CS269I: Incentives in Computer Science
Lecture #19: Time-Inconsistent Planning*

Tim Roughgarden†

December 5, 2016

1 Utility Theory and Behavioral Economics

In almost all of the models that we've studied in this course, we formally defined the *preferences* of the players. To predict how someone acts (e.g., their "best response"), we need to specify what they want. This generally involved specifying utilities or payoffs for the players, and assuming that each player acts to maximize her own utility.[1] (If there is uncertainty, then she maximizes her expected utility with respect to an appropriate distribution.) This is the approach of classical utility theory.

Even the purest of utility theorists don't believe that a typical individual could articulate her entire utility function (and possibly prior distribution) on demand. Rather, the assumption is that players behave *as if* they were maximizing utility, whatever the internal decision-making mechanism might be. How could we test this hypothesis?

When a player really is maximizing some utility function, several properties of their behavior follow. For instance:

1. *Transitivity.* If a player prefers action A to action B, and action B to action C, then she also prefers action A to action C.

2. *Monotonicity.* If we give a player additional actions, she can only be better off. (She could always ignore the new options and do the same thing as before, if desired.)

These are properties that can be tested. And it turns out that, across a range of different experiments, the behavior of people consistently fails these tests. Such empirical findings motivate a reconsideration of the classical utility model.

---

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

[1]In the first couple of weeks we worked with abstract rankings, but these can still be thought of as utility functions of a sort.

Behavioral economics, which has been increasingly in vogue in this century, proposes models of behavior to explain systematic deviations from the predictions of classical utility theory. It is then interesting to ask about the implications of these models for policy, system design, and so on. *Prospect theory*, developed by Kahneman and Tversky [3, 7], is a major part of behavioral economics.[2] The theory includes several parts, including the assumption that people tend to overestimate small probabilities, and "loss aversion," meaning that people tend to value losses (relative to some reference point) more than the analogous gains.

In this lecture we'll focus on a different systematic deviation from classical utility theory: *time-inconsistency*.

2 A Model of Time-Inconsistent Planning

2.1 Procrastination

We next outline a "theory of procrastination." It's easiest to start with an example, so we'll recount an autobiographical story by Akerlof [1]. He was spending a year in India, and he had an errand to do—shipping a friend's box back to the States. This was not a trivial thing to do (long lines at the post office, etc.), but he had to do it. Every morning when he woke up, he had every intent of shipping the box *tomorrow*. Then tomorrow would come, and the logic would repeat itself. The result was months of delays before Akerlof shipped the box.

Here's a simple model that can explain what happened. Suppose the cost of delaying shipping was 1 unit per day of delay. Suppose the cost of shipping the box was $c > 1$. What's the optimal (minimum-cost) thing to do? The optimization problem is to choose the delay $t \geq 0$ to minimize $t + c$. (Not exactly a hard optimization problem—just set $t = 0$.) But somehow Akerlof did not choose this optimal solution, and instead incurred cost linear in the time horizon.

The key assumption for today's lecture is *present bias*: the cost incurred on the same day of a decision is *perceived* as b times as large as its actual cost, where $b > 1$ is the degree of present bias. This model explains Akerlof's behavior: at a given day $t \geq 0$, the incremental cost of shipping on that day would be perceived as $b \cdot c$, while the incremental cost of delaying one day and shipping the next day is $b \cdot 1 + c = b + c$, which is less for many choices of b, c (e.g., $b = 2$ and $c = 3$). Note that in the second case, the cost of shipping (tomorrow) is just c, and not $b \cdot c$—the player really does consider the current day as special, and naively assumes that in the future she will be perfectly rational.

For computer scientists, it's natural to illustrate Akerlof's story using a graph. (For example, in AI, planning problems are often modeled as graph search.) In Figure 1, each vertex represents a current state, and each edge an action. The edges around the boundary of the "fan" correspond to delaying for a day, and each edge to t corresponds to shipping the box. Edges are labeled with their (actual) costs. When the player is at a given vertex v, she computes a shortest path to t, but with a twist: edges going out of v are scaled up by a factor

[2] Balaji Prabhakar mentioned this in his guest lecture (Lecture #11), and the "nudges" in his mechanisms for social good are based in part on insights from prospect theory.

of b before the computation. If the resulting shortest path P has (v, w) as its first edge, then the player moves to w and does the same computation anew. Akerlof's example shows that the player might change her mind about which path to take after traversing the first hop (e.g., delaying again tomorrow, rather than shipping tomorrow as originally planned). This never happens with conventional shortest-path computations (subpaths of shortest paths are again shortest paths).

