SYMBOLIC SYSTEMS 100: Introduction to Cognitive Science

Dan Jurafsky and Daniel Richardson
Stanford University
Spring 2005

May 17, 2005: Human Decision Making

IP Notice: Some slides from David Beaver's lectures last year, some from Stephen Stich at Rutgers at http://www.rci.rutgers.edu/~stich, some from Norman Fenton, and some from Gert Gigerenzer's own slides.

Aristotle

Slide from Stephen Stich

Two Views on Human Rationality

- **Aristotle (384 - 322 B.C.) claimed that humans are rational animals.**
  - **For 2300 years philosophers have (more or less) agreed.**
  - **Aristotle was aware that people's judgments, decisions & behavior are not *always* rational.**
  - **A modern rendition of Aristotle's claim is that all normal humans have the *competence* to be rational.**
    - **The correct principles of reasoning & decision making are in our minds, even if we don't always use them.**

Utility

John Stuart Mill
1806 - 1873

- Economics and many other fields depend on a model of human decision making.
- Classical view: maximize profit, or (a la Mill) maximize utility.
- Bernoulli argued that we become progressively more indifferent to larger gains: a graph of actual profit against perceived value would be curved.

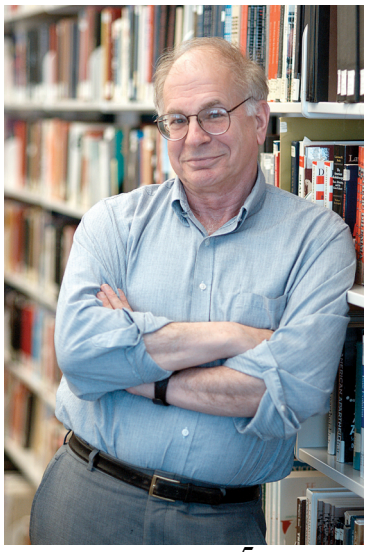
Daniel Bernoulli
1700 - 1782

Slide from David Beaver

Heuristics and Biases

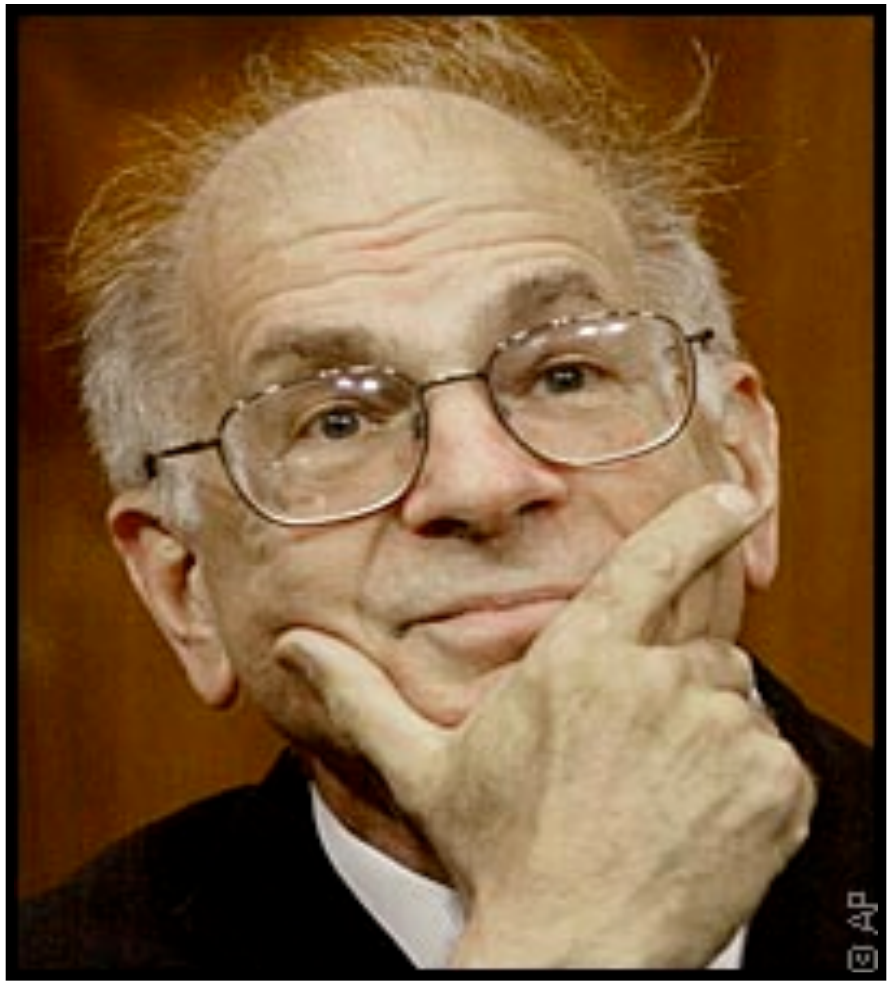
- Tversky and Kahneman dramatically extended Bernoulli's account, and provided extensive psychological evidence for their model.
- What they showed is that humans depart in multiple ways from a classical picture of rational behavior.

