Lecture 6:
CS 6306 / INFO 6306:
Advanced Human Computation

Game Theoretic Approaches

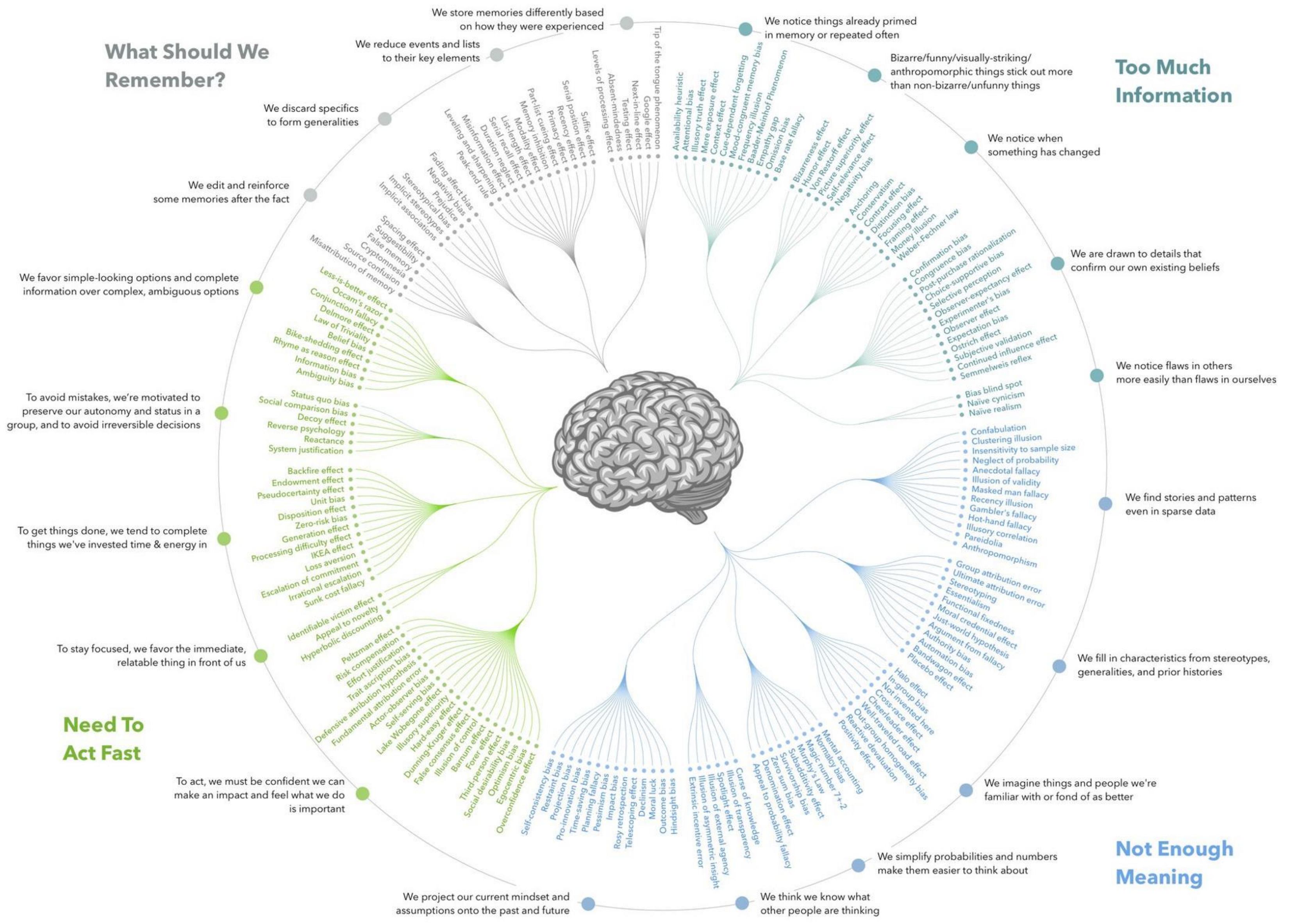
Administrivia

- Online resources:
    - Submission of reaction papers to Piazza
    - Submission of all else via CMS
- No class October 4
- Future weeks

Cognitive bias cheat sheet

https://betterhumans.coach.me/cognitive-bias-cheat-sheet-55a472476b18

COGNITIVE BIAS CODEX, 2016



ALGORITHMIC LAYOUT + DESIGN BY JM3 · JOHN MANOOGIAN III // CONCEPT + METICULOUS CATEGORIZATION BY BUSTER BENSON // DEEP RESEARCH BY WIKIPEDIANS FAR + WIDE

This Week's Readings

- Required readings:
  - Faltings, B., Jurca, R., Pu, P. and Tran, B.D., 2014. "Incentives to counter bias in human computation." In *Second AAAI Conference on Human Computation and Crowdsourcing*.
  - Ghosh, A., 2013. "Game theory and incentives in human computation systems." In *Handbook of Human Computation* (pp. 725-742). Springer New York.

