t}) = y|_c \ B) = v$$

where $a(\bar{t})$ is a random attribute, B is a conjunction of literals, r is the name of the random selection rule used to generate the values of $a(\bar{t})$, $v \in [0, 1]$, and y is a possible value of $a(\bar{t})$.

It is read as:
if the value of $a(\bar{t})$ is generated by rule r, and B holds, then the probability of the selection of y for the value of $a(\bar{t})$ is v.

In addition, it indicates the potential *existence of a direct causal relationship between B and the possible value of $a(\bar{t})$.*

Biased Dice: Pr-atom

$$pr(roll(D) = 6 \mid_c owner(D) = mike) = \frac{1}{4}.$$

"The probability of Mike's die rolling a 6 is $\frac{1}{4}$."

- The possible worlds of the two stories about rolling dice are the same, but now P-log can compute probabilistic measures adjusting for this new information.
- Briefly, to compute the measure of a possible world in which $roll(d_1) = 6$, we use $\frac{1}{4} * \frac{1}{6}$ instead of $\frac{1}{6} * \frac{1}{6}$.
- For worlds where $roll(d_1) \neq 6$, our belief in such outcomes is $\frac{(1-\frac{1}{4})}{5} = \frac{3}{20}$. So the measure of each such world is

$$\frac{3}{20} \times \frac{1}{6} = \frac{1}{40}.$$

Observations and Intentions

P-log also allows us to record observations of the results of random experiments:

$$obs(a(\bar{t}) = y)$$

$$obs(a(\bar{t}) \neq y)$$

and the results of *deliberate intervention* in experiments:

$$do(a(\bar{t}) = y)$$

For example:

- $obs(roll(d_1) = 6)$ says that the random experiment consisting of rolling the first die shows 6
- $do(roll(d_1) = 6)$ says that, instead of throwing the die at random, it was deliberately put on the table showing 6

Incorporating the Knowledge: Formal Semantics

Translating the Atoms:

$$obs(a(\bar{t}, y))$$

$$\neg obs(a(\bar{t}, y))$$

$$do(a(\bar{t}, y)).$$

New Rules:

- Eliminate worlds that do not correspond to observations:

$$\leftarrow obs(a(\bar{t}, y)), \neg a(\bar{t}, y)$$

$$\leftarrow \neg obs(a(\bar{t}, y)), a(\bar{t}, y)$$

- Set values for intervened-on attributes:

$$a(\bar{t}, y) \leftarrow do(a(\bar{t}, y))$$

- Break the indifference default to cancel randomness:

$$intervene(a(\bar{t})) \leftarrow do(a(\bar{t}, y))$$

Dynamic Range

- Sometimes our experiments are such that our sample changes.
- Example: What is the probability of drawing two aces in succession?
- If we draw a card from a deck and then draw another card without replacing the first, our sample has changed.
- This means that we need to be able to represent a dynamic range.

Aces in Succession

$$card = \{1 \ldots 52\}.$$

$$ace = \{1, 2, 3, 4\}.$$

$$try = \{1, 2\}.$$

$$draw : try \rightarrow card$$

Can't use $random(draw(T))$ because we are not drawing from the same deck in the first draw as we are in the second. Instead, we use

$$random(draw(T) : \{C : available(C, T)\}).$$

Aces in Succession: Defining the Range

available(C, T) changes based on the try:

$$\begin{array}{lcl} available(C, 1) & \leftarrow & card(C). \\ available(C, T+1) & \leftarrow & available(C, T), \\ & & draw(T) \neq C. \end{array}$$

Aces in Succession: Defining the Attribute of Interest

Defining *two_aces* will allow us to get the probabilistic measure that we're after:

$$
\begin{aligned}
\mathit{two\_aces} \quad \leftarrow \quad & \mathit{draw}(1) = Y1, \\
& \mathit{draw}(2) = Y2, \\
& 1 \leq Y1 \leq 4, \\
& 1 \leq Y2 \leq 4.
\end{aligned}
$$

Note that because of the dynamic range of our selection, the two cards chosen by the two draws can not be the same.

Possible worlds of the program are of the form

$$W_k = \{\mathit{draw}(1) = c_1, \mathit{draw}(2) = c_2, \dots\}$$

where $c_1 \neq c_2$.

Representing Knowledge in P-log

- **Q:** Why P-log? After all, we can compute the probabilities of these simple examples without it.
- **A:** The use of P-log can substantially clarify the modeling process.

