Class Mechanics

• Instructor: Prof. Haym Hirsh, cs.cornell.edu/~hirsh
• TA: Eric Wang, ericwang0701@gmail.com, ericwang.info
• Course website: cs.cornell.edu/courses/cs5306
• Discussion: piazza.com/cornell/fall2017/cs5306
• Coursework submission: gradescope.com (entry code 9ZRR3V)
• Exams:
  • Prelim (tentative): 10/17/2017, in class
  • Final: 12/12/2017, 9:00AM - 11:30AM
Coursework

• Prelim: 15-20%
• Final: 15-20%
• Assignments (~5): 10-20%
• Projects (2): 45-55%
• Extra credit: Used if you are near the boundary between grades

• READINGS!
Judgment under Uncertainty: Heuristics and Biases

Biases in judgments reveal some heuristics of thinking under uncertainty.

Amos Tversky and Daniel Kahneman

Many decisions are based on beliefs concerning the likelihood of uncertain events such as the outcome of an election, the Gulf of a defendant, or the future value of the dollar. These beliefs are usually expressed in statements such as “I think that . . . .” or “odds are . . . .” or “it is unlikely that . . . .” and so forth. Occasionally, beliefs concerning uncertain events are expressed in numerical form as odds or subjective probabilities. What determines such beliefs? How do people assess the probability of an uncertain event or the value of an uncertain quantity? This article shows that people rely on a limited number of heuristic principles which reduce the complex task of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to serious and systematic errors.

The subjective assessment of probability exemplifies the subjective assessment of physical quantities such as distance or size. These judgments are all based on data of limited validity, which are presented according to heuristic rules. For example, the apparent distance of an object is determined partly by its clarity. The more shrouded the object is seen, the closer it appears to be. This rule has some validity, because in any given scene the more distant objects are seen less clearly than nearer objects. However, the reliance on this rule leads to systematic errors in the estimation of distance. Specifically, distances are often overestimated when visibility is poor because the contours of objects are blurred. On the other hand, distances are often underestimated when visibility is good because the contours of objects are sharp. Hence, it appears that people rely on a variety of principles that allow them to form judgments about probabilities but lead to serious errors in the judgmental process.

Representativeness

Many of the probabilistic questions which people are concerned belong to one of the following types: What is the probability that object A belongs to class B? What is the probability that event A originates from process B? What is the probability that person X generates event A? In answering such questions, people typically rely on the representativeness heuristic, in which probabilities are evaluated by the degree to which a specific similarity matches the general probability of the event. For example, if A is highly representative of B, the probability that A originates from B is judged to be high. On the other hand, if A is not similar to B, the probability that A originates from B is judged to be low. For an illustration of judgment by representativeness, consider the following story: A psychologist has been described by his former neighbor as follows: "Steve is a very shy and withdrawn, unusually helpful, but with little interest in people, or in the world of reality. A neat and tidy soul, he has a need for order and routine, and a passion for details." How do people assess the probability that Steve is engaged in a particular occupation from a list of possibilities (for example, farmer, salesman, airline pilot, engineer, physician)? How do people order these occupations from most to least likely? In the representation heuristic, the probability that Steve is a librarian, for example, is reversed by the degree to which he is representative of, or similar to, the stereotype of a librarian. Indeed, research with problems of this type has shown that people order the occupations by probability and by certainty is exactly the same way (1). This approach to the judgment of probability leads to serious errors, because similarity, or representativeness, is a reflection of several factors that should affect judgments of probability. Inaccuracy in prior probability of outcomes. One of the factors that have an effect on representativeness, but should have a major effect on probability, is the prior probability, or base-rate frequency, of the outcomes. In the case of Steve, for example, the fact that there are many more farmers than librarians in the population should enter into any reasonable estimate of the probability that Steve is a librarian rather than a farmer. Considerations of base-rate frequency, however, do not affect the similarity of Steve to the stereotypes of librarians and farmers. If people evaluate probability by representativeness, therefore, prior probabilities will be neglected. This hypothesis was tested in an experiment where prior probabilities were manipulated (1). Subjects were shown brief personal descriptions of several individuals, allegedly sampled at random from a group of 100 professional engineers and lawyers. The subjects were asked to assess, for each description, the probability that it belonged to an engineer rather than to a lawyer. In one experimental condition, subjects were told that the group from which the descriptions had been drawn consisted of 10 engineers and 30 lawyers. In another condition, subjects were told that the group consisted of 50 engineers and 20 lawyers. The odds that any particular descriptive characteristic belongs to an engineer rather than to a lawyer should be higher in the first condition, when there is a majority of engineers, than in the second condition, where there is a majority of lawyers. Specifically, it can be shown by applying Bayes' rule that the ratio of these odds should be 1.33, or 5:3, for such descriptive characteristics. In a simple violation of Bayes' rule, the subjects in the two conditions produced results
Kahneman and Twersky

