Budgeted Social Choice: A Framework for Multiple Recommendations in Consensus Decision Making

Tyler Lu, Craig Boutilier

Department of Computer Science,

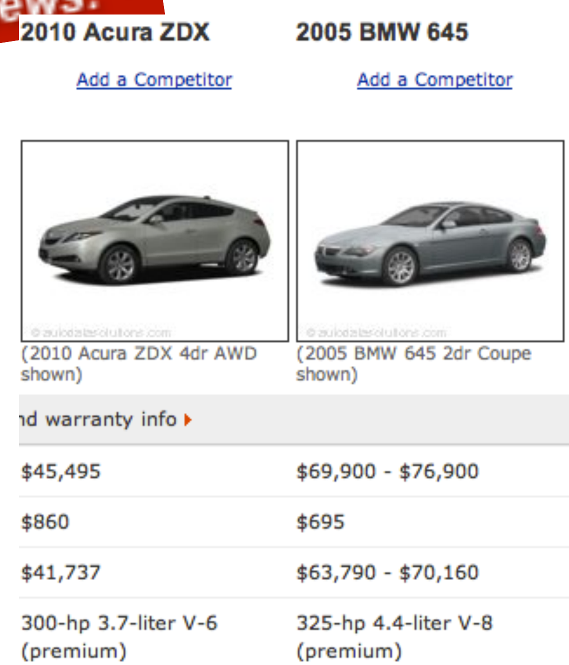
University of Toronto

VELUT ARBOR fiVO

VELUT ARBOR fiVO

Background

- Lots of preference data generated nowadays
  - Search clicks, movie ratings, product purchases, …



Background

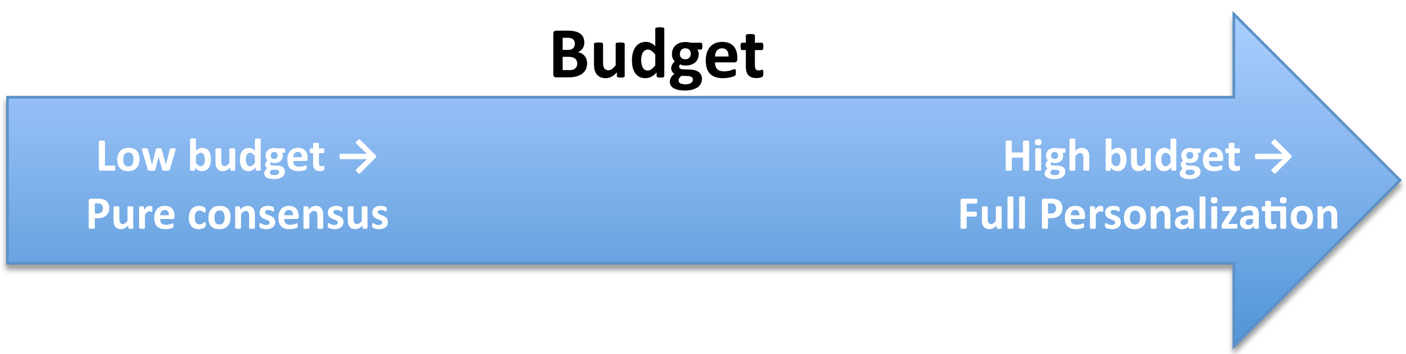
- Recommender systems facilitate personalized product suggestions

Motivation

- However sometimes cannot make personalized recommendations
  - Privacy concerns, lack of data
  - Limits on inventory, factory production limits
  - Public projects (e.g. new bus routes; park location)
- More generally, *constraints* on the number of recommendations ("items") that can be offered

Budgeted Social Choice

- **Providing a middle ground**: assume a budget; consensus decision must not exceed budget.
  - Can build 2 to 4 new bus routes given $1 million
  - Can configure at most 5 different product lines
- Allows for a spectrum of problems:



