CIS 700 Differential Privacy in Game Theory and Mechanism Design February 7, 2014

Lecture 4

Lecturer: Rachel Cummings *Scribe: Rachel Cummings*

The Empirical Implications of Privacy-Aware Choice

1 Introduction

Today we'll be talking about a very different version of privacy than we have seen so far in this class. This notion of privacy is different from differential privacy in two crucial ways. First, we're going to be looking at individuals rather than databases. If we wanted to apply differentially private algorithms to a database of size one (i.e. for a single individual), we would have to output either pure noise or a constant function, thus losing all accuracy. Differential privacy requires that the output cannot be too sensitive to the input of a single individual, but if the individual is the entire database, then the output of the algorithm cannot be sensitive to the input. For this reason, it doesn't make sense to talk about differential privacy for a single individual. Second, today we're not going to add any noise to the observed data. There are many environments "in the wild" where there is no opportunity to add noise. For example, when browsing online, the NSA (presumably) knows our identities and tracks our browsing behavior. When shopping on Amazon, one must be logged in with credit card information and a mailing address. Thus Amazon can identify each shopper individually and records our exact purchases. Similarly at the grocery store checkout, the clerk can physically see each shopper and observes the shopper's purchases exactly. In these examples, there is no mechanism by which to add noise, because choices are directly observed.

A rational individual who knows that she is being watched might change her shopping or browsing behavior to obscure her true preferences. This is exactly the phenomenon we want to explore today: How do people change their behavior when they have concerns for privacy and know that they're being watched? Can an observer learn about agents' true preferences by observing choices in this setting? We will see that a rough answer to the latter question is "no." The observer is not able to reverse engineer the agents' preferences, without making very strong assumptions on the structural properties of preferences.

2 Model

We begin with a model of choices and observations in this privacy-aware setting.

Definition 1 (Choice Instance) *A choice instance $(X, \mathcal{A}, c)$ consists of:*
X*: finite set of alternatives*
$\mathcal{A}$*: collection of subsets of* X
$c : \mathcal{A} \to X$ *deterministic choice function s.t* $c(A) \in A$

As an example of a choice instance, today Alice is offered a menu of goods $A_1 = \{a, b\}$, and she picks $c(A_1) = a$. Tomorrow, she's offered a different (possibly overlapping) menu of goods $A_2 = \{a, b, c\}$, and picks $c(A_2) = b$. This example corresponds to the choice instance $(X, \mathcal{A}, c)$ where $X = \{a, b, c\}$, $\mathcal{A} = \{A_1, A_2\}$ and choice function c where $c(A_1) = a$ and $c(A_2) = b$.

In the standard economic setting (here called the *privacy-oblivious* setting), agents are not concerned with privacy. Instead they have preferences over objects in X, and select their most preferred object from menu A. An observer could safely infer that for all $y \in A \setminus \{c(A)\}$ that were available but not chosen, $c(A) \succ y$. For those of you with an economics background, you might notice that Alice's choices in the above example violate the Weak Axiom of Revealed Preference (if she is privacy-oblivious). That is, based upon her choices, we must conclude that Alice is irrational. From $c(A_1) = a$, we conclude that $a \succ b$, and from $c(A_2) = b$, we conclude that $b \succ a$. This is not possible for a rational agent.

However, if Alice is concerned with privacy, her choices can be rationally explained. Imagine that the alternatives in X are new cameras, where a is a disposable camera, b is a digital "point-and-shoot" camera, and c is a fancy SLR camera[1]. Imagine also that Alice doesn't like newfangled technology, so her true preferences over cameras are $a \succ b \succ c$, but she doesn't want her preferences to be known because she is afraid of being called a Luddite. Perhaps when faced with menu A_1, Alice chose her a because she truly prefers that camera. When option c was added in menu A_2, Alice found it would be too embarrassing to have the observer think that she preferred $a \succ c$, so she chose b because she didn't mind that the observer thinks she prefers $b \succ c$. With this story about Alice's privacy concerns, she appears to be rational. This suggests that a new model is required to describe the preferences of privacy-aware agents, to formally capture the intuition described above.