- Additional readings:
  - Bachrach, Yoram, Thore Graepel, Gjergji Kasneci, Michal Kosinski, and Jurgen Van Gael. "Crowd IQ: aggregating opinions to boost performance." In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pp. 535-542.
  - Dasgupta, A. and Ghosh, A., 2013. "Crowdsourced judgement elicitation with endogenous proficiency." In *Proceedings of the 22nd international conference on World Wide Web* (pp. 319-330). ACM.
  - Easley, D. and Ghosh, A., 2013. "Incentives, gamification, and game theory: an economic approach to badge design." In *Proceedings of the fourteenth ACM conference on Electronic commerce* (pp. 359-376). ACM.
  - Ghosh, A. and McAfee, P., 2012. "Crowdsourcing with endogenous entry." In *Proceedings of the 21st international conference on World Wide Web* (pp. 999-1008). ACM.
  - Jain, S., Chen, Y. and Parkes, D.C., 2009. "Designing incentives for online question and answer forums." In *Proceedings of the 10th ACM conference on electronic commerce* (pp. 129-138). ACM.
  - Jain, S. and Parkes, D.C., 2013. "A game-theoretic analysis of the ESP game." *ACM Transactions on Economics and Computation*, 1(1), p.3.
  - Kamar, E. and Horvitz, E., 2012. "Incentives and truthful reporting in consensus-centric crowdsourcing." Technical report, MSR-TR-2012-16, Microsoft Research.
  - Moshfeghi, Y., Rosero, A.F.H. and Jose, J.M., 2016. "A Game-Theory Approach for Effective Crowdsource-Based Relevance Assessment." *ACM Transactions on Intelligent Systems and Technology (TIST)*, *7*(4), p.55.
  - Naroditskiy, V., Jennings, N.R., Van Hentenryck, P. and Cebrian, M., 2014. "Crowdsourcing contest dilemma." *Journal of The Royal Society Interface*,*11*(99), p.20140532.
  - Pickard, G., Pan, W., Rahwan, I., Cebrian, M., Crane, R., Madan, A. and Pentland, A., 2011. Time-critical social mobilization. *Science*,*334*(6055), pp.509-512.
  - Shah, N.B. and Zhou, D., 2015. "Double or nothing: Multiplicative incentive mechanisms for crowdsourcing." In *Advances in Neural Information Processing Systems* (pp. 1-9).
  - Shah, N.B., Zhou, D. and Peres, Y., 2015. "Approval voting and incentives in crowdsourcing." In *International Conference on Machine Learning (ICML)*.
  - Snehalkumar (Neil) S. Gaikwad, et al, 2016. "Boomerang: Rebounding the Consequences of Reputation Feedback on Crowdsourcing Platforms." In *Proceedings UIST: ACM Symposium on User Interface Software Technology.*
  - Ugander, J., Drapeau, R. and Guestrin, C., 2015. "The Wisdom of Multiple Guesses." In *Proceedings of the Sixteenth ACM Conference on Economics and Computation* (pp. 643-660). ACM.
  - Wu, W., Daskalakis, C., Kaashoek, N., Tzamos, C. and Weinberg, M., 2015, March. "Game theory based peer grading mechanisms for MOOCs." In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale* (pp. 281-286). ACM.

