Social Choice / Making Group Decisions

Michal Jakob and Michal Pěchouček

Agent Technology Center, Dept. of Computer Science and Engineering, FEE, Czech Technical University

AE4M36MAS Autumn 2012 - Lect. 9

OTEVŘENÁ INFORMATIKA

Based partially on slides of Kevin L. Brown and Michael Rovatsos

Motivation

- Sorting product reviews
- Ranking commenters
- Determining joint trip destinations
- => aggregating individual preferences

Outline

- Formalization
- Basic voting schemes
- Paradoxes
- Desirable properties
- Arrow’s theorem

Social Choice

- **Social Choice Theory** is concerned with group decision making (basically analysis of mechanisms for voting)
- Out settings
  - a set of **agents**
  - a set of **outcomes**
  - agents have **preferences** across them
- Task is either to derive a globally acceptable preference ordering, or determine a winner

Formal Setting

- $N = \{1,2,\dots,n\}$ is a set of **agents** (voters)
- $O = \{o_1, o_2, \dots, o_m\}$ is a set of **outcomes** (alternatives / candidates)
- L is a set of all total orderings over O
- each agent i has **preferences** over O: a (non strict) total ordering $\succcurlyeq_i \in L$ over the set O s.t. $o_k \succcurlyeq_i o_l$ if agent i weakly prefers o_k to o_l
- Tuple $[\succcurlyeq] \in \mathrm{L}^n$ is a **preference profile**

Social Choice Function

Definition (Social choice function)

Assume a set of agents $N = \{1, 2, \ldots, n\}$, and a set of outcomes (or alternatives, or candidates) O. Let $L_{\text{-}}$ be the set of non-strict total orders on O. A social choice function (over N and O) is a function $C : L_{\text{-}}^{n} \mapsto O$.

- A social choice function takes the voter preferences and selects *one* outcome
- Example: presidential election.

Social Welfare Function

Definition (Social welfare function)

Let N, O, L_{-} be as above. A social welfare function (over N and O) is a function $W : L_{-}^{n} \mapsto L_{-}$.

- A social welfare function takes the voter preferences and produces a social preference order
- Example: beauty contest.

Non-Ranking Voting Schemes

- Agents give votes to one or more candidates
- **Plurality**
  - pick the outcome which is preferred by the most people
- **Cumulative voting**
  - distribute e.g., 5 votes each
  - possible to vote for the same outcome multiple times
- **Approval voting**
  - accept as many outcomes as you "like"

Anomalies with Plurality

- Outcomes A, B, C and 100 agents with following preferences
  - 40% agents prefer A
  - 30% agents prefer B
  - 30% agents prefer C
- With plurality, A gets elected even though a *clear majority* (60%) prefer another candidate!

Ranking Voting Schemes

- Non-ranking schemes do not take into account agent’s full preference structure / ignore voter’s preference orders
- **Plurality with elimination ("instant runoff")**
  - everyone selects their favorite outcome
  - the outcome with the fewest votes is eliminated
  - repeat until one outcome remains
- **Borda**
- **Pairwise elimination**

Borda

1. assign each outcome a number
2. the most preferred outcome gets a score of $n-1$, the next most preferred gets $n-2$, down to the n^{th} outcome which gets 0
3. then sum the numbers for each outcome, and choose the one that has the highest score

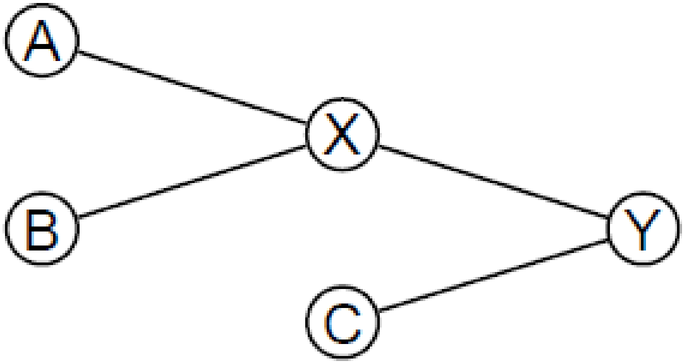
Borda Example

Politics example:

