CS269I: Incentives in Computer Science
Lecture #6: Incentivizing Participation*

Tim Roughgarden†

October 12, 2016

1 Some Simple Models of Participation

Last lecture, we discussed the broken incentives in Gnutella, where a Prisoner's Dilemma-type scenario gives users a strong incentive to free ride. This is one version of a broader issue, that of incentivizing *participation*. (In Gnutella, we can equate uploading with participating.) As you can imagine, incentivizing high levels of participation is a first-order concern for many systems and businesses. What are some examples?[1]

Last lecture, we saw that the Prisoner's Dilemma provides a simple way of thinking about the conflict between individual incentives and the collective good, and through its lens we can recognize this conflict in many real-world situations. We next go through some analogous "mathematical stories" to help us think about obstacles to achieving high levels of participation.

1.1 Coordination Games

The Prisoner's Dilemma shows that when participation is costly, it can be a dominant strategy for every player to not participate.[2] But even when individual incentives are squarely aligned with the collective good, it can be tricky to achieve Pareto optimality. This point is illustrated by the following *coordination game*.

Consider two best friends, Alice and Bob. There's a party tonight, and each wants to go if and only if the other one goes. And it's a good party, so they'll have more fun going out than staying home. The situation can be summarized with the following payoff matrix:

†Department of Computer Science, Stanford University, 474 Gates Building, 353 Serra Mall, Stanford, CA 94305. Email: `tim@cs.stanford.edu`.

[1] E.g., social networks, meet-ups, product adoption, etc.

[2] Recall that a *dominant strategy* is one that always maximizes a player's payoff, no matter what other players do.

	party	stay home
party	$2, 2$	$-1, 0$
stay home	$0, -1$	$0, 0$

Recall that in each entry of the matrix, the first number is the payoff to Alice and the second is the payoff to Bob.

Is this coordination game just another version of the Prisoner's Dilemma, in disguise?[3] To see that it isn't, note that neither player has a dominant strategy—the best course of action depends on what the other player does (each player wants to copy the other).

Recall that a *Nash equilibrium* is an outcome in which each player is playing a best response to the other's strategy—equivalently, neither player can strictly increase his or her payoff by unilaterally switching to the other strategy. The Prisoner's Dilemma only had one Nash equilibrium (with both players defecting), but here we have two: (party, party) and (stay home, stay home). The first of these is Pareto optimal, while the second is not. As we'll see, this is often a general issue in incentivizing participation: multiple equilibria, with wildly varying participation levels across them.

1.2 Technology Adoption and Network Cascades

The coordination games one encounters in real life often involve more than two players. One example is in the decision of whether or not to adopt a new technology, such as joining a new social network. One simple way to model the preference of a player i is to assume that the player prefers to adopt the technology if and only if at least k_i of her friends also adopt the technology. (In the two-player coordination game, $k_i = 1$ for both players.) Typically, such a technology adoption game will have at least two equilibria, one with all players adopting and one with all players not adopting.

One interesting phenomenon in such models is *network cascades*, which explain how ideas and products can "go viral" and why it's so hard to predict which ideas and products will catch on. This has been a hot topic in social networks over the past decade (e.g., look up the "influence maximization" problem), but we won't have time to cover this (see CS224W instead). But to quickly illustrate the point, consider the friendship network in Figure 1, with vertices (i.e., people) labeled with their k_i-values. What happens if we somehow coerce the leftmost vertex v into adopting the technology (e.g., by giving them a free copy or other incentives)? Nothing: since no one else will adopt until at least two friends have done so, the idea never catches on. But suppose we lower k_w from 2 to 1—now what happens? Now the technology goes viral—after v starts using it, it gets adopted by w, then by x, then by y, and then by z. Thus tiny changes to a system can have a radical effect on the ultimate outcome.

[3]For those keeping score, the only difference is that the $(0, -1)$ entries here were $(3, -1)$ entries in the Prisoner's Dilemma.