Amos Tversky
1937 - 1996



Daniel Kahneman

Slide from David Beaver



Daniel Kahneman

Winner of the 2002 Nobel Prize in Economics

Slide from Stephen Stich

The Kahneman and Tversky Program

- Many have interpreted the Kahneman and Tversky program as showing that Aristotle was wrong.
  - This "heuristics & biases" program claims that
    - most people *do not* have the correct principles for reasoning & decision making
    - we get by with much simpler principles ("heuristics & biases") which sometimes get the right answer, and sometimes do not

Slide from Stephen Stich

Experimental Studies of Human Reasoning

- **The Conjunction Fallacy**
- **Base Rate Neglect**
- **Overconfidence**
- **Framing**

Slide from Stephen Stich

An intuition from Gigerenzer

Which US city has more inhabitants,
San Diego or San Antonio?

Americans:

62%

correct

Germans:

100%

correct

Slide from Gigerenzer

Recognition Heuristic

If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value.

Ecological Rationality

The heuristic is successful when ignorance is systematic rather than random, that is, when lack of recognition correlates with the criterion.

Slide from Gigerenzer

Experimental Studies of Human Reasoning

- **The Conjunction Fallacy**
- Base Rate Neglect
- Overconfidence
- Framing

Slide from Stephen Stich

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

(a) Linda is a teacher in elementary school.
(b) Linda works in a bookstore and takes Yoga classes.
(c) Linda is active in the feminist movement.
(d) Linda is a psychiatric social worker.
(e) Linda is a member of the League of Women Voters.
(f) Linda is a bank teller.
(g) Linda is an insurance sales person.
(h) Linda is a bank teller and is active in the feminist movement.

Slide from Stephen Stich

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

(a) Linda is a teacher in elementary school.
(b) Linda works in a bookstore and takes Yoga classes.
(c) Linda is active in the feminist movement.
(d) Linda is a psychiatric social worker.
(e) Linda is a member of the League of Women Voters.
(f) Linda is a bank teller.
(g) Linda is an insurance sales person.
(h) Linda is a bank teller and is active in the feminist movement.

Naïve subjects: (h) > (f) 89%

Slide from Stephen Stich

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

(a) Linda is a teacher in elementary school.
(b) Linda works in a bookstore and takes Yoga classes.
(c) Linda is active in the feminist movement.
(d) Linda is a psychiatric social worker.
(e) Linda is a member of the League of Women Voters.
(f) Linda is a bank teller.
(g) Linda is an insurance sales person.
(h) Linda is a bank teller and is active in the feminist movement.

Naïve subjects: (h) > (f) 89%; Sophisticated subjects: 85%

Slide from Stephen Stich

The Conjunction Fallacy: another version

- John is 19, wears glasses, is a little shy unless you talk to him about Star Trek or Lord of the Rings, and stays up late most nights playing video games.
- Which is more likely:
  1. John is a CS major who loves contemporary sculpture and golf, or
  2. John loves contemporary sculpture and golf?
- Once again, on problems like this, people tend to pick (1) above (2), in contradiction to probability theory.