The Monty Hall Problem

Monty's show involves a player who is given the opportunity to select one of three closed doors, behind one of which there is a prize. Behind the other two doors are empty rooms. Once the player has made a selection, Monty is obligated to open one of the remaining closed doors which does not contain the prize, showing that the room behind it is empty. He then asks the player if he would like to switch his selection to the other unopened door, or stay with his original choice. Does it matter if he switches?

Representing the General Knowledge of the Domain

$$
\begin{array}{l}
doors = \{1, 2, 3\}. \\
open, selected, prize : doors. \\
\\
\neg can\_open(D) \leftarrow \; selected = D. \\
\neg can\_open(D) \leftarrow \; prize = D. \\
can\_open(D) \leftarrow \; not \; \neg can\_open(D). \\
\\
random(prize). \\
random(selected). \\
random(open : \{X : can\_open(X)\})
\end{array}
$$

Recording What Happened

$$obs(selected = 1).$$
$$obs(open = 2).$$
$$obs(prize \neq 2).$$

Computing the Probabilistic Measures

- Knowing the laws and the observations, the player must now decide whether to switch.
- To decide, compute the probability of the prize being behind door 1 and of the prize being behind door 3.
- To do that, consider the possible worlds of the program and their measures. Then sum up the measures of the worlds in which the prize is behind door 1. Do the same for those with prize behind door 3.

Possible Worlds Given the Observations

$$W_1 = \{selected = 1, prize = 1, open = 2, can\_open(2), can\_open(3)\}.$$

$$W_2 = \{selected = 1, prize = 3, open = 2, can\_open(2)\}.$$

In W_1 the player would lose if she switched; in W_2 she would win.

Note that the possible worlds contain information not only about where the prize is, but which doors Monty can open.

This is the key to correct calculation!

The probabilistic measure of a possible world is the product of likelihoods of the random events it is comprised of. It follows that

$$\hat{\mu}(W_1) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{18}$$
$$\hat{\mu}(W_2) = \frac{1}{3} \times \frac{1}{3} \times 1 = \frac{1}{9}.$$

Normalization gives us:

$$\mu(W_1) = \frac{1/18}{1/18 + 1/9} = \frac{1}{3}$$

$$\mu(W_2) = \frac{1/9}{1/18 + 1/9} = \frac{2}{3}.$$

Finally, since $prize = 1$ is true in only W_1,

$$P_{\Pi_{monty1}}(prize = 1) = \mu(W_1) = \frac{1}{3}.$$

Similarly for $prize = 3$:

$$P_{\Pi_{monty1}}(prize = 3) = \mu(W_2) = \frac{2}{3}.$$

Changing doors doubles the player's chance to win.

Death of a Rat

Consider the following program Π_{rat} representing knowledge about whether a certain rat will eat arsenic today, and whether it will die today.

$$
\begin{aligned}
&arsenic, death : boolean.\\
&random(arsenic).\\
&random(death).\\
&pr(arsenic) = 0.4.\\
&pr(death \mid_c arsenic) = 0.8.\\
&pr(death \mid_c \neg arsenic) = 0.01.
\end{aligned}
$$

- The rat is more likely to die if it eats arsenic.
- Eating arsenic has a causal link with death.

Intuition

- Seeing the rat die raises our suspicion that it has eaten arsenic.
- Killing the rat (with a gun) does not affect our degree of belief that it ate arsenic.
- Does this play out in P-log?

Death of a Rat: Possible Worlds

$$
\begin{array}{lll}
W_1 : & \{arsenic, death\}. & \hat{\mu}(W_1) = 0.4 \times 0.8 = 0.32 \\
W_2 : & \{arsenic, \neg death\}. & \hat{\mu}(W_2) = 0.4 \times 0.2 = 0.08 \\
W_3 : & \{\neg arsenic, death\}. & \hat{\mu}(W_3) = 0.6 \times 0.01 = 0.006 \\
W_4 : & \{\neg arsenic, \neg death\}. & \hat{\mu}(W_4) = 0.6 \times 0.99 = 0.594
\end{array}
$$

Since the unnormalized probabilistic measures add up to 1, they are the same as the normalized measures. Hence,

$$P_{\Pi_{rat}}(arsenic) = \mu(W_1) + \mu(W_2) = 0.32 + 0.08 = 0.4.$$

Death of a Rat: Computing Probabilities with $obs(death)$

- Program $\Pi_{rat} \cup \{obs(death)\}$ has two possible worlds, W_1 and W_3, with unnormalized probabilistic measures as above.
- Normalization yields

$$P_{\Pi_{rat} \cup \{obs(death)\}}(arsenic) = \frac{0.32}{0.32 + 0.006} = 0.982.$$

- The observation of death raised our degree of belief that the rat had eaten arsenic.