When an agent (Alice) makes a privacy-aware choice, she believes that the observer is going to make inferences about her preferences based upon the observed choice. The agent believes herself to be smarter than the observer, and incorporates the observers inferences when making choices. That is, she picks $c(A) \in A$ such that $\forall y \in A$,

$$(c(A),\ \text{info revealed by } c(A)) \succeq (y,\ \text{info revealed by } y)$$

We formalize the idea of "information revealed" as the inferences made by an observer who will naively make inferences as in the standard economic (privacy-oblivious) model. Let $\mathbb{B}(X) = 2^{X \times X}$ be the collection of all sets of pairwise comparisons of elements from X. Then $\mathbb{B}(X)$ is the set of all preferences that can be inferred by the observer. If we want to talk about some specific collection of binary preferences being inferred, use $B \in \mathbb{B}(X)$, and we'll think of B as a binary relation representing preferences.

$$(x, y) \in B \quad \iff \quad x \succ y$$

In the paper ([CEW14]), we generalize this method of inference to incorporate Alice's higher order beliefs about the observer's reasoning. For the purpose of this lecture, we restrict to the specific form of inferences described above. One point to emphasize here is that the method of inference described is only Alice's *beliefs* about the way in which the observer will make inferences. The observer could actually be as sophisticated as you or me, when making inferences about Alice.

Definition 2 (Privacy Preference) *A privacy preference is a linear order $\succsim$ over pairs of outcomes and privacy $X \times \mathbb{B}(X)$*

Let's go back to the example, and see exactly the form these privacy preferences will take when Alice chooses $b = c(A_2)$ from $A_2 = \{a, b, c\}$. Alice believes the observer will infer *privacy-oblivious* preferences:

$$b \succ a \text{ and } b \succ c$$

Alice's *privacy-aware* preferences must be:

$$(b, \{(b, a), (b, c)\}) \succ (a, \{(a, b), (a, c)\})$$

and

$$(b, \{(b, a), (b, c)\}) \succ (c, \{(c, a), (c, b)\})$$

3 Rationalizability

We want Alice to be rational, which means that her privacy preferences should be rational preferences. However, we don't get to see her true privacy preferences. Instead, we get to see her choices and must infer privacy preferences. We now define what it means for Alice to make rational choices. Informally, Alice's choices are rational when we can find rational privacy preferences that are optimized by the observed choices.

[1]Here we do not consider the effect of prices on choice, so imagine that all cameras are the same price.

Definition 3 (Rationalizable) *A choice instance* $(X, \mathcal{A}, c)$ *is rationalizable (via privacy preferences) if there is a privacy preference* $\succsim$ *such that* $\forall A \in \mathcal{A}$*, if* $x = c(A)$ *and* $y \in A \setminus \{x\}$*, then*

$$(x, \{(x, z) : z \in A \setminus \{x\}\}) \succ (y, \{(y, z) : z \in A \setminus \{y\}\})$$

or for shorthand,

$$(x, B(x, A)) \succ (y, B(y, A))$$

We say that $\succsim$ **rationalizes** $(X, \mathcal{A}, c)$.

This leads to the question: When can choice instances be rationalized via privacy preferences? One might imagine that the flexibility and higher dimensionality of privacy preferences would allow us to explain all (or nearly all) choices. For this reason, we impose natural assumptions on the structural properties of privacy preferences. During this lecture, we will consider privacy preferences that satisfy first monotonicity, then separability, and finally additive separability.

4 Monotonicity

Monotone privacy preferences are those in which the agent always prefers more privacy to less privacy when paired with a fixed object.

Definition 4 (Monotone Privacy Preferences) *A binary relation* $\succeq$ *over* $X \times \mathbb{B}(X)$ *is a monotone privacy preference if* $B \subsetneq B'$ *implies that* $(x, B) \succ (x, B')$*.*

When $B \subsetneq B'$, then under B' the observer makes the same inferences as under B, plus some additional inferences. Thus B corresponds to fewer inferences being made about the agent's preferences, and more privacy for the agent. We now ask, when can choices be rationalized with monotone privacy preferences? Proposition 5 says that all choices can be rationalized with monotone privacy preferences. This result could be interpreted in one of two ways. First, if the observer knows that the agent is rational, then monotonicity of her privacy preferences is not testable because he can never observe choices that will contradict monotonicity. Second, if the observer knows that the agent has monotone privacy preferences, then rationality is not testable because again, he will never see choices that are not rationalizable.