Experimental Studies of Human Reasoning

- The Conjunction Fallacy
- **Base Rate Neglect**
- Overconfidence
- Framing

Slide from Stephen Stich

A panel of psychologists have interviewed and administered personality tests to 30 engineers and 70 lawyers [70 engineers and 30 lawyers], all successful in their respective fields. On the basis of this information, thumbnail descriptions of the 30 engineers and 70 lawyers [70 engineers and 30 lawyers], have been written. You will find on your forms five descriptions, chosen at random from the 100 available descriptions. For each description, please indicate your probability that the person described is an engineer, on a scale from 0 to 100.

Slide from Stephen Stich

Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles.

Slide from Stephen Stich

Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles.

Dick is a 30-year-old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues.

Slide from Stephen Stich

Base rate and representativeness: can people do the math?

- With no personality sketch, simply asked for the probability that an unknown individual was an engineer
  - subjects correctly gave the responses 0.7 and 0.3
- When presented with a totally uninformative description
  - the subjects gave the probability to be 0.5
- Kahneman & Tversky concluded that when no specific evidence is given, prior probabilities are used properly; when worthless evidence is given, prior probabilities are ignored.

(Material adapted from N. Fenton)

Slide from David Beaver

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?

____%

Slide from Stephen Stich

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?

____%

Harvard Medical School: "2%" = 18%; "95%" = 45%

Slide from Stephen Stich

Why is the correct answer 2%?

- Think of a population of 10,000 people.
- We would expect just 10 people in this population to have the disease (1/1000 x 10,000 = 10)
- If you test everybody in the population then the false positive rate means that, in addition to the 10 people who do have the disease, another 500 (5% of 10,000) will be wrongly diagnosed as having it.
- In other words only about 2% of the people diagnosed positive (10/510) actually have the disease.
- When people give a high answer like 95% they are ignoring the very low probability (i.e. rarity) of having the disease. In comparison the probability of a false positive test is relatively high.

Experimental Studies of Human Reasoning

- The Selection Task
- The Conjunction Fallacy
- Base Rate Neglect
- **Overconfidence**
- Framing

Slide from Stephen Stich

In each of the following pairs, which city has more inhabitants?

- (a) Las Vegas (b) Miami

  How confident are you that your answer is correct?

  50% 60% 70% 80% 90% 100%

- (a) Sydney (b) Melbourne

  How confident are you that your answer is correct?

  50% 60% 70% 80% 90% 100%

- (a) Hyderabad (b) Islamabad

  How confident are you that your answer is correct?

  50% 60% 70% 80% 90% 100%

Slide from Stephen Stich

Experimental Studies of Human Reasoning

- The Selection Task
- The Conjunction Fallacy
- Base Rate Neglect
- Overconfidence
- **Framing**

Slide from Stephen Stich

Imagine the U.S. is preparing for the outbreak of an unusual Asian disease which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the program are as follows:

Slide from Stephen Stich

- **If program A is adopted , 200 people will be saved.**
- **If program B is adopted, there is a 1/3 probability that 600 people will be saved and a 2/3 probability that no people will be saved.**

Slide from Stephen Stich

- If program A is adopted , 200 people will be saved.
- If program B is adopted, there is a 1/3 probability that 600 people will be saved and a 2/3 probability that no people will be saved.

Slide from Stephen Stich

- **If program C is adopted, 400 people will die.**
- **If program D is adopted, there is a 1/3 probability that nobody will die and a 2/3 probability that 600 people will die.**

Slide from Stephen Stich

- If program C is adopted, 400 people will die.

- If program D is adopted, there is a 1/3 probability that nobody will die and a 2/3 probability that 600 people will die.

Slide from Stephen Stich

- If program A is adopted , 200 people will be saved.
- If program B is adopted, there is a 1/3 probability that 600 people will be saved and a 2/3 probability that no people will be saved.
- If program C is adopted, 400 people will die.
- If program D is adopted, there is a 1/3 probability that nobody will die and a 2/3 probability that 600 people will die.

