Definitions:

A **mechanism** is a procedure for making a decision/choosing an outcome or taking an action that interacts with "strategic" agents i.e. have preferences over the output

A mechanism is **strategy proof** (truthful, incentive compatible) if honesty/following the rules is always the best policy.
That is, lying cannot make you better off.

An outcome is **Pareto optimal** if in any other outcome, at least one agent is worse off.

Matching & Allocation

Office allocation:

n agents, each with an office, each with preference order on all offices

How to rearrange them to make people happy?

Top Trading Cycle Algorithm (TTCA)

while ∃ agents remaining

each remaining agent points to her favorite remaining office

Claim: ∃ a cycle

Reallocate according to some cycle

Delete those agents from graph.

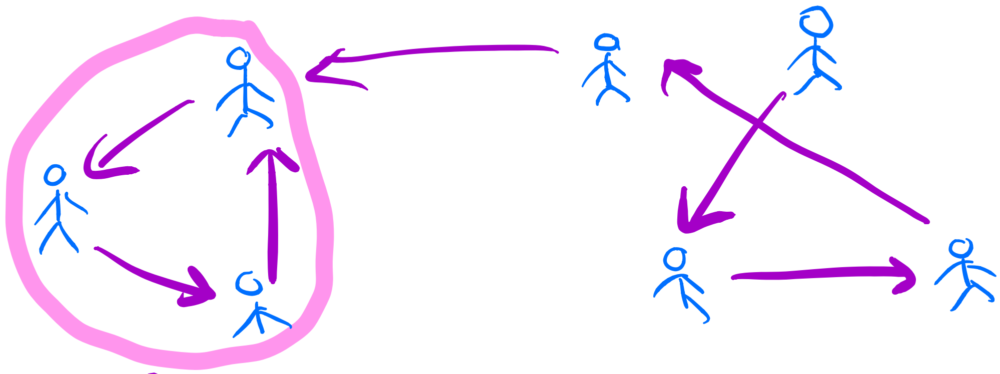
Thm: It is in each agent's best interest to tell the truth, no matter what anyone else does.

aka truthful, strategyproof, incentive-compatible.

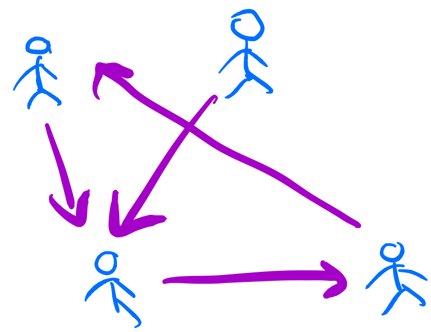
Proof: Fix agent i; suppose he was allocated in the j^{th} round (in cycle C_j) if he was truthful.

Claim: i cannot get any office in $C_1, \ldots, C_{j-1}$ because none point to i or to any office in $C_j, \ldots, C_{last}$.

But he gets his favorite office in $C_j, \ldots, C_{last}$ so no point in lying



C_1



If an agent is in C_k
he gets his favorite office
among all offices
in $C_k \dots C_m$

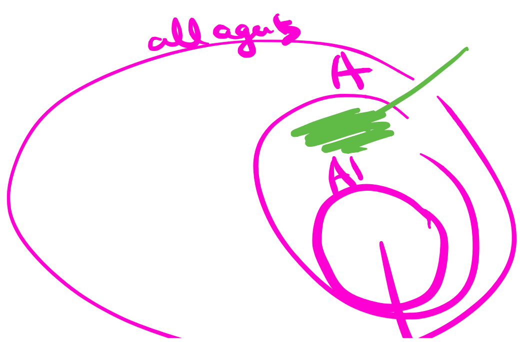
<u>Thm:</u> Allocation is <u>stable:</u>
impossible for subset of agents to go off &
reallocate on their own and all do at least
as well as with TTCA (with one strictly better)

By contradiction
<u>Pf</u> Suppose A subset of agents that all doing
at least as well, with at least one strictly
better.

