MO758/MC758
Teoria dos Jogos Algorítmica

Rafael C. S. Schouery
rafael@ic.unicamp.br

Universidade Estadual de Campinas

1º semestre/2018

Projeto de Mecanismos sem Dinheiro

Projeto de Mecanismos sem Dinheiro

Teorema [Gibbard-Satterthwaite]: Se f é uma função de escolha social à prova de estratégia sobre A, com $|A| \geq 3$, então f é uma ditadura.

- O teorema de Gibbard-Satterthwaite é um fator limitante no projeto de mecanismos
- Mas existem situações onde ele não se aplica
- Assim, podemos criar mecanismos interessantes mesmo com essa limitação
- O VCG e outros mecanismos utilizam dinheiro para incentivar os jogadores
- Porém, nem sempre é interessante ou possível usar dinheiro nos mecanismos

Problemas de alocação

Vamos considerar dois problemas de alocação:

- alocação de casas:
  - um grupo quer trocar de casas entre si
  - cada pessoa tem uma ordem de preferência na casa que ficará
- emparelhamento estável:
  - devemos casar um grupo de homens com um grupo de mulheres
  - cada pessoa tem uma ordem de preferência sobre o grupo do sexo oposto

Uma pessoa se importa apenas com o seu resultado e não com o resultado global

Alocação de casas

Alocação de bens indivisíveis:

- n agentes, cada um com uma única casa e uma ordem de preferência em todas as casas
- Objetivo: realocar as casas de um modo apropriado

As ordens de preferência são arbitrárias, mas a ordem sobre as alocações é restrita:

- Qualquer alocação que deixa o agente com uma mesma casa é equivalente

O teorema de Gibbard-Satterthwaite não se aplica!

Alocação de casas

Numere os agentes de 1 a n

- Uma alocação a é uma permutação de 1 a n
  - a_i é a casa que fica com o agente i
- A: conjunto de todas as alocações
- $\succ_i$: ordem de preferência do agente i
  - $x \succ_i y$ significa que i prefere a casa x do que a casa y

Objetivo: encontrar uma alocação estável das casas aos agentes

Estável:

- nenhum grupo de agentes pode se unir e trocar as casas entres eles
- obtendo uma alocação melhor para eles do que a alocação sugerida

Coalizão

Para $S \subseteq [n]$, seja $A(S) = \{z \in A : z_i \in S, \forall i \in S\}$

- $A(S)$: alocações onde agentes de S trocam casas entre si

S é uma coalizão bloqueadora para uma alocação a em A se existe z em $A(S)$ tal que

- existe $j \in S$ com $z_j \succ_j a_j$
- para todo $i \neq j$ em S, ou $z_i \succ_i a_i$ ou $z_i = a_i$

Núcleo: conjunto de alocações sem coalizão bloqueadora

Algoritmo TTC - top trading cycle

Grafo orientado:

- um vértice para cada agente
- arco de i para j se a casa de j é a favorita de i

Cada vértice tem grau de saída 1

- há exatamente um circuito em cada componente
- Circuitos são vértices-disjuntos

Seja $N = [n]$ e N_1 o conjunto de agentes em circuitos

Execute as trocas de casas dos circuitos em N_1

Repita: novo grafo com agentes de $N \setminus N_1$, e arco de i para j se a casa de j é a favorita de i em $N \setminus N_1$

Pare quando não sobrarem mais vértices

Núcleo

Teorema: O núcleo do problema da alocação de casas consiste exatamente de uma alocação

Prova: Vamos mostrar que se uma alocação está no núcleo ela precisa ser a devolvida pelo algoritmo TTC

Alocação em que $i \in N_1$ não recebe sua casa favorita tem N_1 como coalizão bloqueadora

Assim, qualquer alocação no núcleo dá para $i \in N_1$ sua casa favorita da mesma forma que o algoritmo

Qualquer alocação no núcleo precisa dar a casa favorita em $N \setminus N_1$ para um agente em N_2

- se der uma casa melhor, N_1 é coalizão bloqueadora
- se der uma casa pior, N_2 é coalizão bloqueadora

Qualquer alocação no núcleo precisa dar a casa favorita em $N \setminus \cup_{i=1}^{k-1} N_i$ para um agente em N_k

Núcleo

Ou seja, se existe uma alocação no núcleo, ela precisa ser igual a alocação encontrada pelo algoritmo TTC

Mas será que a alocação a do algoritmo está no núcleo?