Humans are subject to “architectural” errors
What Do We Know about People?
What Do We Know about Amazon Mechanical Turk?
What Do We Know about People Using Computers?
Computers Are Social Actors (CASA)

People interact with computers as if they are social actors

Theories about how people interact can apply to how people interact with computers
What Do We Know about People Using Computers?
What Do We Know about Amazon Mechanical Turk?
What Do We Know about What Motivates Turkers?
What Do We Know about Motivation?
Abraham Maslow
Maslow’s Hierarchy of Needs

- **Physiological**
  - breathing, food, water, sex, sleep, homeostasis, excretion

- **Safety**
  - security of: body, employment, resources, morality, the family, health, property

- **Love/belonging**
  - friendship, family, sexual intimacy

- **Esteem**
  - self-esteem, confidence, achievement, respect of others, respect by others

- **Self-actualization**
  - morality, creativity, spontaneity, problem solving, lack of prejudice, acceptance of facts
What Do We Know about What Motivates Turkers?
What Do We Know about What Motivates Open Source Programmers?
“Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects”

K.R. Lakhani and R.G. Wolf

*Perspectives on Free and Open Source Software*, MIT Press, 2005
<table>
<thead>
<tr>
<th>Motivation</th>
<th>% of respondents indicating up to 3 statements that best reflect their reasons to contribute (%)</th>
<th>% volunteer contributors</th>
<th>% paid contributor</th>
<th>Significant difference (t statistic/p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Enjoyment based Intrinsic Motivation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code for project is intellectually stimulating to write</td>
<td>44.9</td>
<td>46.1</td>
<td>43.1</td>
<td>n.s.</td>
</tr>
<tr>
<td>Like working with this development team</td>
<td>20.3</td>
<td>21.5</td>
<td>18.5</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Economic/Extrinsic based Motivations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improve programming skills</td>
<td>41.3</td>
<td>45.8</td>
<td>33.2</td>
<td>3.56 (p=0.0004)</td>
</tr>
<tr>
<td>Code needed for user need (work and/or non-work)*</td>
<td>58.7</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>- Work need only</td>
<td>33.8</td>
<td>19.3</td>
<td>55.7</td>
<td>10.53 (p=0.0000)</td>
</tr>
<tr>
<td>- Non-work need</td>
<td>29.7</td>
<td>37.0</td>
<td>18.9</td>
<td>5.16 (p=0.0000)</td>
</tr>
<tr>
<td>Enhance professional status</td>
<td>17.5</td>
<td>13.9</td>
<td>22.8</td>
<td>3.01 (p=0.0000)</td>
</tr>
<tr>
<td><strong>Obligation/Community based Intrinsic Motivations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Believe that source code should be open</td>
<td>33.1</td>
<td>34.8</td>
<td>30.6</td>
<td>n.s.</td>
</tr>
<tr>
<td>Feel personal obligation to contribute because use F/OSS</td>
<td>28.6</td>
<td>29.6</td>
<td>26.9</td>
<td>n.s.</td>
</tr>
<tr>
<td>Dislike proprietary software and want to defeat them</td>
<td>11.3</td>
<td>11.5</td>
<td>11.1</td>
<td>n.s.</td>
</tr>
<tr>
<td>Enhance reputation in F/OSS community</td>
<td>11.0</td>
<td>12.0</td>
<td>9.5</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
What Do We Know about What Motivates Participation in Online Communities?
Building Successful Online Communities
Evidence-Based Social Design