Slide from Stephen Stich

Program A versus B

- Tversky and Kahneman (1981) found that the majority choice (72%) for an analogue of problem 1 using an anonymous "Asian disease" (not SARS) was answer A.
- The prospect of saving 200 lives with certainty was more promising than the probability of a one-in-three chance of saving 600 lives.

But A + B have the same utility

- The standard way of calculating expected utility:
  - $\Sigma_i P(o_i) \times U(o_i)$
  - where each o_i is a possible outcome, $P(o_i)$ is its probability and $U(o_i)$ is its utility or value.
- On this basis, option B (*a risky prospect*) would be of equal expected value to the first prospect A.
- So in this case subjects are not simply using expected utility: they are risk averse.

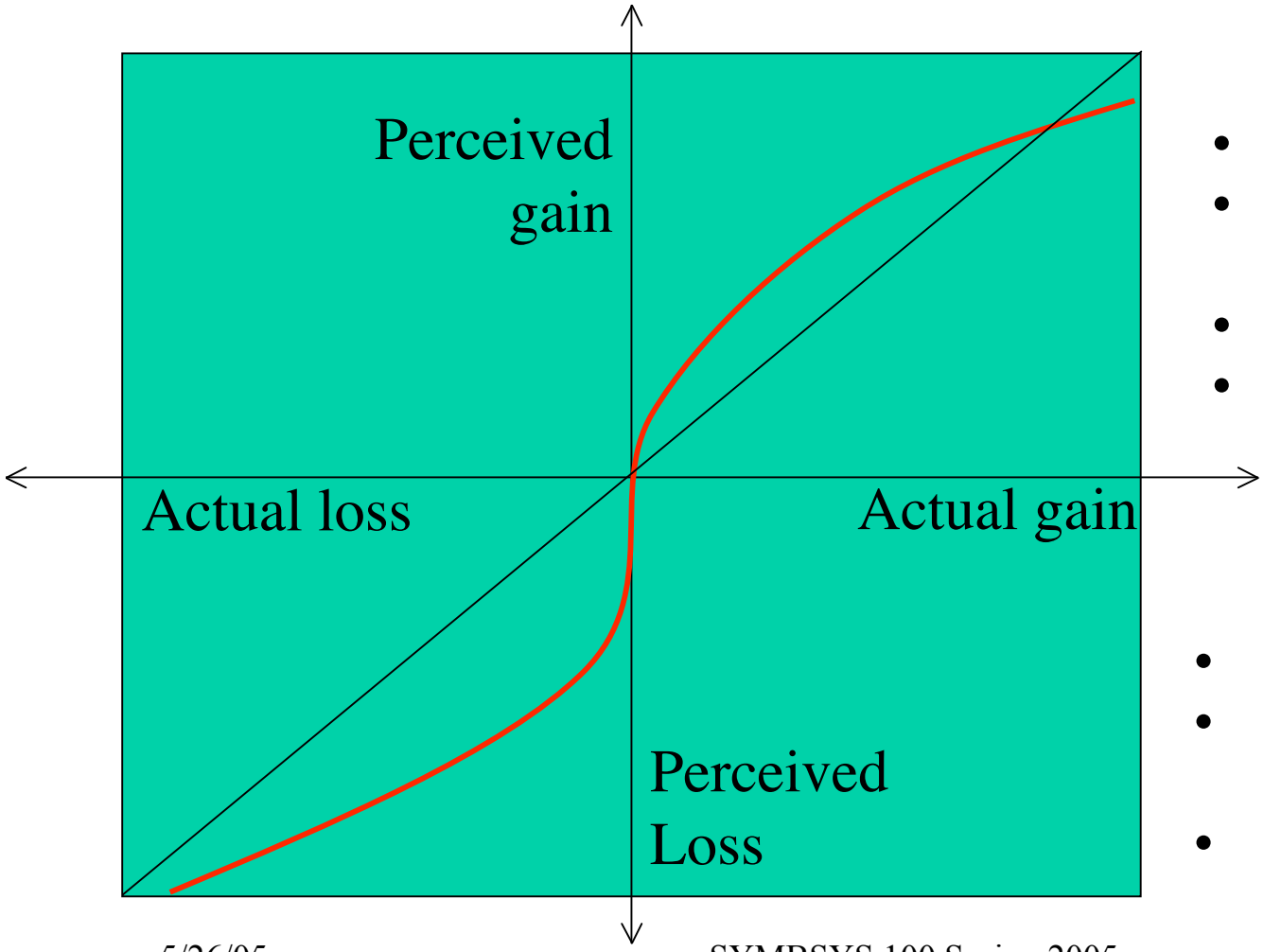
Program C versus D

- The majority of respondents in the second problem (78%) chose the riskier option.
- The certain death of 400 people is apparently less acceptable than the two-in-three chance that 600 people will die.
- Once again, there is no difference between the options (C and D) in standard expected utility.
- We say in this case that people are risk seeking.

Framing

- Tversky and Kahneman's intention was that Problem 1 and Problem 2 are underlyingly identical, but presented in different ways.
- To the extent that they are right, we say that they are different framings of the same problem: the different responses for the two problems are what we call framing effects.
- Yet another way in which people's decisions differ from what would be predicted on the standard view of expected utility.

The Incredible S-curve of Value



- Small gain good,
- Big gain not much better,
- Small loss *terrible*,
- Big loss not much worse.

Thus we are generally:

- Loss averse,
- Risk averse for gains, but
- Risk seeking for losses.

Estimation of Risk

Estimates of Probabilities of Death From Various Causes:

Cause	*Stanford graduate estimates*	*Statistical estimates*
Heart Disease	**0.22**	**0.34**
Cancer	**0.18**	**0.23**
Other Natural Causes	**0.33**	**0.35**
All Natural Causes	0.73	0.92
Accident	**0.32**	**0.05**
Homicide	**0.10**	**0.01**