Note that with only assuming monotonicity, we can't ask about the observer inferring preferences over objects, because these preferences may not be well-defined. For example, it's possible that $(x, B) \succ (y, B)$ while $(y, B') \succ (x, B')$.

Proposition 5 *Any choice instance is rationalizable via monotone privacy preferences.*

Proof Fix a choice instance $(X, \mathcal{A}, c)$. We will construct a graph representing the inferred privacy preferences, and use the graph to find a monotone privacy preference that rationalizes $(X, \mathcal{A}, c)$.

$G = (V, E)$, directed
$V = X \times \mathbb{B}(X)$
E represents inferred preferences over pairs (x, B) from observed choices and monotonicity

By the following lemma, it is sufficient to show that this graph is acyclic.

Lemma 6 (Szpilrajn's Lemma) *If* $C \subseteq D$ *is acyclic, then there is a linear order* $\succeq$ *on* D *such that* $C \subseteq \succeq$*.*

Intuitively, Szpilrajn's Lemma says that as long as the edge-set is acyclic, it can be extended to form a complete preference ordering (think of C as the edge-set E, and D as $V \times V$). The promised linear order $\succeq$ will be a monotone privacy preference that rationalizes $(X, \mathcal{A}, c)$.

From the assumptions placed on the inferred privacy preferences, this graph can have only two types of edges: those coming from the requirement that preferences rationalize choices, and those coming from the assumed monotonicity of preferences.
Rationalizing: $[(x, B(x, A)), (x', B(x', A))] \in E$ if for some $A \in \mathcal{A}$, $x, x' \in A$ and $x = c(A)$
Monotone: $[(x, B), (x, B')] \in E$ for all $x \in X$ if $B \subset B'$

We proceed by proving some facts about this graph.

Fact 7 *Any cycle in this graph cannot contain monotone edges*

First, traversing rationalizing edges keeps the size of B fixed because $B(x, A)$ is the set of pairwise comparisons of x with elements in $A \setminus \{x\}$, and $B(x', A)$ is the set of pairwise comparisons of x' with elements in $A \setminus \{x'\}$ for the same set A. Traversing monotone edges increases the size of B. Thus the size of the B component is non-decreasing along all edges in this graph, and strictly increasing along monotone edges, so any cycle must contain only rationalizing edges.

Fact 8 *Graph cannot have two rationalizing edges in a row*

Assume not. Then the graph contains a path of the following.

$$(x_1, B(x_1, A)) \longrightarrow (x_2, B(x_2, A)) \longrightarrow (x_3, B(x_3, A))$$

Since these are both rationalizing edges, we can conclude that $\exists A$ such that $x_1, x_2, x_3 \in A$, and both $x_1 = c(A)$ and $x_2 = c(A)$. If $x_1 \neq x_2$, then this violates uniqueness of choice. Then it must be that $x_1 = x_2$. However, then $B(x_1, A) = B(x_2, A)$ and $(x_1, B(x_1, A)) \succ (x_2, B(x_2, A)) = (x_1, B(x_1, A))$ which is a contradiction because a pair cannot be strictly preferred to itself.

With these two facts, we are done and the graph cannot contain a cycle. Thus all choices can be rationalized with monotone privacy preferences. ■

This is a happy result for our friend Alice, but is a bad result for the observer because, even with this restriction on preferences, the observer can't tell the difference between rational and irrational agents! Perhaps if the observer hopes for testable implications, he should consider stronger assumptions on the form of privacy preferences. We now look at separable privacy preferences. This structural assumption is nice for the observer because it ensures that preferences over objects are well-defined.

5 Separability

Separable privacy preferences require that the agent has a preference ordering over X that must be respected by privacy preferences. This is precisely the ordering that observer hopes to learn.