Background:
Homo economicus

- People are "agents" who seek to optimize (the expected value of) some "utility function" and are presumed to be "rational" = have the capacity to make all necessary inferences to find the optimal action
- Contrast with cooperative behavior, behavioral economics
- Allows theorems about behavior
  - Results are true for optimal agents, so must be true for sub-optimal agents
  - Assumptions not always valid, but provide guiding directions

Game Theory: Foundations

- Agents use “strategies” that say what to do in every situation – can be probabilistic
- Game: A set of actions and rewards for each agent
- Equilibrium:
  - Given all agents’ strategies, no agent will do better by changing strategies
  - Strategy may be probabilistic
  - Gives best-case scenario for other agents following their own self-interested strategies
- Zero-sum games:
  - Outcomes are net zero – there’s a loser for every winner
  - Best-studied example

Game Theory: Foundations

- John von Neumann:
  - 1928: “Minimax Theorem” for two-person zero-sum games
    If *a* is Player 1’s best possible outcome for the worst of Player 2’s strategies, then:
    - Player 2’s best possible payout for the worst (to Player 2) of Player 1’s strategies is –*a*
    - There is a (mixed) strategy for achieving it
  - 1944: *Theory of Games and Economic Behavior*, with Oskar Morgenstern
- John Nash:
  - Every finite game (finite players, finite pure strategies) has at least one “Nash” equilibrium

Game Theory: Foundations

- Mechanism design: Purposeful creation of “games”
  - What “incentives” will get agents to behave in some desired fashion
  - Examples:
    - How much to pay on AMT
    - Leader boards (*cf* “Reconstructing the world in 3D: bringing games with a purpose outdoors”)
- Incentives:
  - Financial vs social psychological
  - External vs internal
- Incentive-compatible mechanism design:
  - Each agent’s best outcome comes from acting consistent with the mechanism designer’s goals – incentives elicit desired information

Game Theory and Incentives in Human Computation Systems

"The Web is increasingly centered around contributions by its users […] ranging from labeling and categorization of images and other content […], to answering questions on online Q&A forums […], all the way to peer-grading homework assignments in online education. But while some human computation systems consistently attract high-quality contributions, other seemingly similar ones suffer from junk or low quality contributions, and yet others fail due to too little participation. How can we design incentives in these systems to elicit desirable behavior from potential participants?"

Game Theory and Incentives in Human Computation Systems

- What to “incentivize”:
  - (Continuing) Participation
  - Effort
  - Truthfulness
- Can prove theorems that say, for example, that putting in maximal effort or divulging private information truthfully yields maximal reward
  - Example: Vickerey auctions

Game Theory and Incentives in Human Computation Systems

- Example 1: DARPA Balloon Challenge
  - Pickard, G., Pan, W., Rahwan, I., Cebrian, M., Crane, R., Madan, A. and Pentland, A., 2011. Time-critical social mobilization. *Science*,*334*(6055), pp.509-512.
  - Incentivized finding balloons and recruiting participants
    - \$N for finding balloon, \$N/2 for recruiting someone who found a balloon, \$N/4 for recruiting the person who recruited the person who found the balloon, etc.

Proposition 1. *The recursive incentive mechanism is never in deficit (i.e. never exceeds its budget).*

Theorem 3. *Under the monotonic diffusion assumption, if each node can recruit less than $\sqrt{k}$ nodes, where k is the number of all recruited nodes that are not descendant of this node when all nodes recruit, then all nodes recruiting is a Nash equilibrium.*

Game Theory and Incentives in Human Computation Systems

- Example 2: Games with a Purpose
  - Jain, S. and Parkes, D.C., 2013. "A game-theoretic analysis of the ESP game." *ACM Transactions on Economics and Computation*, 1(1), p.3.
  - Existing design incentives (fast matching) lead to less-specific tags

Theorem 3. $((L, s_1^{\downarrow}), (L, s_2^{\downarrow}))$ *is a strict ordinal Bayesian-Nash equilibrium of the complete ESP game under match-early preferences, for every distribution over* U*, except the uniform distribution. Moreover,* $(L, s_1^{\downarrow})$ *is a strict ordinal best-response to* $(H, s_2^{\downarrow})$ *for every distribution over* U*, except the uniform distribution.*

Theorem 5. $((H, s_1^{\uparrow}), (H, s_2^{\uparrow}))$ *is a Bayesian-Nash equilibrium of the complete ESP game for Zipfian distributions over* U *with* $s \leq 1$ *and any additive utility function that satisfies rare-words preferences and any multiplicative utility function that satisfies rare-words preferences with* $r \geq 2$.