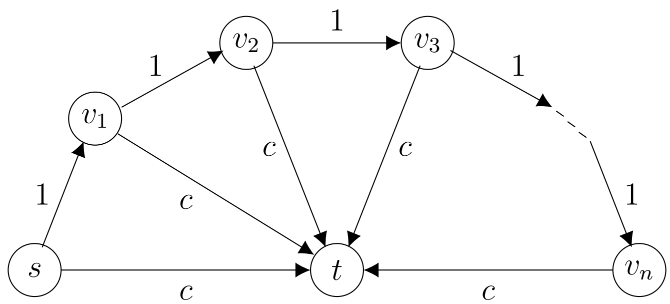

Figure 1: Graph model of Akerlof's story.

2.2 The Model

Here's the general model [4]:

- There is a directed acyclic graph G, with a unique source vertex s and a unique sink vertex t.
- Each edge e of G has a nonnegative cost c_e.
- At each vertex v, the player chooses the v-t path P that minimizes

$$b \cdot (\text{cost of next hop}) + (\text{cost of all subsequent hops}),$$

  and then traverses the first hop of P (and then repeats the same computation).

A related concept in behavioral economics is "hyperbolic discounting," which combines present bias with exponential discounting. For simplicity, we assume no discounting in this lecture (just present bias), but the lessons learned extend also to models of hyperbolic discounting.

2.3 Choice Reduction

Next we use the model to show how choice reduction can be beneficial to a player. Consider the graph in Figure 2. The story is: you're taking a three-week course, and you need to complete two projects by the end of the course. Each "column" of the graph corresponds to a week of the course, and each "row" indicates the number of projects finished to-date. The cost of doing zero projects in a given week is 1; the cost of doing one project is 4; and the cost of doing both projects in the same week is 9. The minimum cost of a path is 9 (doing one project in each of two of the weeks). What will a player with present bias $b = 2$ do?

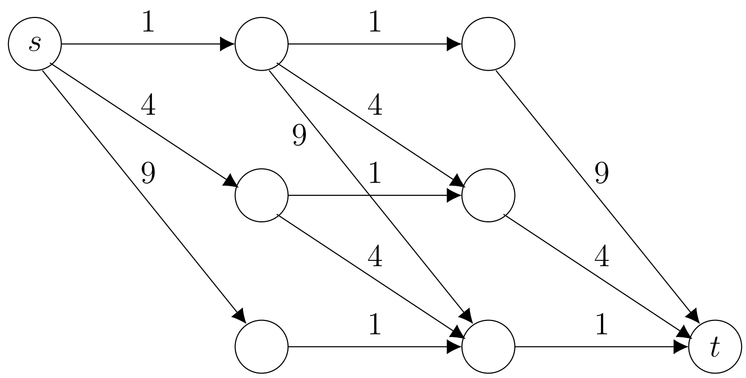

Figure 2: Graph model of a three-week course with two projects.

The first week, it would clearly be crazy to do both projects (the perceived cost would be 20). Doing one project now and one project later has perceived cost $2 \cdot 4 + 4 + 1 = 13$. Since doing nothing now and one project in each of the second and third weeks has perceived cost $2 \cdot 1 + 4 + 4 = 10$, the player will not do any projects in the first week. In the second week, it is again crazy to contemplate doing both projects. Doing one project now and the other project next week seems reasonable, with perceived cost $2 \cdot 4 + 4 = 12$. But procrastinating again (and doing both projects in the last week) has perceived cost only $2 \cdot 1 + 9 = 11$, so this is what the player will do.

The instructor of the course has an easy option to prevent such procrastination: just require that at least one of the projects be completed by the end of the second week. This has the effect of removing the upper-right vertex from the graph in Figure 2. Now, the player will again procrastinate during the first week, but will follow a shortest path, completing one project in each of the second and third weeks. Such beneficial choice reduction cannot arise with a classical utility-maximizing player.

2.4 Task Abandonment

For our next example, we extend the basic model to include a nonnegative reward at the sink t. (One could also allow intermediate rewards, if desired.) The player is no longer forced to take a path all the way to t; she does so only if the consequent reward outweighs her perceived cost. The player can stop at any time to avoid incurring any further costs.