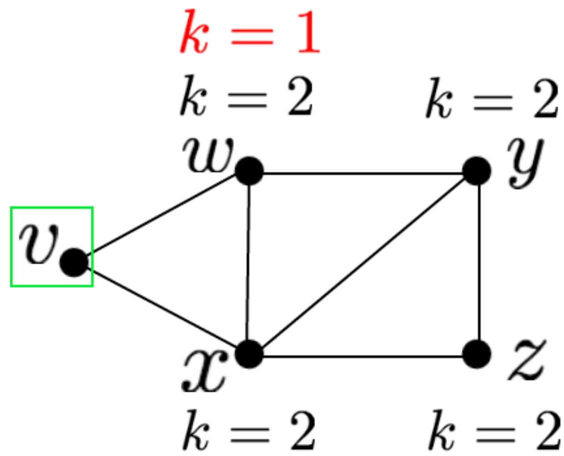

Figure 1: When $k_w = 2$, v's adoption has no ripple effects. When $k_w = 1$, v's adoption triggers a cascade, with all vertices eventually adopting.

1.3 Underinvestment in a Public Good

In our last simple model illustrating challenges in incentivizing participation, the payoff structure more closely resembles the Prisoner's Dilemma than the coordination games considered above. The issue will not be coordinating between multiple equilibria; rather, there will be a direct conflict between individual incentives and the collective good, with no Pareto-optimal equilibrium.

A *public good* is something that benefits everyone. Traditionally, one might think about infrastructure, like a park or a highway, or about the environment. In the context of this class, think about Wikipedia. Whenever anybody writes a good Wikipedia article, the whole world benefits. But writing a good Wikipedia article is a time-consuming activity for the individual. How does this tension play out?

Here's a simple model:[4]

- There are n participants.
- Each participant i chooses an investment level $x_i \geq 0$. (E.g., how much time to spend writing Wikipedia articles, or how many files to upload to a peer-to-peer network.)
- Participant i suffers a cost of x_i^2.
- Participant i enjoys a benefit of $\sum_{j=1}^{n} x_j$.

Thus a contribution by an individual benefits everyone (linearly) and costs only that individual (quadratically).

[4]The qualitative conclusions we'll draw remain valid for a wide range of models.

Assume that an individual sets her investment level x_i to maximize her benefit minus her cost, $\sum_{j=1}^{n} x_j - x_i^2$. To identify the optimal investment level, we take the derivative of this expression (i.e., $1 - 2x_i$), set it to 0, and solve to obtain $x_i = \frac{1}{2}$.[5,6] If all individuals do this, then everyone enjoys a utility of $\frac{n}{2} - \frac{1}{4}$.

The outcome of individual optimization is nowhere near Pareto-optimal. To see this, consider the problem of maximizing the collective utility

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{n} x_j - x_i^2 \right).$$

By symmetry, there is an optimal solution in which all of the x_i's are equal to a common value, call it y. Taking the derivative with respect to y and setting it equal to zero (or, equalizing the marginal collective cost with the marginal collective benefit) yields

$$\underbrace{\frac{\partial}{\partial y} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} y \right)}_{=n^2} = \underbrace{\frac{\partial}{\partial y} \left(\sum_{i=1}^{n} y^2 \right)}_{=2ny}$$

and hence $y = n/2$. Note that this solution has every individual investing n *times as much* as the individually optimal level. At this collectively optimal investment level, every participant earns utility $n \cdot \frac{n}{2} - (\frac{n}{2})^2 = \frac{n^2}{4}$, far more than in the individually optimal outcome as n grows large.

1.4 Network Effects

One buzzword for what's going on in all of these examples is *network effects.*

Vocabulary Lesson

network effect (n.): the effect that one user of a good or service has on the value of that good or service to its other users.

All of our examples today have *positive* network effects, meaning that the value of a good or service increases with the number of other users. Next lecture we'll talk about congestion, which is an example of negative network effects.

[5] Actually, this is only the "first-order condition," which in general is a necessary but not sufficient condition for optimality. But since the expression $(\sum_{j=1}^{n} x_j) - x_i^2$ is concave (with negative second derivative), the first-order optimality conditions are also sufficient for global optimality.

[6] Or, as an economist might put it, choose the investment level that equalizes the marginal cost and marginal benefit. ("Marginal" is just economics-speak for "derivative of.")