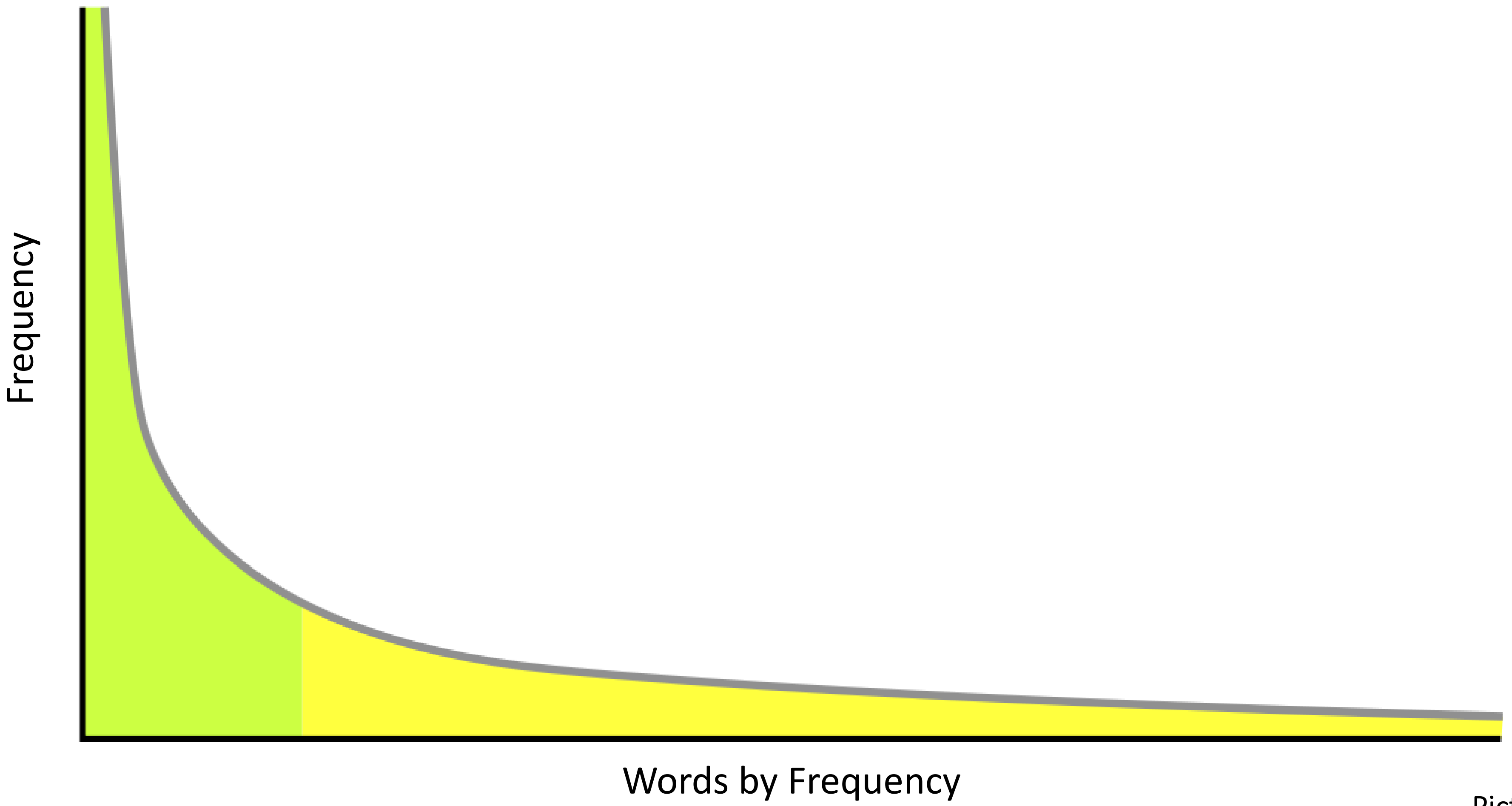
Game Theory and Incentives in Human Computation Systems

- Example 2: Games with a Purpose
  - Jain, S. and Parkes, D.C., 2013. "A game-theoretic analysis of the ESP game." *ACM Transactions on Economics and Computation*, 1(1), p.3.
  - Existing design incentives (fast matching) lead to less-specific tags

Theorem 3. *$((L, s_1^{\downarrow}), (L, s_2^{\downarrow}))$ is a strict ordinal Bayesian-Nash equilibrium of the complete ESP game under match-early preferences, for every distribution over U, except the uniform distribution. Moreover, $(L, s_1^{\downarrow})$ is a strict ordinal best-response to $(H, s_2^{\downarrow})$ for every distribution over U, except the uniform distribution.*