Consider the example in Figure 3, and again assume that $b = 2$. Think of the first edge as the decision of whether or not to join a gym, and the second edge as whether or not to actually go to the gym. At the beginning, because the player gets a great deal on a gym membership, the perceived cost of going to t is 6, less than the reward of 7 of getting into sick shape. After the player gets to v, however, the reality of actually going to the gym sets in, and the perceived cost of 8 of going outweighs the reward, and the player stops at v. Again, such task abandonment (even as the underlying cost and reward structure remains the same) cannot arise with a classical utility-maximizing player.

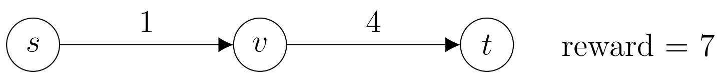

Figure 3: Present bias can lead to task abandonment.

3 The Cost of Present Bias

In all of our examples, present bias leads a player to traverse a path more costly than the optimal path. In this section we focus on the question: how much worse can the chosen path be, compared to a shortest path? This question is reminiscent of our work on the price of anarchy (Lecture #7), where we studied how bad the outcome of selfish behavior (by many players) can be, relative to the cost of a socially optimal outcome.

3.1 The Weakly Monotone Case

Define the *cost ratio* of a planning instance as the ratio between the cost of the path traversed by an agent with present bias and that of the shortest path. In Akerlof's original example (Figure 1), the cost of a shortest path is c, while the cost of the path chosen by a present-biased agent grows linearly with the number n of vertices (as roughly $n + c$). Can the cost ratio be worse than this?

We next identify a condition under which the cost ratio is also upper bounded by the number of vertices in the graph.

Definition 3.1 A planning instance G is *weakly monotone* if

$$d(v, t) \leq d(s, t) \tag{1}$$

for all $v \in V$, where $d(v, t)$ denotes the length of a shortest path from v to t.

We emphasize that the shortest path distances in (1) are with respect to the true costs (not perceived costs). Weak monotonicity does not imply that you can't go backward (i.e., find yourself further from t after traversing an edge), just that you can't ever be further from t than where you started. For example, imagine training for a marathon: if you skip some training sessions you will regress, but not beyond the point you were at when you hadn't done any training at all.

All of the examples we've considered thus far are weakly monotone (check this for the example in Figure 2). And weakly monotone instances cannot have a superlinear cost ratio.

Proposition 3.2 ([2]) *In every weakly monotone planning instance, for any present bias $b \geq 1$, the cost ratio is at most the number of vertices n.*

Proof: Fix such an instance and let P be the path chosen by the player. Since P has at most n (really, $n-1$) edges, it is enough to show that each edge of P has cost at most $d(s,t)$. To see why this is true, fix an edge $e = (v,w) \in P$. When the player is at the vertex v, the perceived cost of the shortest v-t path (where by "shortest" we mean w.r.t. the true edge costs) is at most $b \cdot d(v,t)$, with the worst case occurring when only the first edge of the shortest v-t path has a nonzero cost. On the other hand, the perceived cost of following the path P from v to t is at least that of the first hop, namely $b \cdot c_e$. Since the agent chose to follow the path P instead of the shortest v-t path,

$$b \cdot c_e \leq b \cdot d(v,t).$$

Hence $c_e \leq d(v,t)$ and, by weak monotonicity, $c_e \leq d(s,t)$, as desired. ■

3.2 Lower Bound for the General Case

Is every graph weakly monotone? No, and in general the cost ratio can be much worse than in Akerlof's procrastination example. To see this, consider the example shown in Figure 4. Like in Akerlof's example, every day the player plans to procrastinate for one day and then go to t tomorrow. Unlike Akerlof's example, the cost of going to t escalates by a b factor every time the player procrastinates. The cost of a shortest path in this graph is c, while the player chooses a path with cost $b^n c$, and so the cost ratio is exponential in the number of vertices.