1.5 Summary

The examples in this section show that incentivizing participation is both important and difficult. In easier cases like the coordination and technology adoption games, the problem is to help participants coordinate on the Pareto optimal equilibrium. In harder cases like investment in a public good, where individual incentives conflict with the collective good, the problem is often intractable without some kind of regulation or binding collective agreement.

2 Case Study: Badge Design

2.1 Preamble

For the rest of this lecture, we take a more detailed look at a specific way of incentivizing participation: *badges* [1]. Many online systems give active users publicly viewable badges to recognize their contributions. Stack Overflow is a canonical example, and any site that crowdsources its content is a candidate for the deployment of badges.[7]

We focus on badges that reward a specific type of activity—e.g., a badge for answering questions, earned after submitting 10 different answers with at least 10 upvotes each. (Stack Overflow used to use badges in this way; now their system is hilariously labyrinthine.)

Before proceeding to a mathematical model to reason about badges, let's identify what phenomena we want the model to capture. Here are some empirical findings on Stack Overflow [1] and Coursera [2]:

1. Badges increase participation levels. Specifically, badges were introduced as a pilot program in Andrew Ng's machine learning class the third time it was offered. While forum activity was essentially the same in the first two offerings, it was significantly larger in the third offering (despite the number of students enrolled being similar).

2. Badges alter the mix of activities that users undertake, steering activity to the types that are rewarded by badges.

3. Activity accelerates as a user gets close to earning a badge.

On the third point, Figure 2 (taken from [1]) vividly displays this acceleration, as well as the immediate drop-off in activity levels after earning the badge.

Is there a mathematical model that replicates these empirical findings? The trick will be to find a model that is simultaneously simple enough to analyze and rich enough to generate the observed behavior above.

[7] Wikipedia is another interesting example. To be promoted to a top-level editor, the candidate must submit a portfolio of accomplishments (pages created, improved, etc.) reminiscent of a tenure case, and the appointment is then voted on by a committee of editors.

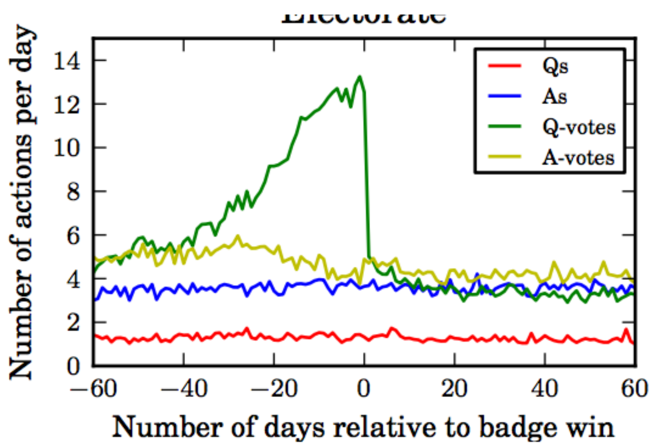

Figure 2: The activity with an associated badge (here, "Q-votes") accelerates as a user get close to earning the badge. (Figure taken directly from Anderson et al. [1].)

2.2 The Model

For this lecture, we'll assume that there is only one badge that a user can earn.[8] We assume that this badge is given after T "successes." A success might be writing an answer that garners at least 10 upvotes, for example.

A user chooses, at each stage i, an investment level $q_i \in [0, 1]$. (E.g., corresponding to the amount of time spent writing an answer.) We interpret q_i as the probability of a success at stage i. That is: the user gets a new success at stage i with probability q_i, and does not get a new success with probability $1 - q_i$. At any given stage, we can associate a user with a "state" $s \in \{0, 1, 2, \ldots\}$ equal to the number of success that the user has had thus far. (Once $s \geq T$, the user gets the badge.)

How do we expect a user to behave (i.e., choose q_i's)? How should we model the preferences of a user? We proceed by making three assumptions. First, we assume that a user has some value v for obtaining the badge.[9] Different users can have different values; this won't matter to us, because we can analyze each user separately.