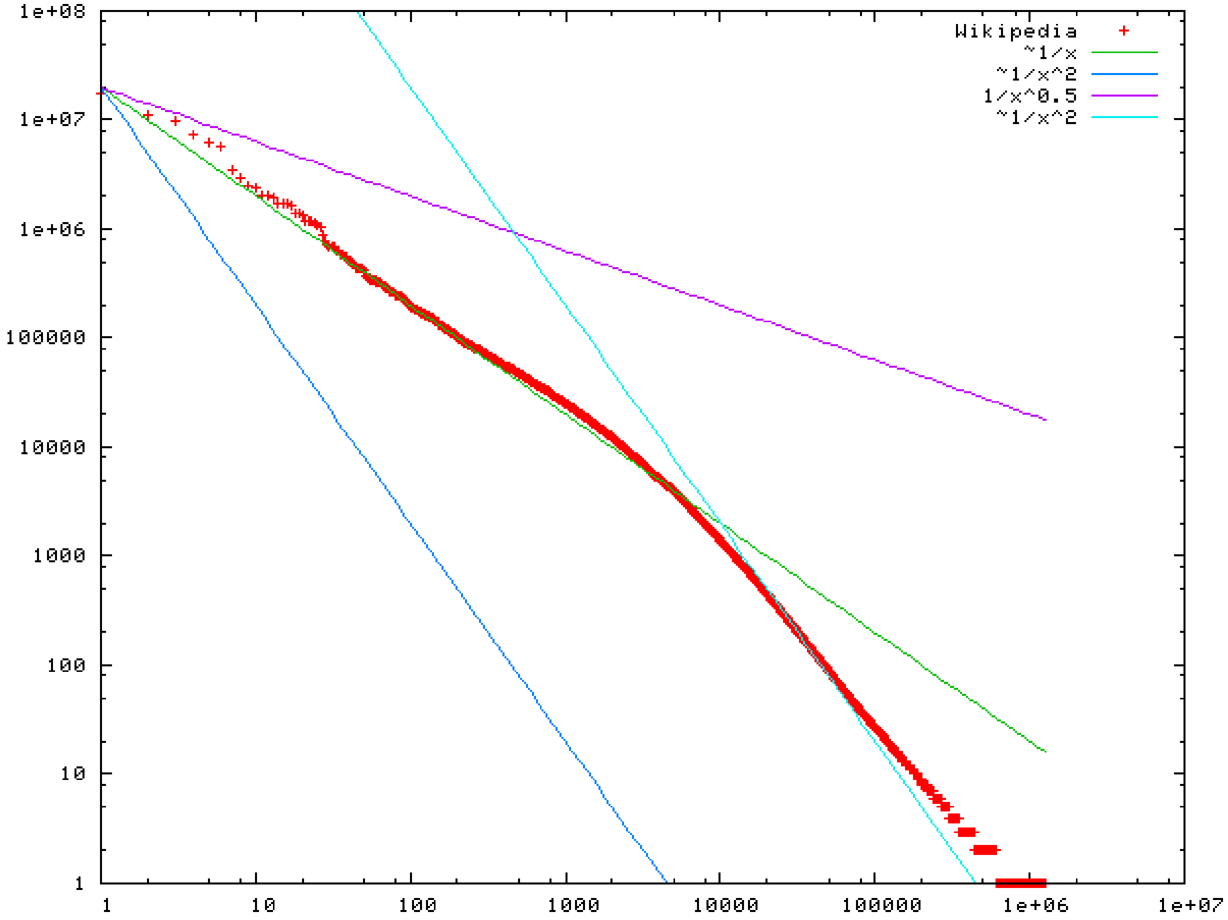
Theorem 5. *$((H, s_1^{\uparrow}), (H, s_2^{\uparrow}))$ is a Bayesian-Nash equilibrium of the complete ESP game for Zipfian distributions over U with $s \leq 1$ and any additive utility function that satisfies rare-words preferences and any multiplicative utility function that satisfies rare-words preferences with $r \geq 2$.*

Zipf's Law



Picture by Hay Kranen / PD

Zipf's Law in Wikipedia Vocabulary (actually, 3 Zipf's Laws)



From Wikipedia, LGPL

Game Theory and Incentives in Human Computation Systems

- Example 3: Correct answers to binary questions
  - Dasgupta, A. and Ghosh, A., 2013. "Crowdsourced judgement elicitation with endogenous proficiency." In *Proceedings of the 22nd international conference on World Wide Web* (pp. 319-330). ACM.
  - Proves there exists one equilibrium where all agents put in maximal effort and answer truthfully