Other Unnatural	**0.11**	**0.02**
All Unnatural	0.53	0.08

(Based on a study by Amos Tversky, Thayer Watkin's report)

Big and small probabilities

- Small probability events are overrated.
- More generally, behavior departs from classical rationality substantially as regards very small and very large probabilities.
- Even when told the frequency of events, subjects tend to view the difference between 0 and 0.001 (or that between 0.999 and 1) as more significant than that between .5 and .5001.

The "Bleak Implications" Hypothesis

- These results have "bleak implications" for the rationality of ordinary people (Nisbett et al.)
- "It appears that people lack the correct programs for many important judgmental tasks…. We have not had the opportunity to evolve an intellect capable of dealing conceptually with uncertainty." (Slovic, Fischhoff and Lichtenstein , 1976)

Slide from Stephen Stich

The "Bleak Implications" Hypothesis

- "I am particularly fond of [the Linda] example, because I know that the [conjunction] is least probable, yet a little homunculus is my head continues to jump up and down, shouting at me – "but she can't just be a bank teller; read the description." … Why do we consistently make this simple logical error? Tversky and Kahneman argue, correctly I think, that our minds are not built (for whatever reason) to work by the rules of probability." (Stephen J. Gould , 1992, p. 469)

Slide from Stephen Stich

The Challenge from Evolutionary Psychology

- The "frequentist" hypothesis
  - Useful information about probabilities was available to our ancestors in the form of *frequencies*
    - E.g. 3 of the last 12 hunts near the river were successful
  - Information about the probabilities of single events was not available
  - So perhaps we evolved a mental capacity that is good at dealing with probabilistic information, but only when that information is presented in a frequency format.

Slide from Stephen Stich

Evolutionary Psychology Applied to Reasoning: Some Experimental Results

- **Making Base Rate Neglect "Disappear"**

Slide from Stephen Stich

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?