Robert E. Kraut
Paul Resnick

with Sara Kiesler, Mora Berke, Yan Chen, Nel Kettner, Joseph Konstan, Yuning Pan, and John Reed
"More than fun and money. Worker Motivation in Crowdsourcing – A Study on Mechanical Turk“

Nicolas Kaufmann, Thimo Schulze, and Daniel Veit

<table>
<thead>
<tr>
<th>Paper</th>
<th>Focus</th>
<th>Intrinsic Motivation</th>
<th>Extrinsic Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Leimeister et al., 2009)</td>
<td>Idea Competitions</td>
<td>Enjoyment Based Motivation</td>
<td>Immediate Payoffs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Community Based Motivation</td>
<td>„Direct compensation“</td>
</tr>
<tr>
<td>(Brabham, 2008)</td>
<td>Content Market</td>
<td>“Creative outlet”; “Fun”; “Produce [content] that I like”; “Passes the time when I am bored”</td>
<td>„Learning“; “Self-Marketing“</td>
</tr>
<tr>
<td>(Brabham, 2010)</td>
<td>Design Competition</td>
<td>“Love of community”; “Addiction to the community”</td>
<td>„Improve skills“; “Earn a reputation”</td>
</tr>
<tr>
<td>(Ipeirotis, 2010)</td>
<td>Mechanical Turk</td>
<td>“Fruitful way to spend free time”; “To kill time”; “Tasks are fun”</td>
<td>„Improve creative skills“; “Get employed as a freelancer”</td>
</tr>
<tr>
<td>(Organisciak, 2008)</td>
<td>Crowdsourcing</td>
<td>Fun; Boredom; achievement (by the action); Interest (curiosity)</td>
<td>Money</td>
</tr>
<tr>
<td>Motivation Type</td>
<td>Component</td>
<td>Value</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------------------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Enjoyment Based Motivation</td>
<td>Skill Variety</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Task Identity</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Task Autonomy</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Direct Job Feedback</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pastime</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>Community Based Motivation</td>
<td>Community Identity</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social Contact</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>Immediate Payoffs</td>
<td>Payment</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>Delayed Payoffs</td>
<td>Signaling</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human Capital Advancement</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Social Motivation</td>
<td>Action Significance by Values</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Action Significance by Norms &amp; Obligations</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Indirect Job Feedback</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>
“Social Desirability Bias in Reports of Motivation for US and India Workers on Mechanical Turk”

Judd Antin and Aaron Shaw

CSCW 2011
“Breaking Monotony with Meaning”

D. Chandler and A. Kapelner

More people were induced to work for a meaningful task.
“Financial Incentives and the `Performance of Crowds’”

W. Mason and D. J. Watts

Proceedings of the First Workshop on Human Computation, 2009
Instructions

At the beginning of a task, you will be presented with a list of images taken from traffic cameras. An example list is shown below:

Your goal is to render the list chronologically from left to right and top to bottom. The sorted list is shown below:

Notice that in the sorted row, the truck on the right moves away from the camera, and the blue car on the left approaches the camera. To correctly sort the photos, you need to determine the flow of the traffic.

To reorder a list, click and drag a photo to the position it belongs. The other photos will move accordingly. Once you believe the list is in the correct order, click on the "Submit" button at the top of the page.

If you do not want to complete any more tasks, click on the "Finished" button at the bottom of the page.

(This button will not be available in the next 3 practice examples)

Click here to practice.
To select a word, first click on the first letter of the word, then click on the last letter of the word. If you are correct, it will turn red and the word will appear to the right of the puzzle.

For each puzzle you will see a set of possible words and their category. Not all of the words listed are in the puzzle! In addition, the number of words in each puzzle changes. The list of possible words follows:

ACHIEVE, ATTAIN, BUILDING, CHAIR, COMPETE, GREEN, LAMP, MASTER, MUSIC, PLANT, STAPLE, STEREO, STRIVE, SUCCEED, TURTLE

For this practice puzzle, you will have to find at least 8 words to continue.

