Welfare Engineering in Multiagent Systems

Ulle Endriss[1] and Nicolas Maudet[2]

[1] Department of Computing, Imperial College London
Email: ue@doc.ic.ac.uk

[2] LAMSADE, Université Paris-Dauphine
Email: maudet@lamsade.dauphine.fr

Talk Overview

- Resource allocation by negotiation in multiagent systems
  definition of our basic negotiation framework
- Behaviour profiles of individual agents
  how do agents decide whether or not to accept a deal?
- Measuring social welfare
  what are optimal outcomes from the viewpoint of society?
- Welfare engineering
  how can we make agents negotiate socially optimal outcomes?
- Results for and discussion of concrete notions of social welfare
  utilitarianism, egalitarianism, Lorenz optimality, ...
- Conclusion

Resource Allocation by Negotiation

- Finite set of *agents* $\mathcal{A}$ and finite set of *resources* $\mathcal{R}$.
- An *allocation* A is a partitioning of $\mathcal{R}$ amongst the agents in $\mathcal{A}$.
  Example: $A(i) = \{r_3, r_7\}$ — agent i owns resources r_3 and r_7
- Every agent $i \in \mathcal{A}$ has a *utility function* $u_i : 2^{\mathcal{R}} \to \mathbb{R}$.
  Example: $u_i(A) = u_i(A(i)) = 577.8$ — agent i is pretty happy
- Agents may engage in negotiation to exchange resources in order to benefit either themselves or society as a whole.
- A *deal* $\delta = (A, A')$ is a pair of allocations (before/after).
  An agent may or may not find a particular deal *acceptable*.

Possible Agent Behaviour Profiles

An agent i may or may not accept a particular deal $\delta = (A, A')$. Here are some examples for possible *acceptability criteria:*

rational (selfish) agent	$u_i(A) < u_i(A')$
rational but cooperative agent	$u_i(A) \leq u_i(A')$
rational and demanding agent	$u_i(A) + 10 < u_i(A')$
masochist	$u_i(A) > u_i(A')$

disciple of agent *guru*	$u_{guru}(A) < u_{guru}(A')$
team worker (for team T)	$\sum_{j \in T} u_j(A) < \sum_{j \in T} u_j(A')$

Example for a Protocol Restriction

no more than two agents to be involved in any one deal	$\lvert\mathcal{A}^\delta\rvert \leq 2$ where $\mathcal{A}^\delta = \{i \in \mathcal{A} \mid A(i) \neq A'(i)\}$

Social Welfare

A *social welfare ordering* formalises the notion of a society's "preferences" given the preferences of its members (the agents).

▶ The *utilitarian* social welfare $sw_u(A)$ of an allocation of resources A is defined as follows:

$$sw_u(A) \quad = \quad \sum_{i \in \mathcal{A}} u_i(A)$$

That is, anything that increases average (and thereby overall) utility is taken to be socially beneficial.

▶ Under the *egalitarian* point of view, on the other hand, social welfare is tied to the welfare of a society's weakest member:

$$sw_e(A) \quad = \quad min\{u_i(A) \mid i \in \mathcal{A}\}$$

Utilitarianism versus Egalitarianism

- In the multiagent systems literature the utilitarian viewpoint (i.e. social welfare = sum of individual utilities) is usually taken for granted.
- In philosophy/sociology/economics not.
- John Rawls' *"veil of ignorance"* (*A Theory of Justice*, 1971):

> *Without knowing what your position in society (class, race, sex, ...) will be, what kind of society would you choose to live in?*

- Reformulating the *veil of ignorance* for multiagent systems:

> *If you were to send a software agent into an artificial society to negotiate on your behalf, what would you consider acceptable principles for that society to operate by?*

- *Conclusion:* worthwhile to investigate egalitarian (and other) social principles also in the context of multiagent systems.

Welfare Engineering

- Different applications induce different measures of social welfare for artificial societies:
  - "pure" *e*-commerce $\longrightarrow$ utilitarian
  - sharing of jointly owned resources $\longrightarrow$ egalitarian
  - ...
- Given some social welfare ordering, we want to "engineer" appropriate (local) behaviour profiles for individual agents to ensure convergence towards a (globally) optimal state.