Suponha que não e seja S é uma coalizão bloqueadora para a

- Existe z em $A(S)$ tal que:
  - existe $j \in S$ com $z_j \succ_j a_j$
  - para todo $i \neq j$ em S, ou $z_i \succ_i a_i$ ou $z_i = a_i$

Escolha $j \in N_k$ com $z_j \succ_j a_j$ e k mínimo

- Isto é, $z_i = a_i$ para todo $i \in N_r \cap S$ com $r < k$

Núcleo

Escolha $j \in N_k$ com $z_j \succ_j a_j$ e k mínimo

- Isto é, $z_i = a_i$ para todo $i \in N_r \cap S$ com $r < k$

Como j não ficou com a casa z_j, ela foi alocada antes do turno k (caso contrário, j apontaria para z_j no turno k e não para a_j)

Assim, $z_j \in N_r \cap S$ e z_j recebe a mesma casa em a e em z

- Isto é, a casa para qual z_j aponta no circuito do passo r também está em S
- De fato, cada um dos elementos desse circuito precisam estar em S
- Em particular, o que apontava para z_j
  - ele não pode ter piorado e, pela escolha de k, não melhorou, então ainda leva z_j
  - contradição com o fato que j leva z_j

TTC é à prova de estratégia

Teorema: O mecanismo TTC é à prova de estratégia

Prova: Seja

- π : as reais preferências do agentes
- a : alocação devolvida pelo TTC com entrada π
- π' : preferências quando um agente $j \in N_k$ mente
- a' : alocação devolvida pelo TTC com entrada π'

Se o TTC não é à prova de estratégia, então $a'_j \succ_j a_j$

Como j, só faz parte de um circuito no turno k

- $a_i = a'_i$ para todo $i \in \bigcup_{r=1}^{k-1} N_r$
- e $a'_j \in N \setminus \bigcup_{r=1}^{k-1} N_r$

Mas, o algoritmo dá para j a melhor casa de $N \setminus \bigcup_{r=1}^{k-1} N_r$

Casamentos estáveis

Temos:

- H : conjunto de homens
- M : conjunto de mulheres
- Cada homem tem uma ordem de preferência sobre M
- Cada mulher tem uma ordem de preferência sobre H
- Adicione homem/mulher fictício para representar a possibilidade de ficar solteiro

Assim $|H| = |M|$

Emparelhamento de H em M: alocação de homens a mulheres

Casamentos estáveis

Um emparelhamento é instável se existem homens h e h' e mulheres m e m' tais que

- m é emparelhada com h, mas $m' \succ_h m$ e
- m' é emparelhada com h', mas $h \succ_{m'} h'$

Neste caso, h e m' preferiam se casar um com o outro

O par (h, m') é um par bloqueador

Um emparelhamento é estável se não tem par bloqueador

Casamentos estáveis: exemplo

Ordens de preferências para $n = 3$:

$$\begin{array}{cccccc}
\succ_{h_1} & \succ_{h_2} & \succ_{h_3} & \succ_{m_1} & \succ_{m_2} & \succ_{m_3} \\
m_2 & m_1 & m_1 & h_1 & h_3 & h_1 \\
m_1 & m_3 & m_2 & h_3 & h_1 & h_3 \\
m_3 & m_2 & m_3 & h_2 & h_2 & h_2
\end{array}$$

O emparelhamento $\{(h_1, m_1), (h_2, m_2), (h_3, m_3)\}$ é instável, pois (h_1, m_2) é um par bloqueador

Já o emparelhamento $\{(h_1, m_1), (h_2, m_3), (h_3, m_2)\}$ é estável

Dada as listas de preferências de todos, existe emparelhamento estável?

- Como encontrá-lo, se existe?

Algoritmo da aceitação postergada

Versão com proposta masculina

Primeira rodada:

- Cada homem propõe à primeira mulher de sua lista
- Cada mulher que recebeu mais de uma proposta recusa todas, exceto a do homem preferido entre os candidatos
  - Para esse, posterga a sua resposta

Nova rodada:

- Cada homem que teve sua proposta recusada propõe à próxima mulher de sua lista
- Cada mulher com mais de uma proposta recusa todas, exceto a do homem preferido entre os candidatos
  - Eventualmente recusa proposta recebida em rodada anterior

O processo termina em não mais que n^2 rodadas

Algoritmo no exemplo

$\succ_{h_1}$	$\succ_{h_2}$	$\succ_{h_3}$	$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$
m_2	m_1	m_1	h_1	h_3	h_1
m_1	m_3	m_2	h_3	h_1	h_3
m_3	m_2	m_3	h_2	h_2	h_2

Rodada 1:

m_1: $h_2 \quad h_3$

m_2: h_1

m_3:

