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 Intentions
  - Dynamic Range
- Representing Knowledge in P-log

Reading

- Read Chapter 11, *Probabilistic Reasoning*, in the KRR book. Don't get caught up in the syntax. Do pay attention to new constructs. Focus on the big concepts: random attributes, causal probabilities, observations, intentions, dynamic range, etc. It is important to understand what is being modeled, that it *can* be modeled, and that the agent can use logical and probabilistic reasoning together.

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 associated with each world.
- **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 possible worlds of Ω to the set of real numbers 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 $\frac{9}{10}$.

Semantics of P-log

- Any P-log program can be translated into a regular ASP program.
- This translation gives us the logical semantics.
- $\tau(\Pi)$ stands for the "translation of P-log program Π into ASP."
- The probabilistic semantics is defined on the answer sets of these programs.

Translation of a P-log Program

For every attribute $a(\bar{t})$ with $range(a) = \{y_1, \ldots, y_n\}$, mapping τ

- represents the sort information by a corresponding set of atoms; e.g.
  $s = \{1, 2\}$ is turned into facts $s(1)$ and $s(2)$;
- replaces every occurrence of an atom

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

  by

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

  and expands the program by rules of the form

  $$\neg a(\bar{t}, Y_2) \leftarrow a(\bar{t}, Y_1), Y_1 \neq Y_2;$$

- replaces every occurrence of $a(\bar{t}, true)$ and $a(\bar{t}, false)$ by $a(\bar{t})$ and $\neg a(\bar{t})$ respectively, and removes double negation $\neg\neg$, which might have been introduced by this operation;

Translation of a P-log Program, cont.

- replaces every rule of the form

$$random(a(\bar{t})) \leftarrow body$$

  by

$$a(\bar{t}, y_1)\ or\ \ldots\ or\ a(\bar{t}, y_n) \leftarrow body,\ not\ intervene(a(\bar{t}))$$

  where *intervene* is a new predicate symbol;
  (Note: P-log actually allows more-general random selection rules which require one more rule.)
- grounds the resulting program by replacing variables with elements of the corresponding sorts.
- P-log has a few more features. We'll see their translation later.

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:

$$
\begin{array}{l}
die = \{d_1, d_2\}. \\
score = \{1,2,3,4,5,6\}. \\
person = \{mike, john\}. \\
\\
roll : die \rightarrow score. \\
random(roll(D)). \\
\\
owner : die \rightarrow person. \\
high : boolean.
\end{array}
$$

Dice: Rules

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

$$owner(d_1) = mike.$$
$$owner(d_2) = john.$$

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

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

```
die(d1).
die(d2).
score(1..6).
person(mike).
person(john).
roll(D,1) | roll(D,2) | roll(D,3) |
   roll(D,4) | roll(D,5) | roll(D,6) :-
     not intervene(roll(D)).
-roll(D,Y2) :-  roll(D,Y1), Y1 != Y2.
owner(d1,mike).
owner(d2,john).
-owner(D,P2) :- owner(D,P1), P1 != P2.
high :- roll(d1, Y1), roll(d2,Y2), (Y1 + Y2) > 6.
-high :- roll(d1,Y1),  roll(d2,Y2),  (Y1 + Y2) <= 6.
```

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.,

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

(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, this intuition is captured by the following definition: events A and B are independent (with respect to probability function P) if $P(A \wedge B) = P(A) \times P(B)$.
- 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

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

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

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

$$draw : try \to card$$

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 she would like to switch her selection to the other unopened door, or stay with her original choice. Does it matter if she switches?

Representing the General Knowledge of the Domain

$$doors = \{1, 2, 3\}.$$
$$open, selected, prize : doors.$$

$$\neg can\_open(D) \leftarrow \ selected = D.$$
$$\neg can\_open(D) \leftarrow \ prize = D.$$
$$can\_open(D) \leftarrow \ not \ \neg can\_open(D).$$

$$random(prize).$$
$$random(selected).$$
$$random(open : \{X : can\_open(X)\})$$

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,

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

The Spider Bite

- Two kinds of poisonous spiders in Stan's location: creeper and spinner.
- Equally common bites locally; spinner bites more common worldwide.
- Experimental antivenom treats both bites, but effectiveness questionable.
- Stan notices bite but not spider.
- Doctor decides based on bite that it's a creeper or spinner and turns to data on antivenom.