Utilitarian and Egalitarian Systems

Previous results (Sandholm 1998, E. *et al.* 2003):

- *Cooperative rationality* (no agent accepts a loss; one agent requires a profit) is an appropriate behaviour profile in societies where *Pareto optimal* allocations are desirable.
- *Individual rationality* (every agents requires a profit—after compensatory payments) is an appropriate behaviour profile in societies where maximising *utilitarian* social welfare is desired.
- *Equitability* (local improvement of minimal utility) is an appropriate behaviour profile in *egalitarian* agent societies.

Our "sufficiency theorems" typically have the following form:

> *Any sequence of deals conforming to behaviour profile X will eventually result in an allocation of resources that is optimal according to the social welfare ordering Y.*

Necessity of Complex Deals

In general, very complex deals (involving any number of resources or agents) may be *necessary* to guarantee optimal outcomes (given the agent behaviour profiles from before).

Improved Results for Restricted Domains

For example (E. *et al.* 2003):

- Cooperatively rational *one-resource-at-a-time* deals suffice to guarantee maximal utilitarian welfare in *0-1 scenarios* (single resources have utility 0 or 1 and utility functions are additive).

Note that we have no such results for egalitarian agent societies.

Lorenz Optimality

We are now going to look at a compromise between the utilitarian and the egalitarian definitions of social welfare ...

Technical Preliminaries

Every allocation A gives rise to an *ordered utility vector* $\vec{u}(A)$: compute $u_i(A)$ for all $i \in \mathcal{A}$ and present results in increasing order.

Example: $\vec{u}(A) = \langle 0, 5, 20 \rangle$ means that the weakest agent enjoys utility 0, the strongest utility 20, and the middle one utility 5.

Lorenz Optimal Allocations of Resources

Let A and A' be allocations of resources for a society with n agents. Then A is *Lorenz dominated* by A' iff we have

$$\sum_{i=1}^{k} \vec{u}_i(A) \quad \leq \quad \sum_{i=1}^{k} \vec{u}_i(A')$$

for all $k \in \{1..n\}$ and that inequality is strict in at least one case.

Discussion:

- Note that for $k = 1$ that sum is equivalent to the *egalitarian* and for $k = n$ to the *utilitarian* social welfare.
- What kind of local behaviour profile would guarantee Lorenz optimal negotiation outcomes?

Negotiating Lorenz Optimal Allocations

We can prove a new sufficiency theorem:

- In *0-1 scenarios*, any sequence of *simple Pareto-Pigou-Dalton deals* will eventually result in a *Lorenz optimal* outcome.

The class of "simple Pareto-Pigou-Dalton deals" has the following features (see paper for details):

- Any deal involves only *two agents* and *one resource.*
- Any deal is either inequality-reducing but mean-preserving (so-called *Pigou-Dalton transfer*) or *cooperatively rational.*

Note that seemingly more general results from the economics literature do not apply to our discrete negotiation spaces.

Elitist Agent Societies

We may define the *elitist* social welfare $sw_{el}(A)$ of an allocation of resources A as follows:

$$sw_{el}(A) \quad = \quad max\{u_i(A) \mid i \in \mathcal{A}\}$$

Discussion:

- Appropriate if it is in the system designer's interest that at least one agent succeeds (whatever happens to the rest).
- Technically similar to the egalitarian case.

Reducing Envy

An allocation of resources A is called *envy-free* iff the following holds for all pairs of agents $i, j \in \mathcal{A}$:

$$u_i(A(i)) \quad \geq \quad u_i(A(j))$$

Discussion:

- Envy-freeness would be desirable where self-interested agents are expected to collaborate over longer periods of time.
- Note that envy-free allocations do not always exist.
- Still, we could rate social welfare in terms of the number of agents without envy (or the overall "degree" of envy).
- However, it is not possible to define a *local* acceptability criterion that ensures envy reduction, because a deal could always affect the envy of agents not involved in it.

Conclusion

- We have argued that a whole *range of social welfare orderings* are relevant to multiagent systems (not just utilitarian/Pareto).
- We have put forward *welfare engineering* as the process of finding agent behaviour profiles that ensure socially optimal negotiation outcomes for a given social welfare ordering.
- We have put previous results for utilitarian and egalitarian agent societies into the context of this general perspective.
- We have proved a new result for artificial societies where *Lorenz optimal outcomes* are desirable.
- We have also discussed *elitist agent societies* and the idea of *reducing envy* in an agent society.