Rodada 2:

m_1: $h_2 \cancel{h_2} \quad h_3$

m_2: h_1

m_3: h_2

Emparelhamento produzido: $\{(h_1, m_2), (h_2, m_3), (h_3, m_1)\}$

Estabilidade do emparelhamento

Teorema. O emparelhamento produzido pelo algoritmo da aceitação postergada é estável

Prova: Suponha que existe um par bloqueador (h_1, m_1)

- onde h_1 está emparelhado com m_2
- e m_1 está emparelhado com h_2

Então:

- como $m_1 \succ_{h_1} m_2$, h_1 propôs para m_1 antes de propor para m_2
- como h_1 ficou com m_2, m_1 trocou h_1 por um homem melhor
- então $h_2 \succ_{m_1} h_1$

ou seja, (h_1, m_1) não é um par bloqueador

Núcleo

Num emparelhamento estável um par (h, m) não podem melhorar sozinhos

- Nenhum (h, m) prefere sair do mecanismo e se casar

E se considerarmos um grupo que sai do mecanismo?

Um emparelhamento ν domina um emparelhamento μ se existe $S \subseteq H \cup M$ tal que para todo h e m em S, temos

- $\nu(h)$ e $\nu(m)$ pertencem a S
- $\nu(h) \succ_h \mu(h)$ e $\nu(m) \succ_m \mu(m)$

Um emparelhamento μ está no núcleo se e somente se não existe emparelhamento ν que o domina

Teorema. O núcleo do jogo de emparelhamento é o conjunto de todos os emparelhamentos estáveis

Estabilidade do emparelhamento

No exemplo, aplicando a versão da proposta feminina, obtemos o emparelhamento $\{(h_1, m_1), (h_2, m_3), (h_3, m_2)\}$.

$$\begin{array}{cccccc}
\succ_{h_1} & \succ_{h_2} & \succ_{h_3} & \succ_{m_1} & \succ_{m_2} & \succ_{m_3} \\
m_2 & m_1 & m_1 & h_1 & h_3 & h_1 \\
m_1 & m_3 & m_2 & h_3 & h_1 & h_3 \\
m_3 & m_2 & m_3 & h_2 & h_2 & h_2
\end{array}$$

Diferente do obtido pela proposta masculina!

Há alguma diferença significativa entre estes emparelhamentos?

Emparelhamentos ótimos

Emparelhamento ν é ótimo-masculino se não há emparelhamento estável μ tal que existe j em H com $\mu(j) \succ_j \nu(j)$ e para todo h em H, $\mu(h) \succ_h \nu(h)$ ou $\mu(h) = \nu(h)$

Definição de emparelhamento ótimo-feminino é análoga

Emparelhamentos ótimos

Teorema: O algoritmo da aceitação postergada com proposta masculina produz um emparelhamento ótimo-masculino

Prova: Seja μ o emparelhamento encontrado pelo algoritmo e suponha que μ não é ótimo-masculino

Existe emparelhamento estável ν tal que existe j em H com $\nu(j) \succ_j \mu(j)$ e para todo h em H, $\nu(h) \succ_h \mu(h)$ ou $\nu(h) = \mu(h)$

- j propôs primeiro para $\nu(j)$ e foi rejeitado
- escolha j como o primeiro homem que foi rejeitado pela mulher $\nu(j)$

Emparelhamentos ótimos

- j : primeiro homem que foi rejeitado pela mulher $\nu(j)$
- $\nu(j) \succ_j \mu(j)$

Então:

- $\nu(j)$ recebe uma proposta de um homem i que ela prefere
- i prefere a mulher $\nu(j)$ à mulher $\nu(i)$
  - caso contrário, i já teria sido rejeitado por $\nu(i)$ e não teríamos escolhido j
- portanto, $i \succ_{\nu(j)} j$ e $\nu(j) \succ_i \nu(i)$
  - ν não é estável

À prova de estratégia

Teorema. O algoritmo da aceitação postergada com proposta masculina é um mecanismo à prova de estratégia para os homens.

Prova:

- $\pi = (\succ_{h_1}, \succ_{h_2}, \ldots, \succ_{h_n})$
- μ: emparelhamento encontrado pelo algoritmo com preferências masculinas π
- suponha que homem h_1 mente trocando $\succ_{h_1}$ por $\succ_*$
- $\pi^1 = (\succ_*, \succ_{h_2}, \ldots, \succ_{h_n})$
- ν: emparelhamento encontrado pelo algoritmo com preferências masculinas π^1

Vamos mostrar que, se $\nu(h_1) \succ_{h_1} \mu(h_1)$ então ν não é estável

À prova de estratégia

$$\pi = (\succ_{h_1}, \ldots) \qquad \pi^1 = (\succ_*, \ldots)$$

$$\pi \to \mu \qquad \pi^1 \to \nu$$

Seja $R = \{h\colon \nu(h) \succ_h \mu(h)\}$, vamos mostrar que para $h \in R$ onde $m = \nu(h)$ e $h' = \mu(m)$, h' pertence a R