- 43 of $|Ag|$ are left-wing voters: $\omega_L \succ \omega_D \succ \omega_C$
- 12 of $|Ag|$ are centre-left voters: $\omega_D \succ \omega_L \succ \omega_C$
- 45 of $|Ag|$ are right-wing voters: $\omega_C \succ \omega_D \succ \omega_L$

Pairwise Elimination

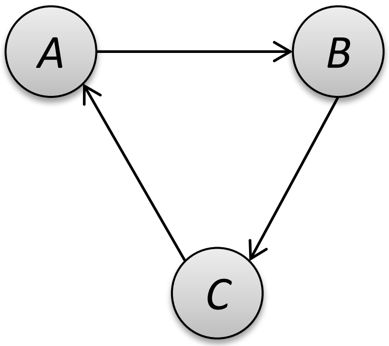
1. in advance, decide a schedule for the order in which pairs will be compared.
2. given two outcomes, have everyone determine the one that they prefer eliminate the outcome that was not preferred, and continue with the schedule



Majority Graphs

- This idea is easiest to illustrate by using a *majority graph*.
- A directed graph with:
  - vertices = candidates
  - an edge (*i, j*) if *i* would beat *j* in a simple majority election.
- A *compact representation of voter preferences*.

agent 1: $A \succ B \succ C$
agent 2: $C \succ A \succ B$
agent 3: $B \succ C \succ A$



Condorcet's Paradox

agent 1: $A \succ B \succ C$
agent 2: $C \succ A \succ B$
agent 3: $B \succ C \succ A$

- There are scenarios in which no matter which outcome we choose the majority of voters will be unhappy with the outcome chosen
- For every possible candidate, there is another candidate that is preferred by a $\frac{2}{3}$ majority of voters!

Condorcet condition

- **Condorcet winner**

  An outcome $o \in O$ *is a* Condorcet winner *if*

  $$\forall o' \in O, \#(o \succ o') \geq \#(o' \succ o)$$

- Condorcet winner does not always exist
  - Sometimes, there's a cycle where A defeats B, B defeats C, and C defeats A in their pairwise runoffs
- **Smith set**

  The Smith set *is the smallest set* $S \subseteq O$ *having the property that*

  $$\forall o' \notin S, \#(o \succ o') \geq \#(o' \succ o).$$

- Smith set always exists

Condorcet example

499 agents: $A \succ B \succ C$
3 agents: $B \succ C \succ A$
498 agents: $C \succ B \succ A$

- What is the Condorcet winner?

  B

- What would win under plurality voting?

  A

- What would win under plurality with elimination?

  C

Sensitivity to Losing Candidate

35 agents: $A \succ C \succ B$
33 agents: $B \succ A \succ C$
32 agents: $C \succ B \succ A$

- What candidate wins under plurality voting?

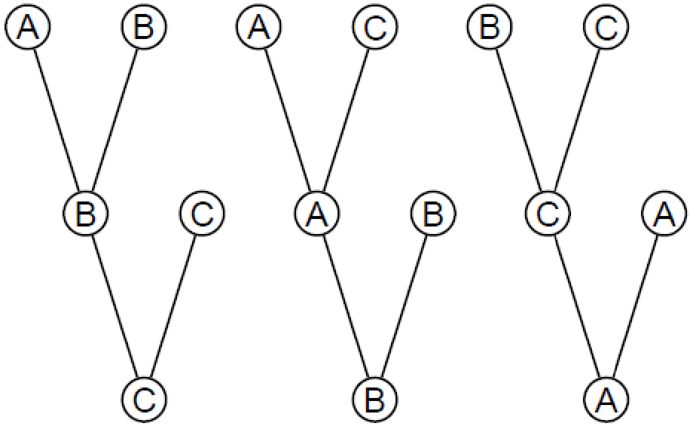
  A
- What candidate wins under Borda voting?

  A
- Now consider dropping C. Now what happens under both Borda and plurality?