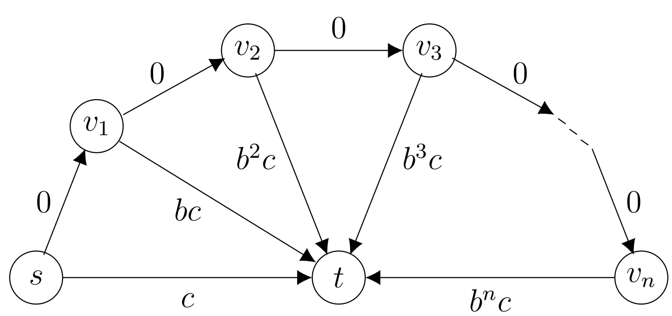

Figure 4: Present bias can lead to an exponential cost ratio.

Several of the other engaging tales spun in Akerlof [1] match up pretty well with the terrifying example in Figure 4, with the running theme being how passivity, complacency, or undue obedience can lead to an escalating crisis and a very costly endgame. He discusses drug addition (with the difficulty of quitting increasing over time), the sociology of LBJ's war room during the escalation of violence in the Vietnam War, the recruitment and control of members in cults like the Moonies and Synanon, and Stalin's takeover of the Bolshevik party.[3]

In light of this example with an exponential cost ratio, what kind of positive news could we hope for? The next result provides at least some consolation: the *only* way that a planning problem can have an exponential cost ratio is if it includes the example in Figure 4. In the (semi-)formal statement, we use "F_k" to refer to the graph in Figure 4, where k denotes the number of non-sink vertices.[4]

Theorem 3.3 ([4]) *If a graph G has an exponential cost ratio, then the graph F_k is "embedded" in it, where $k = \Omega(n)$.*

Note that there's no hope of saying that G must be precisely F_k for some k, since one can add spurious vertices and edges to F_k to get further instances that have an exponential cost ratio. The formal term for "embedded in" is "is a graph minor of," but we'll skip the formal definition, which should be intuitively clear from the proof.

Proof of Theorem 3.3: Consider an instance with n vertices and an exponential cost ratio, meaning cost ratio α^n for some constant $\alpha > 1$. Our task is to exhibit structure analogous to the graph F_k in Figure 4, with k at least a constant fraction of n.[5]

Let P be the path chosen by the player. The first order of business is to show that P must be a long path (i.e., with lots of edges); the path can then play the role of the outer boundary of F_k. We'll do this by first arguing that P contains a vertex that is very far from t (in terms of true shortest-path distance), and then by showing that as the player traverses P, her distance to t can only go up at a bounded rate.

Since the cost of P is at least α^n, there is some edge $e = (x, y) \in P$ with $c_e \geq \alpha^n/n$. (Just because P has at most n, or really $n - 1$, edges.) Since the $\frac{1}{n}$ factor is annoying, let's assume that n is sufficiently large and lower bound α^n/n by β^n for some constant $\beta \in (1, \alpha)$. When the player traverses the edge e, the perceived cost of e is $b\beta^n$; since she chose e over the shortest x-t path, the shortest-path distance $d(x, t)$ must be at least β^n (as its perceived cost is at most b times this). This is a pretty massive violation of the weakly monotone property that we studied in the preceding section.

Next, we claim that traversing an edge of P can only lead the player a b factor away from t. That is, if $(u, v) \in P$, then

$$d(v, t) \leq b \cdot d(u, t). \tag{2}$$

The proof is similar to that of Proposition 3.2. At u, the player perceives the cost of the shortest u-t path as at most $b \cdot d(u, t)$. The u-t path chosen by the player, comprising (u, v) and some path from v to t, has perceived cost at least $d(v, t)$. Since the player chose the latter over the former, inequality (2) holds.

[3] Sample quote: "by acquiescing step-by-step to the crescendo of Stalin's actions, [the remaining party members] were committing themselves to altered standards of behavior" [1, page 15].

[4] See [6] for a quantitatively stronger version of this result.

[5] This bears some resemblance to an argument that we did in Lecture #8, about BGP routing, where we proved that BGP failed to have nice properties only in graphs with a "dispute wheel."

Combining (2) with the fact that the vertex x (identified above) satisfies $d(x,t) \geq \beta^n d(s,t)$, we can conclude that P must be long, with at least $k := \log_b \beta^n = \Omega(n)$ vertices.[6] More generally, for $i = 0, 1, 2, \ldots, k$, it makes sense to define v_i as the last vertex w along P that satisfies $d(w,t) \leq b^i d(s,t)$. (Inequality (2) and the fact that $d(x,t) \geq \beta^n d(s,t)$ imply that there is at least one such vertex for each such i.)