Definition 9 (Separable Privacy Preferences) *A binary relation $\succeq$ over $X \times \mathbb{B}(X)$ is a separable privacy preference if it is monotone and additionally satisfies that for all $x, y \in X$ and $B \in \mathbb{B}(X)$,*

$$(x, B) \succeq (y, B) \Longrightarrow (x, B') \succeq (y, B') \; \forall B' \in \mathbb{B}(X)$$

Here we can ask two questions: First, when can choices be rationalized with separable privacy preferences? Second, now that the underlying preferences over X are well-defined, can the observer infer these preferences from choices? Proposition 10 says that the observer's prospects are just as bleak as with monotonicity alone, despite the stronger assumption.

Proposition 10 *Let $(X, \mathcal{A}, c)$ be a choice instance, and let $\succeq$ be any linear order over X. Then there is a separable privacy preference $\succeq^*$ such that*

1. $\succsim^*$ *rationalizes* $(X, \mathcal{A}, c)$, *and*

2. *the projection of* $\succeq^*$ *onto* X *is well defined and coincides with* $\succeq$.

The first part of this result implies that all choice behavior can be rationalized with separable privacy preferences. The second part of the proposition says that observer is not able to learn the agent's preference ordering over objects. Think of $\succeq$ as the observer's conjecture about the agent's preferences over X. Part (2) says that all conjectured preferences over X are consistent with any choice behavior. That is, for any guess about preferences over X and for any observed choices, the choices will never contradict the guess.

Proof Fix a choice instance $(X, \mathcal{A}, c)$ and a linear order $\succeq$ over X. We will again construct a graph representing inferred privacy preferences and show that the graph is acyclic. By Szpilrajn's Lemma, this is sufficient to complete the proof.

$G = (V, E)$, directed
$V = X \times \mathbb{B}(X)$
$E =$ represents inferred preferences over pairs (x, B) from observed choices, monotonicity, and separability

In this graph, there are three types of edges:
Rationalizing: $[(x, B(x, A)), (x', B(x', A)] \in E$ if for some A, $x, x' \in A$, and $x = c(A)$
Monotone: $[(x, B), (x, B')] \in E$ for all $x \in X$ if $B \subset B'$
Separable: $[(x, B), (x', B)] \in E$ for all $B \in \mathbb{B}(X)$ if $x \succ x'$ according to the fixed order $\succeq$ over X.

As before, we proceed by proving facts about this graph.

Fact 11 *Any cycle in the graph cannot contain monotone edges*

We saw before that traversing a monotone edge strictly increases the size of B, and traversing a rationalizing edge keeps the size of B fixed. Separable edges also keep the size of B fixed because they do not change the set B itself. By the same reasoning as for Fact 7 in the proof of Proposition 5, no cycle can contain monotone edges.

Fact 12 *The graph cannot have two rationalizing edges in a row*

Same as Fact 8 in the proof of Proposition 5.

Fact 13 *The cycle does not have two separable edges in a row (WLOG)*

Consider a cycle that does have two separable edges in a row:

$$(x_1, B) \longrightarrow (x_2, B) \longrightarrow (x_3, B)$$

Then we know that $x_1 \succ x_2$ and $x_2 \succ x_3$. By transitivity, $x_1 \succ x_3$ and these two edges could be replaced by a single transitive edge from (x_1, B) to (x_3, B). If there exists a cycle with two separable edges in a row, then there also exists a cycle without these two edges, which we could consider instead. (Alternatively, we could say that we are looking for the smallest cycle. The smallest cycle would not contain two separable edges in a row, by the same reasoning.) Without loss of generality, we can consider only cycles without two sequential separable edges. The only time this is *not* without loss of generality is when the entire cycle consists of separable edges. However, we have assumed that $\succeq$ is a linear order, so that cannot happen.

From the previous three facts, it must be that cycle has alternating rationalizing and separable edges.

