

Preference aggregation and game theory

Fioravante PATRONE

DIPTTEM, University of Genoa

Paris, 23 June 2010

Workshop on *Game Practice and Preference aggregation*

Summary

Non-strategic point of view

Arrow

Harsanyi and utilitarianism

Interpersonal comparison of utilities

Strategic aspects and game theory

Strategic voting, and G-S

“Small worlds”

Computational issues

A broader picture

Great confusion under the sky: excellent situation!

To put the cart before the horses

Non welfarist approaches

Arrow's impossibility result

Well known: there is no rule that maps \mathcal{P}^n into \mathcal{P} , satisfying Pareto, IIA, non-dictatorship.

Where A is a finite set containing at least 3 elements and \mathcal{P} is the set of all total preorders on A . Implicitly assumed “universality”.

Nice applications (quoted in Roemer '96):

- MacKay (1980), to pentathlon
- May (1954), to the “construction” of personal preferences on multi-attribute alternatives

Positive results:

- Black (1958), for single-peaked preferences (needed an additional dimension)
- May (1952): characterization of simple majority voting over 2 alternatives only

Small detour: “Rawls”

A rule simple to describe. Rank alternatives. Choose the alternative with the best worst individual ranking.

Rawlsian flavor (maxmin).

What violates? Obviously IIA, too close to “Borda”...

Example to see that violates IIA. Cells contain the rankings of alternatives (higher is better). Left: alternative a_2 beats a_1 . Right: the converse is true.

$N \setminus A$	a_1	a_2	a_3
1	3	2	1
2	1	2	3
min	1	2	1

$N \setminus A$	a_1	a_2	a_3
1	2	1	3
2	2	3	1
min	2	1	1

Harsanyi and vNM utilities

Consider n individuals, with vNM preferences on $\Delta(A)$.

Assume that “society” also has vNM preferences satisfying “strong Pareto” and “Independent Prospects”.

Then (Harsanyi 1955), $u_s = \sum_{i=1}^n a_i u_i$, where the positive a_i 's are determined up to a positive constant.

Criticism by Diamond (1967): “independence” property questionable for social preferences:

$$1/2(c, 0) + 1/2(0, c) \sim 1/2(c, 0) + 1/2(c, 0) \sim (c, 0)$$

Utilitarianism? No, see Roemer's *Theories of Distributive Justice*, 1996.

Critical issue for distributive justice

Main actor: Bentham *An Introduction to the Principles of Morals and Legislation*, 1789.

Maximize the sum of utility across individuals.

But (Sidgwick, 1874):

Assuming, then, that the average happiness of human beings is a positive quantity, it seems clear that, supposing the average happiness enjoyed remains undiminished, Utilitarianism directs us to make the number enjoying it as great as possible. But if we foresee as possible that an increase in numbers will be accompanied by a decrease in average happiness or vice versa, a point arises which has not only never been formally noticed, but which seems to have been substantially overlooked by many Utilitarians

Not to say of the difficulties in finding a common scale for measurement of “utility” (felicific, or hedonic calculus?).

Manipulation in voting

It is well known that Borda rule is manipulable.

Example: Graziano Battistini receives the order to stop and wait for Gastone Nencini to allow him to get the time bonus: Tour de France 1960, tenth stage, Mont-de-Marsan - Pau, won by Roger Rivière...

Or the 2007 Brazil GP of Formula 1: thanks to the second pit-stop, Räikkönen can pass Massa and win the championship.

<http://www.matematicamente.it/forum/l-olanda-deve-farsi-battere-dalla-romania-t30433.html>

European soccer championship, 2008: The Netherlands should play to be defeated by Romania

Interesting contribution: Dummett and Farquharson (1961).
Forerunners of Gibbard/Satterthwaite (and explicitly quoted by Gibbard).

Difficulties in implementation

Classical result: Gibbard-Satterthwaite theorem.

Given a social choice rule $f : \mathcal{P}^n \rightrightarrows A$ satisfying CS (or “onto”):
If f is implementable in weak dominant strategies via a strategic game form, then f is dictatorial.

So, for voting (Corollary in Gibbard 1973):

Every voting scheme with at least 3 outcomes is either dictatorial or manipulable.

Of course simple majority is “ok”, *for two outcomes* (May, 52, quoted already)

Vickrey, Clarke and Groves

Vickrey auction (second price sealed bid): well known case in which truthful bidding is weakly dominating.