  B wins

Sensitivity to Agenda Setter

35 agents: $A \succ C \succ B$
33 agents: $B \succ A \succ C$
32 agents: $C \succ B \succ A$



- Who wins pairwise elimination, with the ordering A, B, C?
  C
- Who wins with the ordering A, C, B?
  B
- Who wins with the ordering B, C, A?
  A

Another Pairwise Elimination Problem

1 agent: $B \succ D \succ C \succ A$
1 agent: $A \succ B \succ D \succ C$
1 agent: $C \succ A \succ B \succ D$

- Who wins under pairwise elimination with the ordering A, B, C, D?
  - D
- What is the problem with this?
  - all of the agents prefer B to D – the selected candidate is **Pareto-dominated!**

Desirable Properties

- Pareto Efficiency
- Independence of Irrelevant Alternatives (IIA)
- Nondictatorship

Pareto Efficiency

Definition (Pareto Efficiency (PE))

W is Pareto efficient if for any $o_1, o_2 \in O$, $\forall i \, o_1 \succ_i o_2$ implies that $o_1 \succ_W o_2$.

- when all agents agree on the ordering of two outcomes, the social welfare function must select that ordering

Independence of Irrelevant Alternatives (IIA)

Definition (Independence of Irrelevant Alternatives (IIA))

W is independent of irrelevant alternatives if, for any $o_1, o_2 \in O$ and any two preference profiles $[\succ'], [\succ''] \in L^n$, $\forall i$ $(o_1 \succ'_i o_2$ if and only if $o_1 \succ''_i o_2)$ implies that $(o_1 \succ_{W([\succ'])} o_2$ if and only if $o_1 \succ_{W([\succ''])} o_2)$.

- the selected ordering between two outcomes should depend only on the relative orderings they are given by the agents

- In a Borda count election, 5 voters rank 5 alternatives [*A*, *B*, *C*, *D*, *E*]: 3 voters rank [*A*>*B*>*C*>*D*>*E*]. 1 voter ranks [*C*>*D*>*E*>*B*>*A*]. 1 voter ranks [*E*>*C*>*D*>*B*>*A*].
  - Borda count (*a*=0, *b*=1): *C*=13, *A*=12, *B*=11, *D*=8, *E*=6. ***C* wins.**
- Now, the voter who ranks [*C*>*D*>*E*>*B*>*A*] instead ranks [*C*>*B*>*E*>*D*>*A*]; and the voter who ranks [*E*>*C*>*D*>*B*>*A*] instead ranks [*E*>*C*>*B*>*D*>*A*]. Note that they change their preferences only over the pairs [*B*, *D*], [*B*, *E*] and [*D*, *E*].
  - The new Borda count: *B*=14, *C*=13, *A*=12, *E*=6, *D*=5. ***B* wins**.
- Note that the social choice has changed the ranking of [*B*, *A*] and [*B*, *C*]. The changes in the social choice ranking are dependent on irrelevant changes in the preference profile. In particular, ***B* now wins instead of *C*, even though no voter changed their preference over [*B*, *C*]**.

Nondictatorship

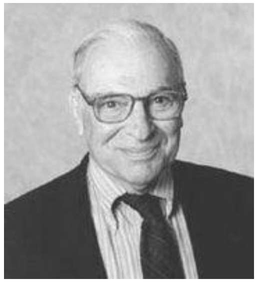
Definition (Non-dictatorship)

W does not have a dictator if $\neg\exists i\, \forall o_1, o_2 (o_1 \succ_i o_2 \Rightarrow o_1 \succ_W o_2)$.

- there does not exist a single agent whose preferences always determine the social ordering.
- We say that W is dictatorial if it fails to satisfy this property.

Arrow's Theorem

- Overall vision in social choice theory: identify "good" social choice procedures
- Unfortunately, a fundamental theoretical result gets in the way



Kenneth Joseph Arrow
Nobel prize in Economics.

Theorem (Arrow, 1951)

Any social welfare function W that is Pareto efficient and independent of irrelevant alternatives is dictatorial.