There is currently nothing to deter a user from exerting maximum effort ($q_i = 1$) to earn the badge as quickly as possible (after T stages). Our second assumption is that a user has a preferred activity level p (which is the same in all stages). Thus a lazy user has p close to 0, and an eager user a p close to 1. We also assume that there is a cost function h, specifying the cost $h(p, q)$ to the user when it plays q instead of p. For a concrete example, you can think of $h(p, q)$ as $(p - q)^2$. (In any case, assume that $h(p, p) = 0$ and $h(p, q) > 0$ for all $q \neq p$.)

Combining the first two assumptions, we see that the utility earned by the user in stage i

[8]See Exercise Set #3 for extensions to multiple badges, both for the same activity (at different levels) and for different activities.

[9]Some people find it hard to believe that anyone ascribes value to a fictitious object like a badge, but empirically many people behave as if badges have non-trivial value.

is

$$\begin{array}{rl} v - h(p,q) & \text{if the user has the badge when stage } i \text{ commences;} \\ -h(p,q) & \text{otherwise.} \end{array}$$

Note that the value v for the badge is accrued every day after receiving it. (Changing the model so that it's a one-time gain would yield similar results.)

Finally, we need to decide how a user reasons about the utility she receives across the different stages. It would be uninteresting and unrealistic to assume that a user is completely myopic, optimizing only the utility at the current stage. This would result in the user setting $q = p$ every day, perhaps inadvertently earning the badge after enough stages. Badges have the potential to influence behavior only if users engage in some kind of long-term planning. If users care equally about utility in all stages, then to first order all strategies are equally good: all of them will lead to the badge eventually (assume $q > 0$ infinitely often), resulting in a long-term average payoff of v. But some strategies will get the badge sooner than others, and all else being equal we expect a user to prefer to get the badge sooner.

We'll model this idea as in our second model for the repeated Prisoner's Dilemma (Lecture #5), where after each stage the game ends with some probability, and players strive to maximize expected payoffs. In the present context, maybe with some probability you get a new job and no longer have any time for Stack Overflow.

Formally, for a *discount rate* $\gamma \in [0, 1]$, or equivalently the probability that the game doesn't end after a given stage, we define the discounted utility of a user as

$$\text{utility now} + \gamma \cdot \underbrace{(\text{utility later})}_{\text{defined recursively}}.$$

Unrolling the recursion, this is

$$\sum_{i=1}^{\infty} [\text{utility at stage } i] \cdot \gamma^{i-1}. \tag{1}$$

This completes the description of the user model. We assume that a user chooses investment levels $q_1, q_2, \ldots$ to maximize her discounted utility (1).

2.3 Optimal Investment Levels

Given the model above, how does a user behave? Note that a user must compromise between two competing desires—she doesn't want to deviate from her preferred investment level p, but she wants to earn the badge sooner rather than later. Intuitively, some tension of this form is necessary to reproduce the empirical findings listed in Section 2.1. Even though our model is quite simple, it is not trivial to determine what a user will do. This is our next goal.

First, some notation. For a state $s \in \{0, 1, 2, \ldots, \}$, indicating the number of successes that a user has had thus far, we define U_s as the discounted utility earned by the user starting at a stage with state s, assuming optimal future play. Note that U_s depends only on the starting state s, and not on the history of how the user arrived at the state. Because everything in our model is history-independent (given the current state), we can also focus on user strategies that are history-independent (i.e., that depend only on the current state).[10]

The plan is to solve for all the U_s's, and along the way the optimal investment level q_s^* at a stage when the user currently has the state s. To begin, suppose $s \geq T$, meaning that the user already has the badge. In this case, there is nothing more to gain, so the optimal strategy is to minimize cost and set $q_s^* = p$. The discounted utility starting from now (under this optimal play) is

$$\begin{aligned} U_s &= v + \gamma v + \gamma^2 v + \cdots \\ &= \frac{v}{1-\gamma}. \end{aligned}$$

Now consider the state $s = T - 1$. Suppose the user chooses the investment strategy q now and then plays optimally at all subsequent stages. Recalling that q is the probability of getting a new success at this stage (moving from state $T-1$ to state T), we have

$$\underbrace{\text{expected discounted utility}}_{\text{playing } q \text{ now, optimally later}} = \underbrace{-h(p,q)}_{\text{utility now}} + \gamma \cdot \underbrace{[qU_T + (1-q)U_{T-1}]}_{\text{expected utility later}}.$$