But if externalities... Example by Jeffrey Ely (in column the utilities attached to alternatives by participants: bidder 2 does not like that a_3 is chosen):

$N \setminus A$	a_1	a_2	a_3
1	1	0	0
2	0	2	-5
3	0	0	3

Truthful bidding is not anymore dominant. So? Well, use Vickrey-Clarke-Groves. It takes into account externalities, that enter into the definition of transfers.

Difficulties from the fact that balance of transfers is not guaranteed.

Numbers, not letters

Rothkopf: *Thirteen Reasons Why the Vickrey-Clarke-Groves Process Is Not Practical.*

Among the 13, the NP completeness of the winner determination problem. But also 1 issue related with exponential effort required by bid preparation and bid communication (and 1 related issue: including bid preparation costs make disappear dominant strategy equilibria).

On the other hand, computational complexity can be useful to prevent manipulation: see the seminal paper by Bartholdi, Tovey *et al.*¹ (rejected by mainstream economic journals, “of course”?). It is worst case analysis. Possibly “most” instances are easy to manipulate (relevant in *prctice*).

¹<http://mat.tepper.cmu.edu/blog/?p=1117>

Less structured mechanisms

Few examples:

- Wikipedia: who is designer, which are the goals? It works. Up to where? The “mechanism” has presently obvious faults:

- no expertise required (good popularization requires strong knowledge)

- too many hands on one article: difficult to improve quality, missing an organicistic approach (to pages, to sectors, to all of the wikipedia)

It will work? Up to where?

Evolution of conventions (Young), of norms. The contractarian approach.

Cooter: “Three Effects of Social Norms on Law: Expression, Deterrence, and Internalization”

Which came first? The chicken or the egg?

Evolution of cooperation (Axelrod and Hamilton): birth of “cooperators” (hard-wired) in a world of defectors. How it can happen? Positive probability of repeated interaction.

- Robustness (famous experiment → TFT)
- Stability (ESS)
- Initial viability (most critical issue, A&H offer few possible paths)

We are the outcome of a long process (evolution, culture): our “preferences” are the temporary outcome of this process. Remind also Cooter (previous slide).

Which came first? The chicken or the egg?

Don't forget that the “teleological” point of view works (if it works) at the appropriate level... Difficult to assign a goal to deoxyribonucleic acid (even if A&H say: *A gene [...] looks beyond its mortal bearer*).

And don't forget that “preferences-goal” is a useful paradigm, allows to work at a very high level of synthesis (like the macro approach in thermodynamics).

Other approaches, like “neuroeconomics”, can work at a more basic level, more similar to hard sciences. The “lie detector” can bypass the big problem of a mechanism designer - not to know preferences of the players (maybe there is some constitutional hurdle to be overcome).

Kidney exchange

Kidney exchange.

Where are preferences? Criteria based on “objective” evaluations: 0-1 compatibility (US point of view), but also via tissue types (grading: European point of view). More debatable is the evaluation of an exchange (sum or \wedge ?)

Analyzed the problem of manipulability of mechanisms used, based also on grading (Villa and Patrone, 2009): the fact that a patient may have more donors opens the door to manipulability (hiding donors may be convenient, to induce a more favorable cycle of exchanges).

Contrary to Roth, Sönmez, and Ünver (2005), who prove that revealing all the set of available donors is a dominant strategy for an efficient mechanism (either deterministic or stochastic).

Where comes from your utility, sir?

Yaari and Bar-Hillel: *On dividing justly*, 1984.

Analysis of some mechanism (among which Bargaining, Nash and Kalai-Smorodinsky, with various disagreement points) that work in the utility space.

Results (of the experiment [*practice?*]) show that suggested divisions depend on whether utilities represent *needs*, *tastes* or *beliefs*. And even a couple of cases, both dealing with needs, get different answers. But all of these cases have *the same set of possible agreements*, in utility space.

Rawls, Dworkin, Sen...

Simply I cannot conclude without mentioning the contributions of Rawls (primary “goods”), Dworkin (equalization of resources), Sen (capabilities, functionings).

Getting (among other things) out of a purely welfarist approach.

Nice book by Roemer, already mentioned.

References:

http://www.diptem.unige.it/patrone/Parigi_23_giu_2010_FP_references.pdf