Our ingredients for the F_k structure are (Figure 5):

- The vertices $v_0, \ldots, v_k$, and t.
- The v_i-v_{i+1} subpaths Q_i of P (for $i = 0, 1, \ldots, k-1$).
- The shortest paths $P_0, \ldots, P_k$ between the v_i's and t (with respect to the true edge costs).

Our choice of the v_i's means that the Q_i's are disjoint paths (except that the endpoint of Q_i equals the starting point of Q_{i+1}). While the P_i's need not be disjoint from each other, we claim that they must be disjoint from the Q_i's (except that the starting point of P_i is the same as the ending point of Q_{i-1} and the starting point of Q_i). The path P_i cannot intersect the path P before v_i, because this would imply a cycle (while the graph G is a DAG). The path P_i also cannot intersect P after v_i: since P_i is a shortest path, the value of $d(w,t)$ only decreases along P_i, but v_i was chosen as the last vertex of P for which $d(w,t) \leq b^i d(s,t)$.

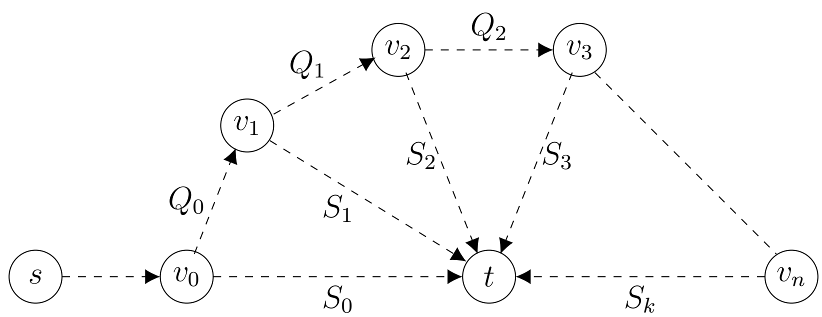

Figure 5: The F_k structure. .

We can now see the sense in which F_k is "embedded" in G: if we contract the edges of each path Q_i to a single edge (between v_i and v_{i+1}), and we contract all of the edges of each P_i except for the first hop, then the result is exactly the graph F_k in Figure 4. ■

4 Sophisticated Players

Thus far we have considered players who are *naive*, meaning not only are they present-biased, but they don't even realize that they are present-biased. (Recall in Akerlof's example, every

[6]The asymptotic notation suppresses a factor of $\log_b \beta$, which is some constant (i.e., independent of n) depending on b and β.

morning the plan was to ship the box tomorrow, as if somehow the perceived cost tomorrow would be less than it is today.) Next we consider *sophisticated* players, who are present-biased but know that they are present-biased and act accordingly. (E.g., whenever you set fictitious intermediate deadlines for yourself, you are being sophisticated and recognizing the future temptation you will have to procrastinate.) Both naive and sophisticated present-biased behavior are common in experiments (and in the real world).

To see the difference between naive and sophisticated players, consider the example in Figure 6, where b is in between c and c^2. Think of the edges (v_0, t), (v_1, t), and (v_2, t) as representing doing your homework Thursday, Friday, or Saturday. We already know what a naive agent will do: since this example is a special case of that in Figure 4, she will procrastinate and do her homework only on Saturday, for the worst-possible cost of c^2. The sophisticated agent does a cost-benefit analysis between doing her homework on Thursday or delaying until Friday, effectively delegating future decisions to her "Friday self." Doing her homework now has perceived cost b. (Remember, while sophisticated, she is still present-biased.) If she delays, she asks: "what will my Friday self do?" Since her Friday self would be choosing between doing her homework (at perceived cost $b \cdot c$) and delaying until Saturday (with perceived cost c^2), she will procrastinate (since $b > c$). This means that on Thursday, the perceived cost of delaying until Friday (and knowing the her Friday self will delay again) is $b \cdot 0 + 0 + c^2 = c^2$. Since $c^2 > b$, the sophisticated agent will do her homework on Thursday to deprive her future self of the option of further procrastination.

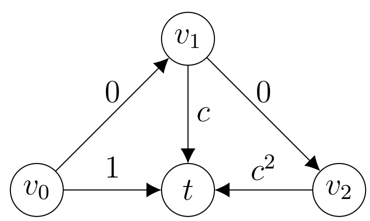

Figure 6: An example that shows the difference between naive and sophisticated players.