____%

Harvard Medical School: "2%" = 18%; "95%" = 45%

Slide from Stephen Stich

1 out of every 1000 Americans has disease X. A test has been developed to detect when a person has disease X. Every time the test is given to a person who has the disease, the test comes out positive. But sometimes the test also comes out positive when it is given to a person who is completely healthy. Specifically, out of every 1000 people who are perfectly healthy, 50 of them test positive for the disease.

Imagine that we have assembled a random sample of 1000 Americans. They were selected by lottery. Those who conducted the lottery had no information about the health status of any of these people.

Given the information above: on average, how many people who test positive for the disease will *actually* have the disease? _____ out of _____.

Slide from Stephen Stich

1 out of every 1000 Americans has disease X. A test has been developed to detect when a person has disease X. Every time the test is given to a person who has the disease, the test comes out positive. But sometimes the test also comes out positive when it is given to a person who is completely healthy. Specifically, out of every 1000 people who are perfectly healthy, 50 of them test positive for the disease.

Imagine that we have assembled a random sample of 1000 Americans. They were selected by lottery. Those who conducted the lottery had no information about the health status of any of these people.

Given the information above: on average, how many people who test positive for the disease will *actually* have the disease? _____ out of _____.

"1 out of 51" = 76%; n = 50 Slide from Stephen Stich

Evolutionary Psychology Applied to Reasoning: Some Experimental Results

- Making Base Rate Neglect "Disappear"
- **Making the Conjunction Fallacy "Disappear"**

Slide from Stephen Stich

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

(a) Linda is a teacher in elementary school.
(b) Linda works in a bookstore and takes Yoga classes.
(c) Linda is active in the feminist movement.
(d) Linda is a psychiatric social worker.
(e) Linda is a member of the League of Women Voters.
(f) Linda is a bank teller.
(g) Linda is an insurance sales person.
(h) Linda is a bank teller and is active in the feminist movement.

Slide from Stephen Stich

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

There are 100 people who fit the description above. How many of them are:

…

(f) bank tellers?

(h) bank tellers and active in the feminist movement?

...

Slide from Stephen Stich

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

There are 100 people who fit the description above. How many of them are:

…

(f) bank tellers?

(h) bank tellers and active in the feminist movement?

...

(h) > **(f)** = 13%

Slide from Stephen Stich

Another response

- Gigerenzer
- Fast and Frugal Heuristics might be a better solution to optimal rational reasonsing in many situations

Two models of reasonsing

- **Optimal probabilistic decision-making**
  - **People are probabilistic reasoners**
  - **People make decisions by complex computation of optimum behavior**
- **Fast and Frugal:**
  - **People have to make inferences with limited time and limited knowledge**
  - **Computing the probabilistically optimal thing to do is hard and slow. And requires lots of complex knowledge**
  - **So maybe people most of the time use heuristics that are:**
    - **Fast (limited time)**
    - **Frugal (limited knowledge)**

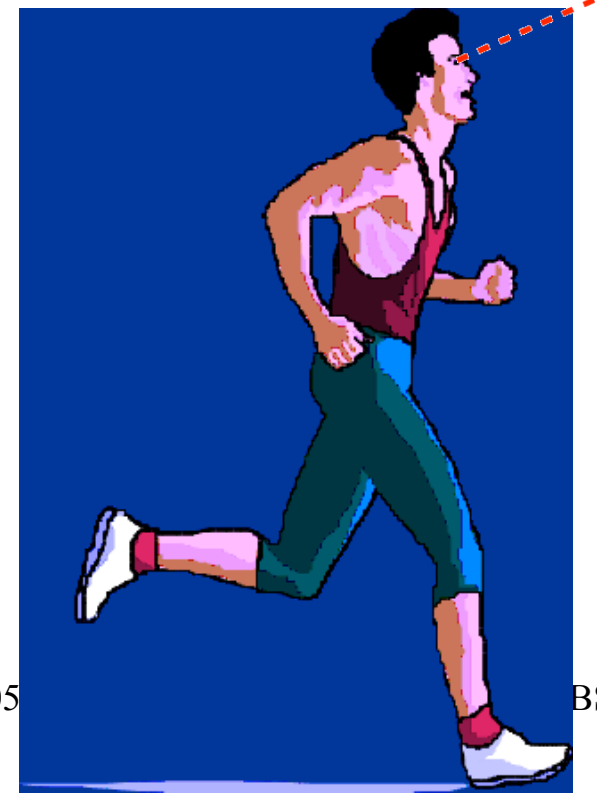
Outfields, baseballs, dogs, and frisbees