- Se $h' = h_1$, nada a fazer (já supomos que $h_1 \in R$)
- Caso contrário, como $m \succ_h \mu(h)$ a estabilidade de μ implica que $h' \succ_m h$
- a estabilidade de ν para π^1 implica que $\nu(h') \succ_{h'} m$
- portanto $h' \in R$
- podemos definir $S = \{m\colon \nu(h) \in R\} = \{m\colon \mu(h) \in R\}$

À prova de estratégia

$$\pi = (\succ_{h_1}, \ldots) \qquad \pi^1 = (\succ_*, \ldots)$$

$$\pi \to \mu \qquad \pi^1 \to \nu$$

$R = \{h\colon \nu(h) \succ_h \mu(h)\}$

$S = \{m\colon \nu(h) \in R\} = \{m\colon \mu(h) \in R\}$

$\mu(m) \succ_m \nu(m)$ para todo $m \in S$

Seja h o último homem em R a fazer uma proposta:

- proposta é feita para $m = \mu(h) \in S$ que rejeitou $\nu(m)$ em alguma iteração anterior
- quando h propõe para m, ela rejeita uma proposta de algum $h' \notin R$ tal que $h' \succ_m \nu(m)$
- Como $h' \notin R$, temos que $m \succ_{h'} \mu(h') \succeq_{h'} \nu(h')$
- Como $h' \neq h_1$, (h', m) é um par bloqueador para ν em π^1

À prova de estratégia (?)

Vimos que o algoritmo da aceitação postergada com proposta masculina é à prova de estratégia para os homens

E para as mulheres?

Voltamos ao exemplo...

Algoritmo no exemplo

$\succ_{h_1}$	$\succ_{h_2}$	$\succ_{h_3}$	$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$
m_2	m_1	m_1	h_1	h_3	h_1
m_1	m_3	m_2	h_3	h_1	h_3
m_3	m_2	m_3	h_2	h_2	h_2

Rodada 1:

m_1: h_2 h_3

m_2: h_1

m_3:

Rodada 2:

m_1: $h_2 \cancel{h_2}$ h_3

m_2: h_1

m_3: h_2

Emparelhamento produzido: $\{(h_1, m_2), (h_2, m_3), (h_3, m_1)\}$

Algoritmo no exemplo

$\succ_{h_1}$	$\succ_{h_2}$	$\succ_{h_3}$	$\succ_{m_1}$	$\succ_{m_2}$	$\succ_{m_3}$
m_2	m_1	m_1	h_1	h_3	h_1
m_1	m_3	m_2	h_3	h_1	h_3
m_3	m_2	m_3	h_2	h_2	h_2

Rodada 1:

$m_1: \quad h_2 \quad h_3$

$m_2: \quad h_1$

$m_3:$

Rodada 2:

$m_1: \quad h_2 \quad h_3 \cancel{h_3}$

$m_2: \quad h_1 \quad h_3$

$m_3:$

Rodada 3:

$m_1: \quad h_2 \quad \cancel{h_3} \quad h_1$

Problema das admissões em universidades

No problema das admissões em universidades:

- Temos um conjunto $C = \{c_1, \ldots, c_n\}$ de universidades
  - Cada universidade c_i tem uma quota q_i
- e temos um conjunto $S = \{s_1, \ldots, s_m\}$ de alunos
- Cada aluno tem uma ordem de preferência sobre as universidades
- Cada universidade c_i tem uma ordem de preferência sobre os subconjuntos de alunos de tamanho q_i

Problema da admissão em universidades

Roth'85: *The college admissions problem is not equivalent to the marriage problem*

Para o problema do casamento estável vale que

- não existe alocação instável que todos os homens preferem a alocação ótima-masculina
- existe mecanismo à prova de estratégia para os homens

Esses resultados não valem para o problema das admissões em universidades

A dificuldade reside no fato que as universidades têm preferências sobre subconjuntos de alunos

Suplementar vs. Complementar

Se considerarmos que os alunos seguem uma suplementariedade:

- Isto é, as preferências das universidades não são sobre subconjuntos
- mas sim de uma preferência sobre os alunos

Então, o problema é equivalente ao problema do casamento estável

- basta dividir uma universidade com cota q_i
- em q_i universidades de cota 1

Suplementar vs. Complementar

Porém, existem situações onde complementariedade existe:

- Casais de médicos em programas de residência
- Casais de crianças em programas de adoção

Ou seja, nem sempre é possível usar o algoritmo da proposta postergada