Fact 14 *Graph cannot have a path of rationalizing, separable, rationalizing edges*

Assume not. Then the graph contains a path of the form, where the first and last edges are due to rationalizing choices:

$$(x_1, B(x_1, A)) \longrightarrow (x_2, B(x_2, A)) \longrightarrow (x_3, B(x_3, A')) \longrightarrow (x_4, B(x_4, A'))$$

The middle edge is due to separability, which requires $B(x_2, A) = B(x_3, A')$. This implies $x_2 = x_3$, so $(x_2, B(x_2, A)) \succ (x_3, B(x_3, A')) = (x_2, B(x_2, A))$ which is a contradiction since a pair cannot be strictly preferred to itself.

Fact 15 *Cycle cannot have size 2*

If it did, then traversing the cycle twice would give a path of rationalizing, separable, rationalizing edges, which we have just shown is not possible.

■

This means that rationality is not testable even assuming separable privacy preferences. Further, it means that for all linear orders over X, all choices can be rationalized with separable privacy preferences that coincide with the ordering. If the observer hopes to learn anything at all from Alice's choices, he should make even stronger assumptions about the form of her privacy preferences.

6 Additive separability

We now move to consider a specific functional form of preferences, where the utility an agent enjoys from a choice is her utility for the object minus her loss from the inferences made by the observer.

Definition 16 (Additive Privacy Preference) *A binary relation $\succeq$ over $X \times \mathbb{B}(X)$ is an additive privacy preference if there are functions $u : X \to \mathbb{R}_+$ and $v : X \times X \to \mathbb{R}_+$ such that $(x, B) \succ (x', B')$ iff*

$$u(x) - \sum_{(z,z') \in B} v(z, z') > u(x') - \sum_{(z,z') \in B'} v(z, z')$$

Think of the function u as the agent's gain from the object and v as her loss from having each pairwise preference revealed. To rationalize choices using additive privacy preferences, we must find functions u and v that are optimized by the observed choices. Once choices are rationalized, the observer can also learn about preferences over X, as they are specified by u.

Proposition 17 *There exist choice instances which are not rationalizable with additive privacy preferences.*

Proof We prove the Proposition with the following example of choices that cannot be rationalized with additive privacy preferences. Define the set of alternatives $X = \{x, y, z, w\}$ and the choice data as follows:
$z = c(\{x, z\})$, $x = c(\{x, y, z\})$, $x = c(\{x, w\})$, $w = c(\{x, y, w\})$, $w = c(\{w, z\})$, $z = c(\{w, y, z\})$.

Assume (towards a contradiction) there exists a pair (u, v) that rationalizes this choice instance. From $z = c(\{x, z\})$, we see

$$u(z) - v(z, x) > u(x) - v(x, z)$$

and from $x = c(\{x, y, z\})$ we see

$$u(z) - v(z, x) - v(z, y) < u(x) - v(x, z) - v(x, y)$$

Therefore $v(z, y) > v(x, y)$.

From the other pairs of observations, $x = c(\{x, w\})$ and $w = c(\{x, y, w\})$ imply that $v(x, y) > v(w, y)$, and $w = c(\{w, z\})$ and $z = c(\{w, y, z\})$ imply that $v(w, y) > v(z, y)$. Combining the inequalities on v, we see a contraction. ■

We finally see testable implications of rationality with additive privacy preferences. We would now like to develop a test for rationality in this setting. The previous example suggests a necessary condition for rationalizability.

Necessary condition

Definition 18 *For $y \in X$, define binary relation R^y by: $x\ R^y\ z$ if there is $A \in \mathcal{A}$ with $z = c(A)$ and $x = c(A \cup \{y\})$.*

That is, we say $x\ R^y\ z$ if z was originally chosen from some set A, but with the addition of y, the agent picks x instead. Since x was always available, the change in choice must have come from privacy concerns. This suggests that the agent doesn't want the observer to infer $z \succ y$ (or that it's not too embarrassing if the observer infers $x \succ y$). In the previous example, it was the case that

$$zR^y xR^y wR^y z$$

Proposition 19 *A choice instance can be rationalized only if R^y is acyclic for all $y \in X$.*

Proof Assume towards a contradiction that for some $y \in X$, R^y contains a cycle.