THEOREM 6. *Suppose $p_i > 1/2$, and tasks are apriori equivalent. Then, the equilibrium where all agents choose $[(1, X)]$ yields maximum reward to each agent.*

THEOREM 8. *Suppose every agent can be a reference rater with some non-zero probability for every other agent, and tasks are apriori equivalent. Then, the only equilibria (symmetric or asymmetric) in which agents mix between $(1, X)$ and any low-effort strategy $[(0, r_{ij})]$ with non-trivial support on $[(0, r_{ij})]$ are those where all agents always use low effort on all tasks.*

Game Theory and Incentives in Human Computation Systems

- Example 3: Correct answers to binary questions
  - Dasgupta, A. and Ghosh, A., 2013. "Crowdsourced judgement elicitation with endogenous proficiency." In *Proceedings of the 22nd international conference on World Wide Web* (pp. 319-330). ACM.
  - Does not require "truth" to assess workers
    - Compares to other workers
    - Assumes multiple tasks
  - Blind agreement is an inferior outcome
  - Does not require eliciting other information from agents
    - *Cf* "Bayesian Truth Serum" - Drazen Prelec
      - Elicit subject's answer
      - Elicit estimate of what subject believes others will answer
      - Weight answers by accuracy of these estimates
      - *Requires asking estimate of what others will answer*

Game Theory and Incentives in Human Computation Systems

- Example 4: Crowdsourcing contests, user-generated content
  - Ghosh, A. and McAfee, P., 2012. "Crowdsourcing with endogenous entry." In *Proceedings of the 21st international conference on World Wide Web* (pp. 999-1008). ACM.

"We use a mechanism with monotone, rank-based rewards in a model with contributors who strategically choose both participation and quality to simultaneously capture a wide variety of crowdsourcing environments, ranging from conventional crowdsourcing contests with monetary rewards such as TopCoder, to crowdsourced content such as in Q&A forums. We first analyze the equilibria of such monotone rank-order mechanisms, and explicitly construct the unique mixed-strategy equilibrium for this mechanism (§3). We then use this construction, which explicitly gives us the equilibrium participation probability and distribution of qualities, to address the question of how to design rewards for each of the two settings previously mentioned."

Game Theory and Incentives in Human Computation Systems

- Example 4: Crowdsourcing contests, user-generated content
  - Ghosh, A. and McAfee, P., 2012. "Crowdsourcing with endogenous entry." In *Proceedings of the 21st international conference on World Wide Web* (pp. 999-1008). ACM.

Theorem 4.1. *Suppose each of the rewards a_i is constrained to lie below some maximum value A_i, $0 \leq a_i \leq A_i$, where $A_1 \geq \ldots \geq A_n$. Then, the choice of rewards $(a_1, \ldots, a_n)$ that optimizes the equilibrium distribution of qualities, and therefore the expected value of any increasing function of the contributed qualities, is*

$$
\begin{aligned}
a_i &= A_i, \quad i = 1, \ldots, n-1; \\
a_n &= \min(A_n, c(0)).
\end{aligned}
$$

Game Theory and Incentives in Human Computation Systems

- Example 4: Crowdsourcing contests, user-generated content
  - Ghosh, A. and McAfee, P., 2012. "Crowdsourcing with endogenous entry." In *Proceedings of the 21st international conference on World Wide Web* (pp. 999-1008). ACM.
  - Tax all entries and give to proceeds to the winner
  - Implementability:
    - If the system can rank the qualities of contributions, optimal outcomes can never be implemented by contests
    - With noise in quality rankings, equilibrium maximizes designer's utility

Game Theory and Incentives in Human Computation Systems

- Example 4: Crowdsourcing contests, user-generated content
  - Ghosh, A. and McAfee, P., 2012. "Crowdsourcing with endogenous entry." In *Proceedings of the 21st international conference on World Wide Web* (pp. 999-1008). ACM.
- Looking ahead – limitations to qualitative nature of many results
  - Multi-dimensional model of quality
    - Quality is not one-dimensional
    - Reward should be for the *set* of contributions
    - Different users receive different value from different contributions
    - Ratings need not be a function of quality, but also strategic value
  - Time-varying analyses
    - Users don't make simultaneous choices, base contributions on what's already there
    - Multi-task user behavior and modeling
    - Designing for sustained participation
  - Relating to human behavior