At the optimal investment level q_{T-1}^*, the left-hand side becomes U_{T-1} (by the definitions). That is, q_{T-1}^* is the argmax in the equation

$$U_{T-1} = \max_{q \in [0,1]} \left\{ -h(p,q) + \gamma \cdot [qU_T + (1-q)U_{T-1}] \right\},$$

or after rewriting

$$U_{T-1} = \max_{q \in [0,1]} \left\{ \frac{-h(p,q) + \gamma q U_T}{1 - \gamma + q\gamma} \right\}.$$

Recalling that we've already solved for U_T (it's $v/(1-\gamma)$), the right-hand side is entirely known except for the independent variable q. Solving for q by hand is not trivial, but one line of Matlab (the `fmincon` function) usually does the trick.

Applying exactly the same reasoning, we have

$$U_{T-2} = \max_{q \in [0,1]} \left\{ \frac{-h(p,q) + \gamma q U_{T-1}}{1 - \gamma + q\gamma} \right\}.$$

Given that we've already solved for U_{T-1}, we can again solve for U_{T-2} with one line of Matlab. Proceeding in this way, we can compute all of $U_T, U_{T-1}, \ldots, U_0$, and the corresponding optimal investment levels (the argmaxes) $q_T^*, q_{T-1}^*, \ldots, q_0^*$.[11]

[10] This is a general property of *Markov decision processes (MDPs)*, of which the current model is a special case.

[11] This can be thought of as backward induction (as in Lecture #5) or a form of dynamic programming.

2.4 Example

Solving the sequence of optimization problems above gives sensible answers. For example, suppose $v = 1$, $T = 5$, $\gamma = .75$, $h(p, q) = (p - q)^2$, and $p = \frac{1}{2}$. Using the backward induction strategy outlined above, we can compute the optimal investment levels and discounted utilities as:

	$s = 0$	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s \geq 5$
U_s	.47	.76	1.19	1.84	2.75	4
q_s^*	60%	66%	74%	84%	97%	50%

Already in this simple example, we can clearly see the same effects that were observed empirically (Section 2.1). Without any badges in the system, investment levels would always be 50%. The badge, as the carrot on the stick, leads to higher investment levels at all of the stages prior to earning the badge. More remarkably, these numbers resemble the acceleration of activity observed as users get close to earning badges (Figure 2). Early on, when the badge is far away (like a mirage in the desert), it exerts only a modest pull on investment levels. As the prospect of earning the badge becomes increasingly real, investment levels rise accordingly, reaching a crescendo of nearly 100% when the badge is just out of reach. As in Figure 2, after earning the badge the activity level falls of a cliff, returning to the preferred level.

2.5 Final Comments

The user model and our method of solving for optimal user behavior can be extended in various ways. For example, there can be multiple badges. One version of multiple badges is having multiple milestones at different levels of the same single activity (think of frequent flyer programs). A second version is to have different badges for different activities. (Then, both the investment level $\mathbf{q}$ and the preferred level $\mathbf{p}$ become vectors, with one component per activity. Similarly, the state is then a vector.) This second version introduces some new and interesting user behavior, with different badges pulling investment levels in different directions at the same time.

Less well understood is the problem of badge design. To optimize a given objective (e.g., some mixture of activity levels over a user population), how should badges be implemented? How many should there be, what should the thresholds T be, etc.? One interesting experimental finding in [1] is that very hard-to-get badges can be surprisingly useful, since they can increase participation levels over a very long period of time. Of course, it can also be a good idea to have intermediate milestones along the way.

References

[1] A. Anderson, D. P. Huttenlocher, J. M. Kleinberg, and J. Leskovec. Steering user behavior with badges. In *Proceedings of the 22nd International World Wide Web Conference (WWW)*, pages 95–106, 2013.

[2] A. Anderson, D. P. Huttenlocher, J. M. Kleinberg, and J. Leskovec. Engaging with massive online courses. In *Proceedings of the 23rd International World Wide Web Conference (WWW)*, pages 687–698, 2014.