- Comes in a variety of flavors depending on the nature of the budget

Our Contributions

- New class of problems bridging personalization vs. group decision making
- Generalization of proportional representation via a budget
- Algorithms & analysis
  - General budgeted social choice
  - Limited Choice/Proportional representation
- Experiments on real data

Model 1: Limited Choice

(Illustrative example of budgeted social choice)

- Alternatives $A = \{a_1, \ldots, a_m\}$
- Preference profile $V = (v_1, \ldots, v_n)$ where v_i is a ranking
- Positional scoring function α assigns rank position to a non-negative score (e.g. Borda), non-increasing
- Given $K \geq 1$, find Φ subset A size at most K
  - Φ is the "recommendation set"

Goal: $$\max_{\Phi} \boxed{\sum_{\ell=1}^{n} \max_{a \in \Phi} \alpha(v_\ell(a))}$$

Score of Φ

$S_\alpha(\Phi)$

Limited Choice Examples

Video rental store <u>must decide what new releases to procure</u>.

Has budget to get ***4 new movies***.

Which 4 to choose??

New releases Feb. 16, 2010

Amreeka
Black Dynamite
Cabin Fever 2
Cairo Station
Coco Before Chanel
Contempt
Crude
…

Decision space Φ

N = # new movies

N choose 4 subsets

Limited Choice Examples

- Which movies to get depends on what customers like

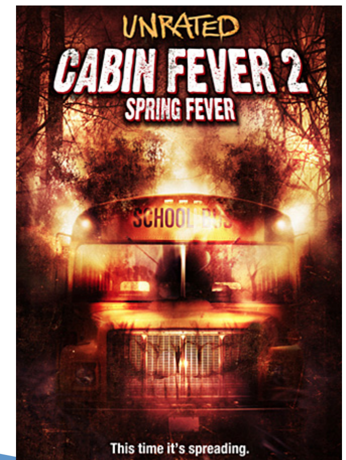
Rich	Craig	Tyler
Cabin Fever 2	Law Abiding Citizen	Hunger
Law Abiding Citizen	Cabin Fever 2	The Lady Killers
Hunger	The Lady Killers	Cabin Fever 2
The Lady Killers	Hunger	Law Abiding Citizen
…	…	…

- Single (social) choice: $K=1$, want to make as many customers as happy as possible
- Personalization: K is large, social choice less of an issue, just get movies people want

Limited Choice Example

- Given what video rental store procures:

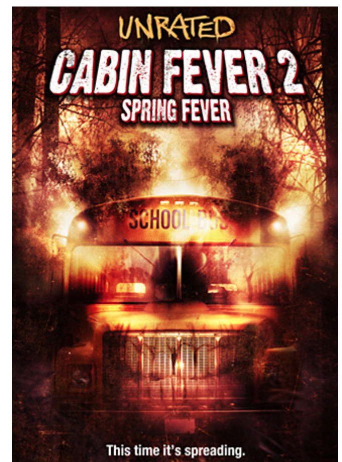
Movies (Φ) =



<u>Craig</u>
Law Abiding Citizen
Cabin Fever 2
The Lady Killers
Lovecraft: Fear of Un..
…

Craig benefits from the ***most preferred***, gets some "satisfaction" e.g. Borda score of 3

Limited Choice Example

Movies (Φ) = {   }

Total Borda score = $S_{\alpha}(\Phi)$ = **Rich's score + Craig's score + Tyler's score**

= 4 + 3 + 3

Observations on Limited Choice Model

- Corresponds to Chamberlin & Courant'83, on proportional representation
- Need not be utilitarian: can allow fairness

$$\max_{\Phi} \min_{\ell} \max_{a \in \Phi} \alpha(v_\ell(a))$$

- **Theorem** If α is the Borda score, given K, x, deciding if there is a slate Φ with $S_\alpha(\Phi) \geq x$ is NP-complete
  - Related but different result in Procaccia et al'08