$$a_1\ R^y\ a_2\ R^y \cdots R^y\ a_k\ R^y\ a_1$$

We saw from the previous example (and the definition of R^y) that

$$x\ R^y\ z \Longrightarrow v(z, y) > v(x, y)$$

Then on this cycle, the following inequalities must hold

$$v(a_1, y) > v(a_2, y) > \cdots > v(a_k, y) > v(a_1, y)$$

This gives a contradiction, so acyclicity of R^y for all $y \in X$ is necessary for choices to be rationalized with additive preferences. ■

Necessary and sufficient condition

While Proposition 19 gives a necessary condition for rationalizability, it is not sufficient. To provide a complete test for rationalizability, we need a necessary and sufficient condition. We now construct such a condition using the linear inequalities that can be inferred from each choice observation.

For notational ease, rename the elements of X to be $\{x_1, \ldots, x_n\}$. When the agent picks $x_i = c(A) \in A$ instead of available alternative $x_j \in A$, the following inequality must hold.

$$u(x_i) - \sum_{z \in A \setminus \{x_i\}} v(x_i, z) > u(x_j) - \sum_{z \in A \setminus \{x_j\}} v(x_j, z)$$

Rearranging terms gives,

$$u(x_i) - u(x_j) + \sum_{z \in A \setminus \{x_j\}} v(x_j, z) - \sum_{z \in A \setminus \{x_i\}} v(x_i, z) > 0$$

The left hand side of this inequality can be written as a dot product of two vectors, where the first vector $\mathbf{u}$ contains the values of $u(\cdot)$ and $v(\cdot,\cdot)$ evaluated on all elements of X. (The first n entries of $\mathbf{u}$ contain the values of $u(x_1),\ldots,u(x_n)$, and the remaining n^2-n entries will contain the values $v(x_i,x_j)$ for $i \neq j$.) This vector can be dotted with another vector of length n^2 containing the appropriate entries from $\{0,1,-1\}$. The inequality requires that the dot product of these two vectors be strictly positive.

We can record all implied inequalities by storing each vector of 0's, 1's, and -1's as a row in a matrix T that has one row for each $x_i = c(A)$ and $x_j \in A \setminus \{c(A)\}$ for each $A \in \mathcal{A}$. We can now represent all inequalities as a matrix inequality:

$$T\mathbf{u} > \mathbf{0} \tag{1}$$

Proposition 20 *A choice instance can be rationalized with additive privacy preferences if and only if there exists a vector* $\mathbf{u}$ *satisfying Equation* (1).

Proof If such a vector $\mathbf{u}$ exists that satisfies this inequality, then there exist functions $u : X \to \mathbb{R}^+$ and $v : X \times X \to \mathbb{R}^+$ such that additive privacy preferences are optimized by the observed choices, and observed choices are rationalizable. Otherwise, choices are not rationalizable. ■

One issue with this condition is that checking it may not be computationally feasible, as the matrix T has super-exponential size (in the worst case, T has $O(n^2 2^n)$ rows). This suggests that checking only the necessary condition (acyclicity of R^y) may serve as a useful approximation if the observer is concerned with run time.

Remark There are two important conclusions to draw from this work. First we see that in the context of revealed preference, agents behave very differently when they have privacy concerns. Privacy-aware agents make choices to optimize different preferences, and they may not always have well defined preferences over objects. Second, we see that there is a difference in what the observer is able to learn from the observed choices. When agents are privacy-oblivious, rationality is testable and the observer can always learn the agent's preferences over objects simply from observing choices. However, in the presence of privacy concerns, the observer can't test for rationality and can't learn anything about the agent's preferences over objects without placing extremely strong restrictions on the form of privacy preferences. This might explain why organizations (such as Google or the NSA) don't like it when their surveillance of our behavior is announced, since this may change us from being privacy-oblivious to privacy-aware. The fact that we know we're being observed and can take action to mask our true preferences may successfully prevent them from learning about us.

Bibliographic Information The contents of this lecture are taken entirely from Cummings, Echenique, Wierman, "The Empirical Implications of Privacy-Aware Choice" [CEW14].

References

[CEW14] Rachel Cummings, Federico Echenique, and Adam Wierman. The empirical implications of privacy-aware choice. *arXiv pre-print arXiv:1401.0336*, 2014.