Antivenom Data

- Of 416 people bitten by creeper worldwide, 312 received antivenom and 104 did not.
- Of those who received it, 187 survived. Of those who didn't, 73 survived.
- The spinner is more deadly and tends to inhabit areas where the treatment is less available.
- Of 924 people bitten by spinner, 168 received the antivenom, 34 of whom survived.
- Of the 756 spinner victims who did not get antivenom, 227 survived.
- Should Stan take the antivenom?

Formalizing the Story for the Doctor

- Boolean attribute *survive* — a random patient survived.
- Boolean attribute *antivenom* — a random patient was administered antivenom
- Attribute *spider* where $spider = creeper$ or $spider = spinner$ indicates which spider bit the person.
- Thus, we have:

$$
\begin{aligned}
survive, antivenom &: boolean.\\
spider &: \{creeper, spinner\}.
\end{aligned}
$$

$$
\begin{aligned}
&random(spider).\\
&random(survive).\\
&random(antivenom).
\end{aligned}
$$

Formalization, cont.

- Bites from the two spiders are equally common in the area, so the doctor assumes:

$$pr(spider = creeper) = 0.5.$$

- Statistical info from the story:

$$pr(antivenom \mid_c spider = creeper) = 312/416 = 0.75$$
$$pr(antivenom \mid_c spider = spinner) = 168/924 = 0.18$$
$$pr(survive \mid_c spider = creeper, antivenom) = 187/312 = 0.6$$
$$pr(survive \mid_c spider = creeper, \neg antivenom) = 73/104 = 0.7$$
$$pr(survive \mid_c spider = spinner, antivenom) = 34/168 = 0.2$$
$$pr(survive \mid_c spider = spinner, \neg antivenom) = 227/756 = 0.3$$

Conditioning on Intentions vs. Observations

- How should the doctor decide whether to administer the antivenom?
- Compare the results of survival with and without antivenom.
- Is the administration of antivenom by the doctor random?

- To calculate the probability of survival with intentional administration of antivenom, add *do*(*antivenom*) to our program.
- This gives us the following possible worlds and measures:

$$W_1 = \{spider = creeper, antivenom, survive\}$$

$$W_2 = \{spider = creeper, antivenom, \neg survive\}$$

$$W_3 = \{spider = spinner, antivenom, survive\}$$

$$W_4 = \{spider = spinner, antivenom, \neg survive\}$$

$$\mu(W_1) = 0.5 \times 0.6 = 0.3 \quad \ll survive$$

$$\mu(W_2) = 0.5 \times 0.4 = 0.2$$

$$\mu(W_3) = 0.5 \times 0.2 = 0.1 \quad \ll survive$$

$$\mu(W_4) = 0.5 \times 0.8 = 0.4$$

- Probability of survival with intentional antivenom is 0.4.

- Now calculate the probability of survival with intentionally not administrating antivenom by $do(\neg antivenom)$ to our program instead.
- This gives us the following possible worlds and measures:

$$W_5 = \{spider = creeper, \neg antivenom, survive\}$$

$$W_6 = \{spider = creeper, \neg antivenom, \neg survive\}$$

$$W_7 = \{spider = spinner, \neg antivenom, survive\}$$

$$W_8 = \{spider = spinner, \neg antivenom, \neg survive\}$$

$$\mu(W_5) = 0.5 \times 0.7 = 0.35 \quad \ll survive$$

$$\mu(W_6) = 0.5 \times 0.3 = 0.15$$

$$\mu(W_7) = 0.5 \times 0.3 = 0.15 \quad \ll survive$$

$$\mu(W_8) = 0.5 \times 0.7 = 0.35$$

- Probability of survival without antivenom is 0.5.

Conditioning on Observations

- Our calculations show that antivenom should **not** be administered.
- Now suppose the doctor decided to treat himself as an observer, instead of a deliberate actor.
- It is common, and wrong, to used the statistics on the chances of something being administered in the calculation when you are acting deliberately.
- The possible worlds do not change, but the measures of $antivenom/\neg antivenom$ are no longer 1, but taken from the likelihood that antivenom is administered.
- If you use these calculations, you will come to the wrong conclusion!

Bayesian Learning

- Common learning problem: Select from a set of models of a random phenomenon by observing repeated occurrences of that phenomenon.
- Bayesian approach to this problem:
  - Begin with a "prior density" on the set of candidate models; i.e., you assume a likelihood.
  - Update it in light of new observations.