- **Shaffer, D.M., S.M. Krauchunas, M. Eddy, and M.K. McBeath. 2004. How dogs navigate to catch Frisbees. *Psychological Science* 15(July):437-441**
- **An intuition for "Fast and Frugal Heuristics"**

Gaze heuristic

Slide from Gigerenzer

Gaze heuristic



Slide from Gigerenzer

Gaze heuristic

Slide from Gigerenzer

Gaze heuristic

Slide from Gigerenzer

Gaze heuristic:
One-reason Decision Making

- ***Predation and pursuit:***

  bats, birds, dragonflies, hoverflies, teleost fish, houseflies

- ***Avoiding collisions:***

  sailors, aircraft pilots

- ***Sports:***

  baseball outfielders, cricket, dogs catching Frisbees

NOTE: Gaze heuristic ignores all causal relevant variables

Slide from Gigerenzer

Gigerenzer Fast and Frugal Heuristics

- **Idea:**
  - **Test to see if fast and frugal heuristics perform as well as more complex optimal models**
  - **Task: predicting some decision from some features of the environment**
  - **"Which of these two cities has more homeless:**
    - **Los Angeles**
    - **New Orleans"**

Cues for predicting homelessness

	LA	Chicago	NY	Norleans
Homeless/million	10,526	6,618	5,024	2,671
Rent control	1	0	1	0
Vacancy rate low	1	1	1	0
High temperature	1	0	1	1
Unemployment	1	1	1	1
Poverty	1	1	1	1
Public Housing bad	1	1	0	0

The Minimalist Heuristic

- **Minimalist**
  - **Step 1: Random Search: Randomly select a cue and look up the cue values of the two objects**
  - **Step 2: Stopping rule: If one obejct has a vue value of one and the other does not, stop search. Otherwise, go back to step 1. If no further cue is found, guess.**
  - **Step 3: Decision rule: Predict that the object with the cue value of one has higher value**

Take the Best

- Instead of random order, try best cue first; if that cue doesn't discriminate, try next one, and so on.
- Step 1: Ordered Search: Select the cue with the highest validity and look up the cue values of the two objects
  - Validity v_i of cue *i* is the number of correct inference divided by the total number of inferences based on cue I.
    - Vi = Ri/(Ri+Wi)
    - LA has cue value of 1 for rent control
    - Chicago has cue value of 0
    - Since LA does hav ea higher homelessness rate, this counts as a right inference.
    - Sum over all pairings of cities

Properties of Heuristics

- **Stop-by-step procedures**
- **Simple stopping rules**
- **One-reason decision making**

How good is fast and frugal?

- **Predict homelessness from cues**
- **Compare with sophisticated models:**
  - **Linear regression**
  - **Bayesian networks**

Trade-off between accuracy and cues looked up in predicting homelessness

Strategy	Avg # of cues looked up	% correct test set = training set	% correct test set != training set
Minimalist	2.1	61	56
Take the Best	2.4	69	63
Multiple regression	6	70	61

When and Why are Frugal Heuristics good?

- **Scarce information**
  - **If there isn't much knowledge, knowledge-based methods fail; heuristics are robust**
- **Highly valid cues are really robust**
  - **Take the Best works well since it relies on a small number of cues for which it has lots of evidence are good.**

Conclusions

- Humans often reply on non-optimal heuristics rather than optimal methods
  - Kahneman and Tversky research program
    - Where do people go wrong, what heuristics:
      - Representativeness
      - Framing
      - Insensitive to Base Rate
  - Gigerenzer
    - Showing that heuristics work pretty well
      - Take the Best
      - Minimalist