Abstract:

In the field of mechanism design, one aims to design mechanisms that on one hand are strategyproof, and on the other hand guarantee certain desiderata of the outcome, such as maximization of welfare, elimination of envy, etc. Despite strong impossibility results, four grand "mechanism design success stories" stand out: In two-sided markets, the Gale-Shapley algorithm provides a stable mechanism that is strategyproof for the proposing side; in markets without transfers, on the domains of single-peaked preferences and of object assignment, there are appealing Pareto-optimal and strategyproof mechanisms such as median voting for the former and top trading cycles for the latter; finally, in markets with transfers, on the domain of quasilinear preferences, the strategyproof VCG (Vickrey-Clarke-Groves) mechanism provides a very general framework for social-welfare maximization.

In this talk, we investigate each of these four success stories, and ask whether one can design appealing mechanisms for the respective desiderata that are not only strategyproof, but also obviously strategyproof, a stronger incentive property that was recently introduced by Li (2015) and has since garnered considerable attention. For stable mechanisms and for quasilinear preferences, we obtain strong negative results (with positive results obtained for very special cases, by us for the former and by Li for the latter). Nonetheless, for single-peaked preferences and for object assignment we show that the class of Pareto optimal social-choice functions that are implementable via obviously strategyproof mechanisms is considerably richer than merely dictatorships, and we fully characterize this class for each of these domains. Our first step toward all of these characterizations is the development of a general "gradual revelation principle" for obviously strategyproof mechanisms, an analog of the (direct) revelation principle for strategyproof mechanisms, which we believe to be of independent interest.

An integrated examination, of all of these negative and positive results, on the one hand reveals that the various mechanics that come into play within obviously strategyproof mechanisms are considerably richer and more diverse than previously demonstrated and can give rise to rather exotic and quite intricate mechanisms in some domains, however on the other hand suggests that the boundaries of obvious strategyproofness are significantly less far-reaching than one may hope in other domains. We thus observe that in a natural sense, obvious strategyproofness is neither "too strong" nor "too weak" a definition for capturing "strategyproofness that is easy to see," but in fact while it performs as intuitively expected on some domains, it "overshoots" on some other domains, and "undershoots" on yet other domains.

Based upon joint work with Itai Ashlagi (2015) and joint work with Sophie Bade (2016).