Game Theory and Incentives in Human Computation Systems

- Example 5: Ranking based on voting
  - N. Alon, F. Fischer, A. Procaccia, and M. Tennenholtz. "Sum of us: strategyproof selection from the selectors." In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge* (TARK), 2011
  - Each agent upvotes other agents
  - Want the *k* most popular agents
  - Agents want to be in the top *k* and vote selections may solely be to achieve it

Game Theory and Incentives in Human Computation Systems

- Example 5: Ranking based on voting

Theorem 3.1. *Let $N = \{1, \ldots, n\}$, $n \geq 2$, and $k \in \{1, \ldots, n-1\}$. Then there is no deterministic SP k-selection mechanism that gives a finite approximation ratio.*

Theorem 4.1. *Let $N = \{1, \ldots, n\}$, $k \in \{1, \ldots, n-1\}$. For every value of m, m-RP is SP. Furthermore,*

1. *2-RP has an approximation ratio of four, and*
2. $(\lceil k^{1/3} \rceil)$*-RP has an approximation ratio of* $1 + \mathcal{O}(1/k^{1/3})$.

Theorem 4.4. *Let $N = \{1, \ldots, n\}$, $n \geq 2$, and let $k \in \{1, \ldots, n-1\}$. No randomized GSP k-selection mechanism can yield an approximation ratio smaller than $(n-1)/k$.*

Game Theory and Incentives in Human Computation Systems

- Example 6: Badges
  - Easley, D. and Ghosh, A., 2013. "Incentives, gamification, and game theory: an economic approach to badge design." In *Proceedings of the fourteenth ACM conference on Electronic commerce* (pp. 359-376). ACM.
  - Design paramaters:
    - Badges for absolute performance vs relative performance
    - Fixed number vs percentage

Game Theory and Incentives in Human Computation Systems

- Example 6: Badges
  - Easley, D. and Ghosh, A., 2013. "Incentives, gamification, and game theory: an economic approach to badge design." In *Proceedings of the fourteenth ACM conference on Electronic commerce* (pp. 359-376). ACM.
  - Design paramaters:
    - Badges for absolute performance vs relative performance
    - Fixed number vs percentage
  - Incentivize:
    - Whether to participate
    - Level of effort

Game Theory and Incentives in Human Computation Systems

- Example 6: Badges

THEOREM 3.2 (EQUILIBRIUM EXISTENCE AND PARTICIPATION). *Consider the mechanism* $\mathcal{M}_\alpha$.

(1) *An equilibrium exists for all values of the standard* α.
(2) *There is a threshold standard* $\alpha_{\max}$ *such that all agents participate with non-trivial effort when* $\alpha \leq \alpha_{\max}$, *and there is no participation for all* $\alpha > \alpha_{\max}$.
(3) *The highest payoff an agent can obtain when the absolute standard is* $\alpha_{\max}$, $\pi(n^*, v, \alpha_{\max})$, *is* w.

THEOREM 5.2. *Suppose that the value of winning depends on the fraction of the population which wins. Consider the relative standards mechanism* $\mathcal{M}^{p}_{\rho}$, *which rewards the top* ρ *fraction of the population. If* $\rho \leq \rho_{\min}(\rho)$ *then there exists a mixed-strategy equilibrium with non-zero participation probability* p *and non-zero effort. If* $\rho > \rho_{\min}(\rho)$ *then there exists a pure-strategy equilibrium.*

Assessment?

- Theorems
- Very limited human subject experiments

Incentives to Counter Bias in Human Computation

- Faltings, B., Jurca, R., Pu, P. and Tran, B.D., 2014. "Incentives to counter bias in human computation." In *Second AAAI Conference on Human Computation and Crowdsourcing*.
- Motivation:
  - Human biases influence outcomes
    - Anchoring on initial answers biases subsequent answers
    - Common beliefs
    - Anchoring values
  - Combines Bayesian Truth Serum and Peer Consistency (matching a random worker)
  - Results:
    - Theory
    - Experiments on AMT

Incentives to Counter Bias in Human Computation

- Theorems:

Proposition 1 *Provided that the payoff of the best* cooperate *strategy is more than γ higher than the best* deceive *strategy, no* heuristic *strategy can be optimal.*

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.*

- Experiments: Count cameras/binoculars/phones/etc in photos
  - Bias: Priming
  - Understandable approximations to game-theoretic incentives
    - Peer confirmation
    - Bayesian truth serum
    - Peer truth serum

Incentives to Counter Bias in Human Computation

Peer Truth Serum: ”We will pay you $0.01, and a bonus if your response matches that of another worker on the same image within +/- 1 tolerance. The bonus is as follows:

- $0.01 if the matching answer is within +/-1 of the current average count.
- $0.06 if the matching answer is something different.”