- Disappointing, basically means we can never achieve combination of good properties without dictatorship

Arrow's Theorem Proof, Step 1

- Proof: We will assume that W is both PE and IIA, and show that W must be dictatorial. Our assumption that $|O| \geq 3$ is necessary for this proof. The argument proceeds in four steps.

Step 1: If every voter puts an outcome b at either the very top or the very bottom of his preference list, b must be at either the very top or very bottom of $\succ_W$ as well.

Consider an arbitrary preference profile $[\succ]$ in which every voter ranks some $b \in O$ at either the very bottom or very top, and assume for contradiction that the above claim is not true. Then, there must exist some pair of distinct outcomes $a, c \in O$ for which $a \succ_W b$ and $b \succ_W c$.

Arrow's Theorem Proof, Step 1

Step 1: If every voter puts an outcome b at either the very top or the very bottom of his preference list, b must be at either the very top or very bottom of $\succ_W$ as well.

Now let's modify $[\succ]$ so that every voter moves c just above a in his preference ranking, and otherwise leaves the ranking unchanged; let's call this new preference profile $[\succ']$. We know from IIA that for $a \succ_W b$ or $b \succ_W c$ to change, the pairwise relationship between a and b and/or the pairwise relationship between b and c would have to change. However, since b occupies an extremal position for all voters, c can be moved above a without changing either of these pairwise relationships. Thus in profile $[\succ']$ it is also the case that $a \succ_W b$ and $b \succ_W c$. From this fact and from transitivity, we have that $a \succ_W c$. However, in $[\succ']$ every voter ranks c above a and so PE requires that $c \succ_W a$. We have a contradiction.

Arrow's Theorem Proof, Step 2

Step 2: There is some voter n^* who is extremely pivotal in the sense that by changing his vote at some profile, he can move a given outcome b from the bottom of the social ranking to the top.

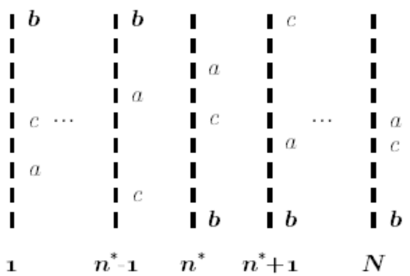
Consider a preference profile $[\succ]$ in which every voter ranks b last, and in which preferences are otherwise arbitrary. By PE, W must also rank b last. Now let voters from 1 to n successively modify $[\succ]$ by moving b from the bottom of their rankings to the top, preserving all other relative rankings. Denote as n^* the first voter whose change causes the social ranking of b to change. There clearly must be some such voter: when the voter n moves b to the top of his ranking, PE will require that b be ranked at the top of the social ranking.

Arrow's Theorem Proof, Step 2

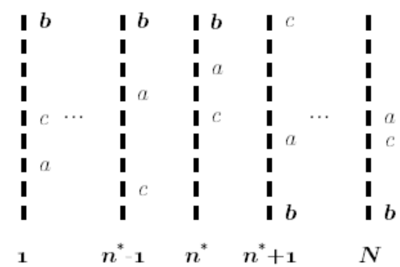
Step 2: There is some voter n^* who is extremely pivotal in the sense that by changing his vote at some profile, he can move a given outcome b from the bottom of the social ranking to the top.

Denote by $[\succ^1]$ the preference profile just before n^* moves b, and denote by $[\succ^2]$ the preference profile just after n^* has moved b to the top of his ranking. In $[\succ^1]$, b is at the bottom in $\succ_W$. In $[\succ^2]$, b has changed its position in $\succ_W$, and every voter ranks b at either the top or the bottom. By the argument from Step 1, in $[\succ^2]$ b must be ranked at the top of $\succ_W$.

Profile $[\succ^1]$:



Profile $[\succ^2]$:

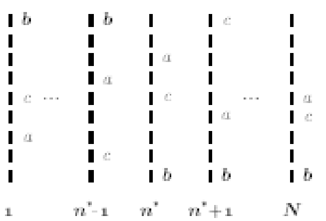


Arrow's Theorem, Step 3

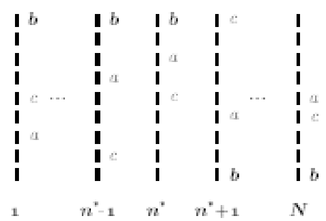
Step 3: n^* (the agent who is extremely pivotal on outcome b) is a dictator over any pair ac not involving b.