The Bayesian Squirrel

- Example from Ray Hilborn and Marc Mangel, *The Ecological Detective*, Princeton University Press 1997.
- A squirrel has hidden its acorns in one of two patches, but can not remember which.
- The squirrel is 80% certain that the food is hidden in Patch 1.
- It knows there is a 20% chance of finding food per day when it is looking in the right patch (and, of course, a 0% chance if it's looking in the wrong patch).

P-log Bayesian Squirrel

- Sorts:

$$patch = \{p1, p2\}.$$
$$day = \{1 \dots n\}.$$

(where n is some constant, say, 5)

- Attributes:

$$hidden\_in : patch.$$
$$found : day \rightarrow boolean.$$
$$look : day \rightarrow patch.$$

Which Attributes Are Random?

- Attribute *hidden_in* is always random:

$$random(hidden\_in).$$

- Attribute *found* is random only if the squirrel is looking for food **in the right patch:**

$$\begin{aligned} random(found(D)) \leftarrow \quad & hidden\_in = P, \\ & look(D) = P. \end{aligned}$$

  Otherwise we have:

$$\begin{aligned} \neg found(D) \leftarrow \quad & hidden\_in = P_1, \\ & look(D) = P_2, \\ & P_1 \neq P_2. \end{aligned}$$

- Attribute $look(D)$ is not random because it is decided by the squirrel's **deliberation.**

Probabilistic Information

$$pr(hidden\_in = p1) = 0.8.$$
$$pr(found(D)) = 0.2.$$

Compute Possible Outcomes of the Next Search for Food

- Add $look(1) = p_1$ to the program.
- Possible worlds and their measures:

$$W_1^1 = \{look(1) = p_1, hidden\_in = p_1, found(1), \ldots\}$$
$$W_2^1 = \{look(1) = p_1, hidden\_in = p_1, \neg found(1), \ldots\}$$
$$W_3^1 = \{look(1) = p_1, hidden\_in = p_2, \neg found(1), \ldots\}$$

$$\mu(W_1^1) = 0.16$$
$$\mu(W_2^1) = 0.64$$
$$\mu(W_3^1) = 0.2$$

- $$P_{\Pi_{sq1}}(hidden\_in = p_1) = 0.16 + 0.64 = 0.8$$
$$P_{\Pi_{sq1}}(found(1)) = 0.16.$$

It's a New Day

- Suppose the squirrel didn't find the nut on day 1.
- This time, it should be a bit less sure that it is in Patch 1.
- We add its observations and intention to the first program:

$$\begin{array}{l} obs(\neg found(1)). \\ look(2) = p_1. \end{array}$$

Possible Worlds for Day 2, Looking in Patch 1

$W_1^2 = \{look(1) = p_1, \neg found(1), hidden\_in = p_1, look(2) = p_1, found(2) \dots\}$

$W_2^2 = \{look(1) = p_1, \neg found(1), hidden\_in = p_1, look(2) = p_1, \neg found(2) \dots\}$

$W_3^2 = \{look(1) = p_1, \neg found(1), hidden\_in = p_2, look(2) = p_1, \neg found(2) \dots\}$

$$\mu(W_1^2) = 0.128/0.84 = 0.152$$
$$\mu(W_2^2) = 0.512/0.84 = 0.61$$
$$\mu(W_3^2) = 0.2/0.84 = 0.238$$

Consequently,

$$P_{\Pi_{sq2}}(hidden\_in = p_1) = 0.762$$

and

$$P_{\Pi_{sq2}}(found(2)) = 0.152$$

Probabilistic Nonmonotonicity

- Notice that the squirrel is now less certain that the nut is in patch 1.
- The only changes to the program were the additions of actions and observations.
- P-log enables this kind of learning because it can represent
  - observations,
  - actions, and
  - conditional randomness.

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.

Summary

You have been introduced to a large variety of approaches to AI:

- Neural Nets and their use in machine learning of pattern recognition.
- Genetic Algorithms and their application to search.
- Logic Programming and its application to modeling nonmonotonic reasoning.
- Action Languages and their application to:
  - reasoning about actions and change
  - planning
  - diagnostics
- Hidden Markov Models and the Viterbi Algorithm and their use in Natural Language Processing.
- P-log, which combines probabilistic and logical reasoning, and its application to modeling Bayesian reasoning and learning.