The formal model is as you would expect: at a given vertex v, the player chooses the next hop $e = (v, w)$ to minimize

$$b \cdot c_e + (\text{cost of what future self will do from } w).$$

This definition is not circular, because G is a DAG. The player's action at every vertex can be determined by working backward from the sink t (along some topological ordering of G's vertices). For example, in Akerlof's original example (Figure 1), a sophisticated player will ship the box immediately, provided the time horizon T is large enough (e.g., at least $b \cdot c$).

We've seen in two examples that sophisticated players make much better decisions than naive ones. Sophisticated players are not perfect, however: if we delete the vertex v_2 from the graph in Figure 6, so that there is only one decision to make, then the player procrastinates (provided $c < b$), and so the cost ratio can be arbitrarily close to the present bias b. It turns out there is no worse example.

Theorem 4.1 ([5]) *For every planning instance, the cost ratio of a sophisticated player is at most her present bias b.*

Proof: We'll show the stronger statement that, for every vertex v of G, the future cost incurred by a sophisticated player starting at v is at most $b \cdot d(v,t)$. We proceed by induction on the vertices of G, starting at the sink t and working backward (along a topological ordering of the vertices). The base case of $v = t$ is trivial. Fix some vertex $v \neq t$, and suppose that the first hop of a shortest v-t path is (v, w). By the inductive hypothesis, the cost the player will incur from w onward is at most $b \cdot d(w,t)$. Thus the perceived cost of choosing the edge (v, w) is at most $b \cdot c_{vw} + b \cdot d(w,t) = b \cdot d(v,t)$. The perceived cost of the path chosen by the player can only be less than this, and the actual cost of this path still less (since $b \geq 1$). Thus the cost incurred from v onward is at most $b \cdot d(v,t)$, completing the inductive step and the proof. ■

Theorem 4.1 is good news—the worst-case case ratio with a sophisticated player (b, independent of n) is far less than that with a naive player (exponential in n). Thus sophisticated players do better than naive players in the worst case—is this true in *every* case? Surprisingly, not only can the path chosen by a sophisticated player be arbitrarily close to b times that of a shortest path, it can even be arbitrarily close to a b factor larger than the path chosen by a naive player.

Consider the problem shown in Figure 7, where x is a large number and ϵ is close to 0. A naive player takes the path $s \to u \to v \to t$ (why?), for a cost of $x + b - \epsilon$. A sophisticated player knows that, if she reaches u, then she will procrastinate and take the path $u \to v \to t$ instead of $u \to t$. Thus from s, the perceived cost of taking the edge (s, u) is $bx + b - \epsilon$. This leads the sophisticated player to instead take the edge (s, w), in order to traverse a path with perceived (and actual) cost $bx + b - 2\epsilon$. As $x \to \infty$ and $\epsilon \to 0$, the cost of the sophisticated player's path approaches b times that chosen by the naive player.

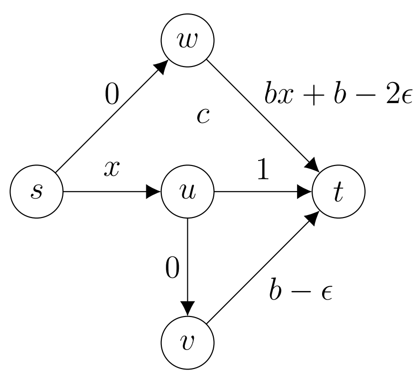

Figure 7: An example where a naive player outperforms a sophisticated player.

References

[1] G. A. Akerlof. Procrastination and obedience. *American Economic Review*, 81(2):1–19, 1991.

[2] N. Gravin, N. Immorlica, B. Lucier, and E. Pountourakis. Procrastination with variable present bias. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, page 361, 2016.

[3] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

[4] J. M. Kleinberg and S. Oren. Time-inconsistent planning: a computational problem in behavioral economics. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, pages 547–564, 2014.

[5] J. M. Kleinberg, S. Oren, and M. Raghavan. Planning problems for sophisticated agents with present bias. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, pages 343–360, 2016.

[6] P. Tang, Y. Teng, Z. Wang, S. Xiao, and Y. Xu. Computational issues in time-inconsistent planning. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017. To appear.

[7] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.