Let $A' \subseteq A$ that get a different office
in separate allocation
than in TTCA

Define k to be index
of the first cycle that
contains any $a \in A'$

All $\tilde{a} \in A$ that are
in $C_1 \dots C_{k-1}$ get



same office under separate
as under TTCA

get something
from TTCA

$\Rightarrow$ a gets an office of some agent in A that is in ~~cycles~~ one of $C_k, \ldots, C_m$
but not same office he got under TTCA

$\Rightarrow$ he likes what he gets under separate less than under TTCA

Outcome is Pareto optimal:

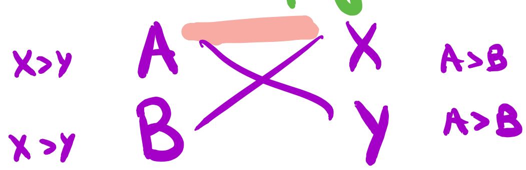
$\forall$ other allocation, there is at least one agent doing worse

Similar proof

Stable Matching

n hospitals (one slot)
n medical students
each hosp has pref order over students
student pref hospitals

Choose matching



A matching is <u>unstable</u> if $\exists\, h \in H,\ s \in S$
s.t.
- h & s are not matched
- h prefers s to $M(h)$
- s prefers h to $M(s)$

(h,s) are a "blocking pair"

Goal: find a stable matching.

Deferred Acceptance (DA) Alg
"hospital proposing"

U - all unmatched hospitals & students (initially all)
T - tentative matches (initially $\emptyset$)

While $\exists$ some $h \in U$ that hasn't yet "proposed" to all students
  h proposes to his favorite s that he hasn't yet proposed to
  if s unmatched or $(s,h') \in T$ but s prefers h to h'
    add (s,h) to T, put $h' \in U$
  else h stays unmatched.

At end, make all matches in T permanent → M

- $O(n^2)$ time to complete
- terminates with perfect matching

Observation: From time student is first proposed to, she is matched forever.

Claim:
There are no unstable (blocking) pairs

Pf: By contradiction. Suppose h-s are unstable pair.

Case 1: h never proposed to s
so he prefers his match.

Case 2: h did propose to s.
she rejected h for h' who she likes more
and the people she is subsequently matched to she likes better.
so she prefers her final match to h

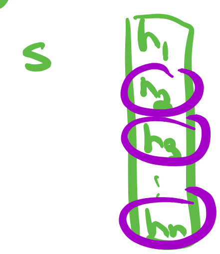
∃ can be multiple stable matchings

X > Y A — X B > A

Y > X B — Y A > B

What can we say about matching produced by DA?

Say that s is an $\boxed{\text{attainable match}}$ of h if $\exists$ stable matching in which they are matched.

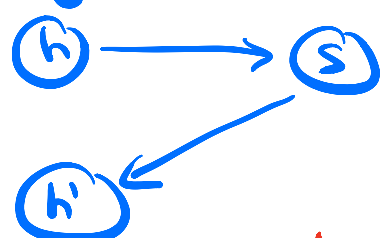


Let best(h) be their favorite attainable match

Claim: Output of DA is $\{(h, \text{best}(h))\}$

Proof by contradiction.

Consider some execution of DA & the first time anyone rejected by best attainable



s is best(h)
rejects h for h'

Since $s = \text{best}(h)$ $\exists$ M' in which h & s are matched. Say h' matched w/ s'

Claim s–h' is blocking pair

s prefers h' to h because she rejected h for h' in DA

h' prefers s to s' because in DA execution h' had not yet been rejected by best attainable

$\Rightarrow$ s is at least as good as h''s best attainable

$\square$

Claim Each s ends up with her worst attainable match.

Pf by contradiction

outcome of DA

h — s

worst attainable h'.

(s is best(h))

suppose h is not worst(s)

∃ stable matching in which h' & s are matched.

h' — s

h —

but this matching isn't stable. $\square$

DA is truthful for the proposers but not for other side.

Extensions

Unacceptable

Roommate matching 2n people

each person has ranking of all 2n-1 others

Quickie probability:

A random variable: a variable (takes one of several possible values) each with associated prob.

toss 2 dice

X : sum of two numbers

2, 3, 4, . . . , 12

$\frac{1}{36}$ $\frac{2}{36}$

$$E(X) = \sum_{\substack{k \text{ s.t.} \\ X \text{ takes} \\ \text{value } k}} k \Pr(X=k)$$

Uniform random vars

X is Uniform on $[a,b]$

$E(X) = \frac{a+b}{2}$

$\boxed{a=0, b=1 \quad [0,1]}$

$\Pr(X < 0.7) = 0.7$

$\Pr(X > 0.7) = 1$

$\Pr(X < x) = \frac{x-a}{b-a}$