Death of a Rat: Computing Probabilities with $do(death)$

- Program $\Pi_{rat} \cup \{do(death)\}$ has the same possible worlds.
- However, $do(death)$ defeats the randomness of death.
- W_1 has unnormalized probabilistic measure 0.4 and W_3 has unnormalized probabilistic measure 0.6. (Same if normalized.)
- Thus,

$$P_{\Pi_{rat} \cup \{do(death)\}}(arsenic) = 0.4.$$

Simpson Paradox

A patient is thinking about trying an experimental drug and decides to consult a doctor. The doctor has tables of the recovery rates that have been observed among males and females, taking and not taking the drug:

Males:		
	fraction_of_popul	recovery_rate
drug	3/8	60%
$\neg$ drug	1/8	70%

Females:		
	fraction_of_popul	recovery_rate
drug	1/8	20%
$\neg$ drug	3/8	30%

Simpson Paradox:

What should the doctor's advice be? If patient is a male, the doctor may attempt to reduce the problem to checking the following inequality involving classical conditional probabilities:

$$P(recover|male, \neg drug) < \; P(recover|male, drug) \tag{1}$$

The inequality fails, and hence the advice is not to take the drug. A similar argument shows that a female patient should not take the drug.

Simpson Paradox:

What if the doctor does not know the patient gender?
Following the same reasoning, the doctor might check whether the following inequality is satisfied:

$$P(recover|\neg drug) < \ P(recover|drug) \tag{2}$$

$P(recovery|drug) = 0.5$ while $P(recovery|\neg drug) = 0.4$.

The drug seems to be beneficial to patients of unknown sex — though similar reasoning has shown that the drug is harmful to the patients of known sex, whether they are male or female!

Simpson Paradox: Pearl's solution

Classical conditional probability (which conditions on the observation of the outcome of a random event) does not faithfully formalize what we really want to know: *what will happen if we do X?*

Condition on doing!

$$P(recover|do(\neg drug)) < \; P(recover|do(drug)) \tag{3}$$

Simpson Paradox:

Pearl develops the mathematics of this new relation and shows that

if we know the person is male then it is better not to take the drug,

the same if we know the person is female,

the same if we do not know the gender.

Simpson Paradox: P-log solution

Program Π:

$$male, recover, drug \quad : \quad boolean$$

$$random(male)$$

$$random(recover)$$

$$random(drug)$$

Simpson Paradox: P-log solution

$pr(male) = 0.5.$

$pr(recover \mid_c male, drug) = 0.6.$
$pr(recover \mid_c male, \neg drug) = 0.7.$
$pr(recover \mid_c \neg male, drug) = 0.2.$
$pr(recover \mid_c \neg male, \neg drug) = 0.3.$

$pr(drug \mid_c male) = 0.75.$
$pr(drug \mid_c \neg male) = .25.$

Simpson Paradox: P-log solution

$$P_{\Pi \cup do(drug)}(recover) = .4$$

$$P_{\Pi \cup do(\neg drug)}(recover) = .5$$

$$P_{\Pi \cup \{obs(male), do(drug)\}}(recover) = 0.6$$

$$P_{\Pi \cup \{obs(male), do(\neg drug)\}}(recover) = 0.7$$

$$P_{\Pi \cup \{obs(\neg male), do(drug)\}}(recover) = 0.2$$

$$P_{\Pi \cup \{obs(\neg male), do(\neg drug)\}}(recover) = 0.3$$

Do not take the drag! Same result – different mathematics.

Other Examples

- The spider bite example, shows that it important to distinguish between deliberately deciding to administer antivenom vs. utilizing the statistic of the administering of antivenom.
- The Bayesian squirrel shows how a dynamic range can be used to model Bayesian learning.
- The wandering robot shows how CR-rules can be used with P-log.

Advantages of P-log

- P-log probabilities are defined with respect to an explicitly stated knowledge base. In many cases this greatly facilitates creation of probabilistic models.
- In addition to logical nonmonotonicity, P-log is "probabilistically nonmonotonic" — addition of new information can add new possible worlds and substantially change the original probabilistic model, allowing for Bayesian learning.
- Possible knowledge base updates include defaults, rules introducing new terms, observations, and deliberate actions in the sense of Pearl.