This Week’s Readings

- Bachrach, Yoram, Thore Graepel, Gjergji Kasneci, Michal Kosinski, and Jurgen Van Gael. "Crowd IQ: aggregating opinions to boost performance." In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pp. 535-542.
- Jain, S., Chen, Y. and Parkes, D.C., 2009. “Designing incentives for online question and answer forums.” In *Proceedings of the 10th ACM conference on Electronic commerce* (pp. 129-138). ACM.
- Kamar, E. and Horvitz, E., 2012. ”Incentives and truthful reporting in consensus-centric crowdsourcing.” Technical report, MSR-TR-2012-16, Microsoft Research.
- Moshfeghi, Y., Rosero, A.F.H. and Jose, J.M., 2016. “A Game-Theory Approach for Effective Crowdsource-Based Relevance Assessment.” *ACM Transactions on Intelligent Systems and Technology (TIST)*, *7*(4), p.55.
- Naroditskiy, V., Jennings, N.R., Van Hentenryck, P. and Cebrian, M., 2014. “Crowdsourcing contest dilemma.” *Journal of The Royal Society Interface*,*11*(99), p.20140532.
- Shah, N.B. and Zhou, D., 2015. “Double or nothing: Multiplicative incentive mechanisms for crowdsourcing.” In *Advances in Neural Information Processing Systems* (pp. 1-9).
- Shah, N.B., Zhou, D. and Peres, Y., 2015. “Approval voting and incentives in crowdsourcing.” In *International Conference on Machine Learning (ICML).*
- Snehalkumar (Neil) S. Gaikwad, et al, 2016. “Boomerang: Rebounding the Consequences of Reputation Feedback on Crowdsourcing Platforms.” In *Proceedings UIST: ACM Symposium on User Interface Software Technology.*
- Ugander, J., Drapeau, R. and Guestrin, C., 2015. “The Wisdom of Multiple Guesses.” In *Proceedings of the Sixteenth ACM Conference on Economics and Computation* (pp. 643-660). ACM.
- Wu, W., Daskalakis, C., Kaashoek, N., Tzamos, C. and Weinberg, M., 2015, March. “Game theory based peer grading mechanisms for MOOCs.” In *Proceedings of the Second (2015) ACM Conference on Learning@ Scale* (pp. 281-286). ACM.

Next Week

- Required readings:
  - Martin, D., Hanrahan, B.V., O'Neill, J. and Gupta, N., 2014. "Being a turker." In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*(pp. 224-235). ACM.
  - Irani, L.C. and Silberman, M., 2013. "Turkopticon: interrupting worker invisibility in Amazon Mechanical Turk." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 611-620). ACM.

- Additional readings:
  - Borokhovich, M., Chatterjee, A., Rogers, J., Varshney, L.R. and Vishwanath, S., 2015. "Improving impact sourcing via efficient global service delivery." In *Proceedings Data for Good Exchange (D4GX).*
  - Brawley, A.M. and Pury, C.L., 2016. "Work experiences on MTurk: Job satisfaction, turnover, and information sharing." *Computers in Human Behavior*, *54*, pp.531-546.
  - Gray, M.L., Suri, S., Ali, S.S. and Kulkarni, D., 2016. "The crowd is a collaborative network." In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 134-147). ACM.
  - Gupta, N., Martin, D., Hanrahan, B.V. and O'Neill, J., 2014. "Turk-life in India." In *Proceedings of the 18th International Conference on Supporting Group Work* (pp. 1-11). ACM.
  - Kokkalis, N., Köhn, T., Pfeiffer, C., Chornyi, D., Bernstein, M.S. and Klemmer, S.R., 2013. "EmailValet: Managing email overload through private, accountable crowdsourcing." In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1291-1300). ACM.
  - Lee, M.K., Kusbit, D., Metsky, E. and Dabbish, L., 2015. "Working with machines: The impact of algorithmic and data-driven management on human workers." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1603-1612). ACM.
  - McInnis, B., Cosley, D., Nam, C. and Leshed, G., 2016. "Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk." In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2271-2282). ACM.
  - Salehi, N., Irani, L.C., Bernstein, M.S., Alkhatib, A., Ogbe, E. and Milland, K., 2015. "We are dynamo: Overcoming stalling and friction in collective action for crowd workers." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1621-1630). ACM.