RANDOM WORDS

MASTER MOZ QK KWL F
RGT RD BUILDING O T
DSUCUWRJBMOPGLC
LQRPETEMOCAEFPF
KZTCVFTBXWQVQAI
SULFJDGWWICGZO
WCENPMALARRESUZL
XVQRXO NTN LNUF NW
YHB A I ENSPLICN E F
EHLD TAVAWISC E KU
NKGKLL TEF YBE E CM
OPR PBSALIPJERNX
UVBAZ J M IDHFDGUV
QPIAWSFUNCHGIX
YKIO T W L ORNOA Q WA
"Labor Allocation in Paid Crowdsourcing: Experimental Evidence on Positioning, Nudges and Prices“

Dana Chandler and John Horton

Proceedings of the Third Human Computation Workshop, 2011
### Uptake by position and treatment

<table>
<thead>
<tr>
<th></th>
<th>Balanced Progress Bars</th>
<th>Progress Bar Imbalance on Non-Focal</th>
<th>Progress Bar Imbalance on Focal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Row</strong></td>
<td><strong>Tr0</strong> 0.36 0.14 0.17</td>
<td><strong>Tr1</strong> 0.40 0.13 0.12</td>
<td><strong>Tr2</strong> 0.56 0.13 0.11</td>
</tr>
<tr>
<td></td>
<td><strong>Tr3</strong> 0.36 0.12 0.12</td>
<td><strong>Tr6</strong> 0.30 0.06 0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Tr4</strong> 0.66 0.06 0.08</td>
<td><strong>Tr4</strong> 0.04 0.02 0.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Tr5</strong> 0.33 0.10 0.11</td>
<td><strong>Tr5</strong> 0.06 0.11 0.29</td>
<td></td>
</tr>
</tbody>
</table>

**Column**

0 1 2
"Cost-Effective HITs for Relative Similarity Comparisons “

M. Wilber, I. Kwak, and S. Belongie

Proceedings of the 2014 Conference on Human Computation
Which food on the right tastes more similar to the one on the left?

Please select the two foods that taste most similar to the food on the left.
Human experiments on foods, 5 dimensional

- Grid 4 choose 2
- Grid 8 choose 4
- Grid 12 choose 4
- Grid 16 choose 4
- Random triplets
- CKL

Generalization Error vs. Number of triplets

Total cost ($) vs. Generalization Error
"Incentives to Counter Bias in Human Computation“

B. Faltings, P. Pu, B.D. Tran, and R. Jurca

Proceedings of the 2014 Conference on Human Computation
Proposition 2 Whenever the agents’ prior belief $Pr(x)$ is equal to the publicly available distribution $R(x)$, the Peer Truth Serum makes truthful reporting a Nash Equilibrium.

Proof: Note that the expected reward for an agent who solves the task, obtains answer $x$ and reports $y$ is:

$$\text{pay}(x, y) = Pr_x(y) \cdot f(y, y, R)$$

The condition for solving the task and truthful reporting is being the best response by a margin greater than $\gamma$ is:

$$\forall x, y, x \neq y : \text{pay}(x, x) - \gamma > \text{pay}(x, y)$$

$$Pr_x(x) f(x, x, R) - \gamma > Pr_x(y) f(y, y, R)$$

$$Pr_x(x) f(x, x, Pr) - \gamma > Pr_x(y) f(y, y, Pr)$$

where $\gamma$ is the cost of effort for solving the task and obtaining answer $x$. If $f(x, x, R) = c/R(x)$ and $\gamma = ce$, the truthfulness condition is just the self-predicting condition 2. The scaling constant $c$ has to be chosen in function of the margin $\epsilon$ that can be assumed in condition 2.

Note that this reward scheme has a very intuitive nature: it rewards answers that go against the biases expressed by $R(x)$, but on the other hand still requires matching another agent’s answer, or else that only two or even one could be considered.
"Honesty in an Online Labor Market“

Siddharth Suri, Daniel G. Goldstein, and Winter A. Mason

“Bystander Intervention in Computer-Mediated Communication”

P.M. Markey

*Computers in Human Behavior* (2000)
“God Is Watching You: Priming God Concepts Increases Prosocial Behavior in and Anonymous Economic Game”

A. Shariff and A. Norenzayan

Readings for Next Time

• “Games with a Purpose”, Luis von Ahn, IEEE Computer, June 2006
  https://dash.harvard.edu/bitstream/handle/1/4000812/Jain_Game.pdf?sequence=2