Observations on Limited Choice Model

- How does LCM compare with general positional score ranking (including Borda)?
- **Theorem** If scores are Borda, then picking the top elements *K* of the Borda ranking is a 1/2-approximation to the LCM-optimal slate (tight bound). For arbitrary positional scoring, then picking top *K* can be at least a factor of *K* worse than LCM-opt.
- *Reason*: Positional ranking biases to *popular* alternatives, while limited choice aims for *diversity* of alternatives

Greedy Algorithm for LCM

- LCM-opt can be formulated as an IP with #vars, #constraints = O(#votes #items) [Potthoff & Brams'98]
- Easy to see that $S_\alpha(\Phi)$ is *submodular*
- Greedy approx. with ratio 1-1/e (Nemhauser et al.'78)
  1. *Φ = empty set*
  2. *Run for K steps:*

     Update Φ with $\arg\max_a S_\alpha(\Phi \cup \{a\})$

Budgeted Social Choice: General Form

<u>What to have for banquet?</u>

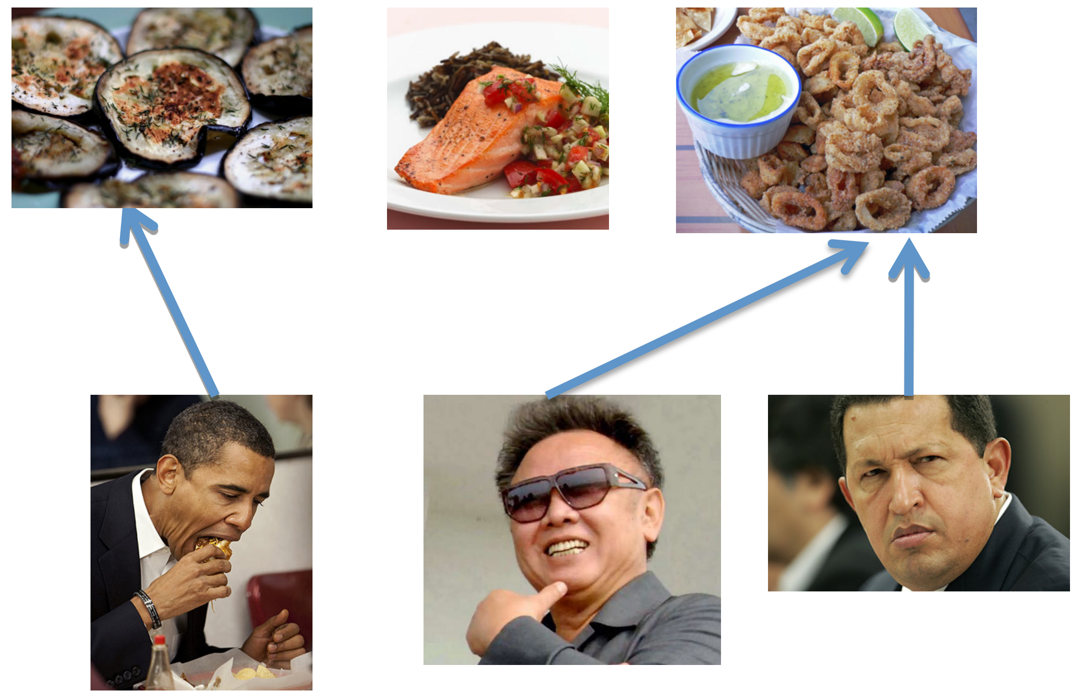
Budget *B* | Alternatives (dishes)

Fixed costs (e.g., equipment, staff needed to cook): t_a
Unit costs (e.g., cost to produce each dish): u_a

- Can't recommend a subset of dishes, because *how many consumed* matters (e.g. if everyone picks the most expensive dish it will deplete budget)
- Instead use a *recommendation function*: an assignment of people to dishes

Budgeted Social Choice: General Form

Alternatives (dishes)



Recommendation function *Φ*:

Cost:

C(Φ) = Fixed + Unit =
$(t_{eggplant} + t_{calamari}) +$
$(u_{eggplant} + 2 \cdot u_{calamari})$

Total score: sum of individual scores (welfare)