We begin by choosing one element from the pair ac; without loss of generality, let's choose a. We'll construct a new preference profile $[\succ^3]$ from $[\succ^2]$ by making two changes. First, we move a to the top of n^*'s preference ordering, leaving it otherwise unchanged; thus $a \succ_{n^*} b \succ_{n^*} c$. Second, we arbitrarily rearrange the relative rankings of a and c for all voters other than n^*, while leaving b in its extremal position.

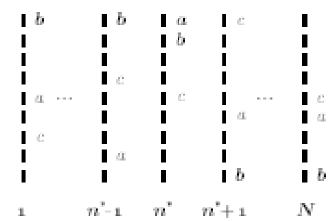
Profile $[\succ^1]$:



Profile $[\succ^2]$:



Profile $[\succ^3]$:

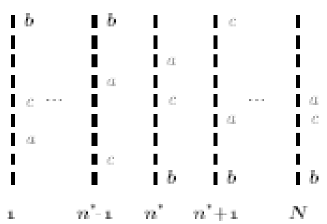


Arrow's Theorem, Step 3

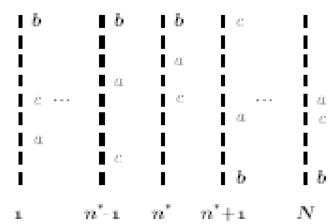
Step 3: n^* (the agent who is extremely pivotal on outcome b) is a dictator over any pair ac not involving b.

In $[\succ^1]$ we had $a \succ_W b$, as b was at the very bottom of $\succ_W$. When we compare $[\succ^1]$ to $[\succ^3]$, relative rankings between a and b are the same for all voters. Thus, by IIA, we must have $a \succ_W b$ in $[\succ^3]$ as well. In $[\succ^2]$ we had $b \succ_W c$, as b was at the very top of $\succ_W$. Relative rankings between b and c are the same in $[\succ^2]$ and $[\succ^3]$. Thus in $[\succ^3]$, $b \succ_W c$. Using the two above facts about $[\succ^3]$ and transitivity, we can conclude that $a \succ_W c$ in $[\succ^3]$.

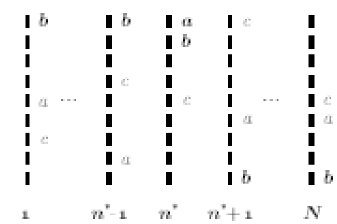
Profile $[\succ^1]$:



Profile $[\succ^2]$:



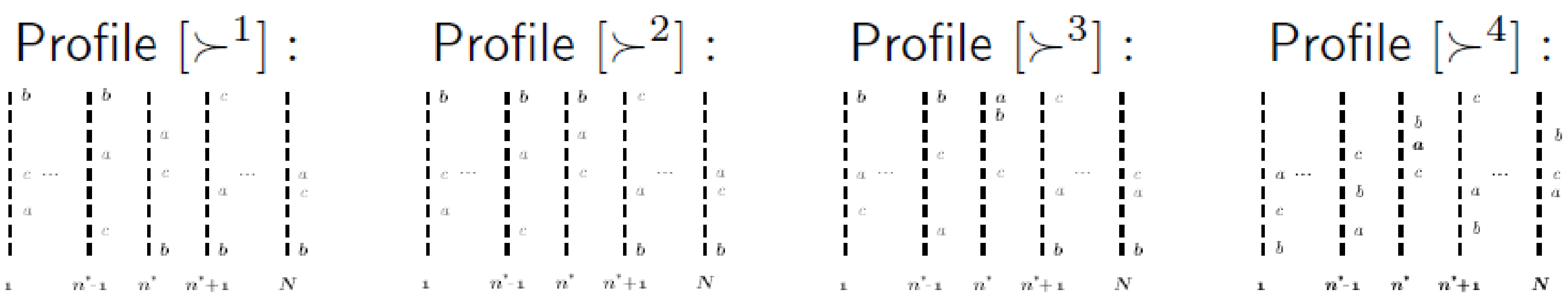
Profile $[\succ^3]$:



Arrow's Theorem, Step 3

Step 3: n^* (the agent who is extremely pivotal on outcome b) is a dictator over any pair ac not involving b.