$S_\alpha(\Phi)$ = α(Barack eggplant rank) + α(Kim calamari rank)
+ α(Hugo calamari rank)

Budgeted Social Choice: General Form

- The goal:

$$\max_{\Phi} S_{\alpha}(\Phi)$$
$$\text{s.t. } C(\Phi) \leq B$$

- **Specializations**
  - Limited choice: $t_a = 1$ and $u_a = 0$, $B = K$
  - Limited choice with costs: fixed cost varies, $u_a = 0$
- Full personalization: if we can afford everyone's favourite item
- We can have "unassigned" agents by adding a dummy item *d* with no costs

Greedy Algorithm for BSC

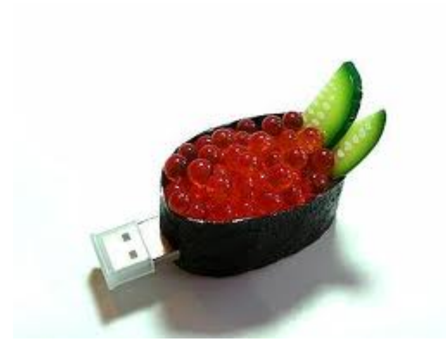
- For each item *a*, sort agents according to preference for *a*
- Find "sweet spot": #agents from sorted list that maximizes ratio of marginal score increase vs. marginal cost increase if they were assigned *a*
- Find the item *a** that maximizes sweet spot ratio and assign *a** to the *i** agents that maximizes the marginal ratio
- Repeat until budget depletes.
  - If minimal fairness required (all agents must be assigned) then do simple backtracking when budget is depleted

Experiments

- Limited Choice
- American Psychological Association 1980 election data – 5 candidates, ~5700 full votes
  - Academics and clinicians on "uneasy terms"
  - K = 2, limited choice gives an academic and clinician as optimal set ("diversity").
  - Greedy is suboptimal (Borda scores) but almost optimal and also gives academic and clinician as solution.

AMERICAN PSYCHOLOGICAL ASSOCIATION

Experiments

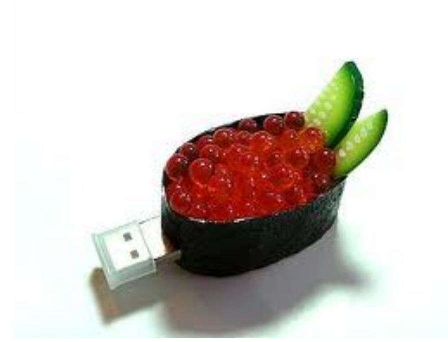


- Sushi dataset – 10 varieties of sushi, 5000 preference rankings from Japan
- Limited choice
  - Tried various *K*
  - Tried exponential, Borda, and cubic α
  - Greedy always finds optimal for all *K* (under 1 sec.)
  - Using Borda & Kemeny rankings gives good approximations
  - CPLEX is slow to solve IP, taking anywhere from 13-90 sec.

Experiments

K	Greedy	Borda	Kemeny	Random	CPLEX (sec.)
2	1.0	1.0	0.932	0.531	49.1
3	1.0	0.986	0.949	0.729	90.38
5	1.0	0.989	0.970	0.813	20.32
7	1.0	1.0	1.0	0.856	13.16

Experiments



- General Budgeted S.C.
  - Randomly generated fixed costs, unit costs were either zero or very small
  - Fixed budget, allowed for 2-5 unique items
  - Greedy is very good within 98-99% of optimal, with runtime 2-5 sec.
  - CPLEX is slow to solve IP, takes 2-5 min.

Conclusions

- Developed a class of problems that range from pure social choice to personalized choices
- Occurs in a variety of real life problems
  - Displaying products/items in electronic commerce
  - Search results, advertising, industrial optimization
- Fast greedy algorithms with excellent approx.
- Future work
  - Dealing with incomplete preferences
    - Using statistical inference/learning, robust inference
  - Trading off social welfare with budget, and other variations