Now construct one more preference profile, $[\succ^4]$, by changing $[\succ^3]$ in two ways. First, arbitrarily change the position of b in each voter's ordering while keeping all other relative preferences the same. Second, move a to an arbitrary position in n^*'s preference ordering, with the constraint that a remains ranked higher than c. Observe that all voters other than n^* have entirely arbitrary preferences in $[\succ^4]$, while n^*'s preferences are arbitrary except that $a \succ_{n^*} c$.

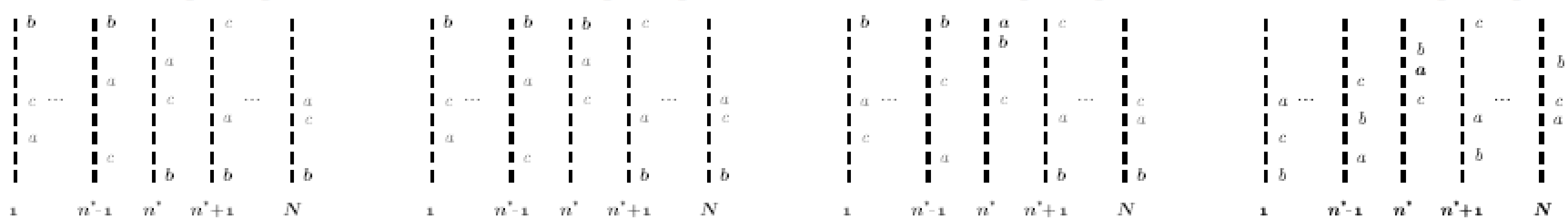


Arrow's Theorem, Step 3

Step 3: n^* (the agent who is extremely pivotal on outcome b) is a dictator over any pair ac not involving b.

In $[\succ^3]$ and $[\succ^4]$ all agents have the same relative preferences between a and c; thus, since $a \succ_W c$ in $[\succ^3]$ and by IIA, $a \succ_W c$ in $[\succ^4]$. Thus we have determined the social preference between a and c without assuming anything except that $a \succ_{n^*} c$.

Profile $[\succ^1]$: Profile $[\succ^2]$: Profile $[\succ^3]$: Profile $[\succ^4]$:



Arrow's Theorem, Step 4

Step 4: n^* is a dictator over all pairs ab.

Consider some third outcome c. By the argument in Step 2, there is a voter n^{**} who is extremely pivotal for c. By the argument in Step 3, n^{**} is a dictator over any pair $\alpha\beta$ not involving c. Of course, ab is such a pair $\alpha\beta$. We have already observed that n^* is able to affect W's ab ranking—for example, when n^* was able to change $a \succ_W b$ in profile $[\succ^1]$ into $b \succ_W a$ in profile $[\succ^2]$. Hence, n^{**} and n^* must be the same agent.

Strategic Manipulation

- We already saw that sometimes, voters can benefit by *strategically misrepresenting their preferences*, i.e., lying – tactical voting.
- Are there any voting methods which are *non-manipulable*, in the sense that voters can *never* benefit from misrepresenting preferences?

The Gibbard-Satterthwaite Theorem

The Gibbard-Satterthwaite Theorem

The only non-manipulable voting method satisfying the Pareto property for elections with more than 2 candidates is a dictatorship.

- In other words, every “realistic” voting method is prey to strategic manipulation . . .
- Fortunatelly, computational complexity of such manipulation can be prohibitive

Conclusions

- Aggregating preferences is a complex problem
- No single best voting mechanism exists
- Weight pros and cons for each particular application
- Reading: [Shoham] – 9.1 – 9.4