When the Universe is Too Big

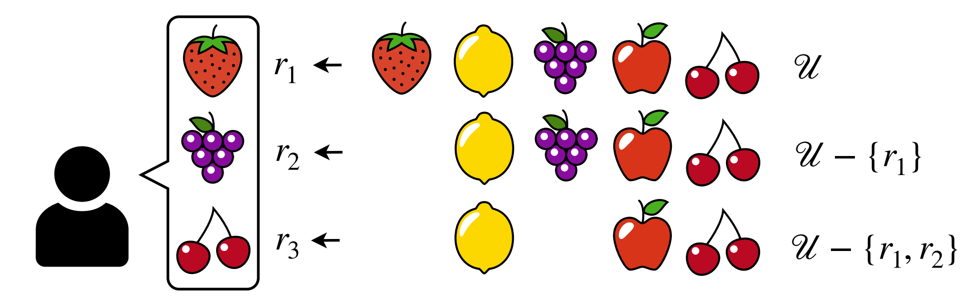
Bounding Consideration Probabilities for Plackett–Luce Rankings

Ben Aoki-Sherwood[1,2,3]
Catherine Bregou[1]
David Liben-Nowell[1]
Kiran Tomlinson[1,4,5]
Thomas Zeng[1,6]

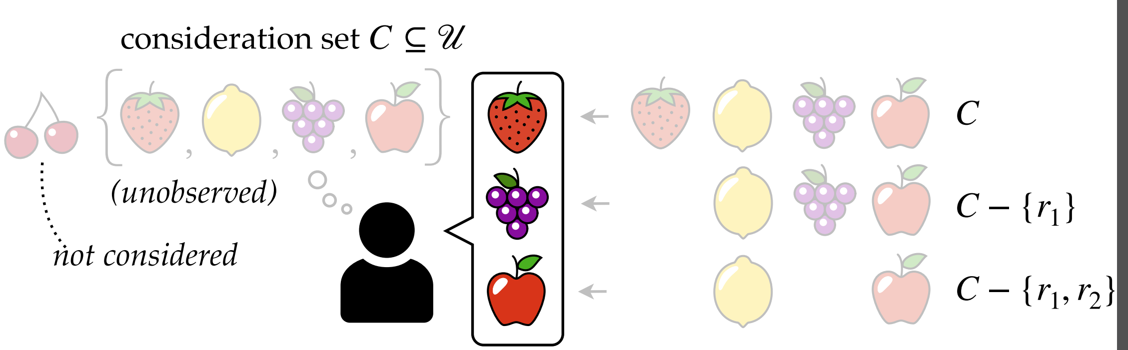
[1]Carleton College
[2]Johns Hopkins Applied Physics Lab
[3]University of Colorado
[4]Cornell University
[5]Microsoft Research
[6]University of Wisconsin

Background and Model

Choice-based ranking: top-k ranking r from universe $\mathscr{U}$



Consider-then-choose ranking: select a consideration set $C \subseteq \mathscr{U}$ to consider, then rank k elements from C.



Consideration Set Model

$$\Pr_C(C) = 0 \quad \text{if } |C| < k$$

$$\Pr_C(C) \propto \left(\prod_{i \in C} p_i\right) \prod_{j \in \mathscr{U} - C} (1 - p_j)$$

Selection Model: Plackett–Luce [1, 2]

$$\Pr_{PL}(r \mid C) = \prod_{i=1}^{k} \frac{\exp(u_{r_i})}{\sum_{j \in \mathscr{C} - \{r_{1 \ldots i-1}\}} \exp(u_j)}$$

Plackett–Luce with Consideration (PL+C)

1. People consider at least k items
2. Each item i is considered independently w.p. p_i
3. Given C, items are ranked by Plackett–Luce based on utility u_i

$$\Pr_{PL+C}(r) = \sum_{C \subseteq \mathscr{U}} \Pr_C(C) \Pr_{PL}(r \mid C)$$

Overview

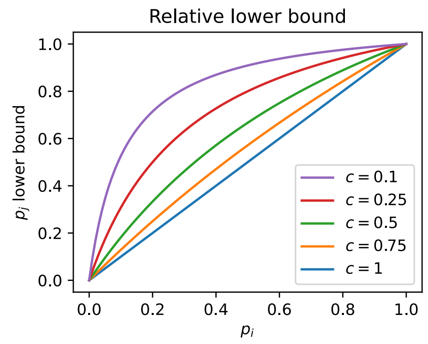
Q: ***Can we tell from rankings what items were considered?***
A: ***No.*** ... and yet, we can still derive:

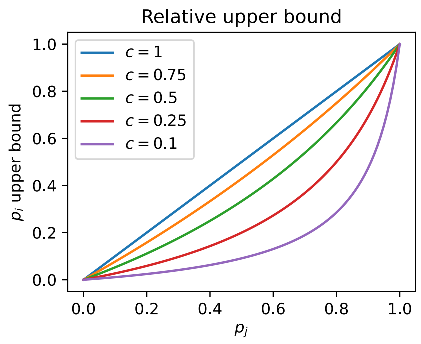
1. **relative bounds** on consideration probabilities, given known utilities.
2. **absolute bounds** on consideration probabilities, given known utilities and a lower bound on expected number of items considered.
3. **algorithms** to tighten our absolute bounds using our relative bounds.
4. **experimental results** based on psychological study data on Americans' perceptions of states' relative contribution to U.S. history [3].

Theoretical Results

Theorem 1: We can't learn consideration probabilities (in general). PL+C consideration probabilities aren't identifiable, even with known utilities.

Theorem 2: We can use "ranking flips" to infer relative consideration. If $u_i > u_j$, but i is ranked in the top-ℓ $c \le 1$ times as often, then $\frac{p_i}{1 - p_i} \le c \cdot \frac{p_j}{1 - p_j}$. Equivalently, $p_i \le \frac{cp_j}{1 - p_j + cp_j}$ and $p_j \ge \frac{p_i}{c - cp_i + p_i}$.





Theory, Continued

If $\sum_{i \in \mathscr{U}} p_i \ge \alpha k$ for $\alpha > 1$ ...

Theorem 3. ... then $\Pr_{PL+C}(\mathscr{R}_{i \le k}) \cdot \left[1 - \left(\alpha e^{1-\alpha}\right)^k\right] \le p_i$.

Theorem 4. ... then $p_i \le \frac{\sum_{j \in \mathscr{U}} \exp(u_j)}{\exp(u_i)} \cdot \left[\Pr_{PL+C}(\mathscr{R}_{i=1}) + \frac{k(\alpha e^{1-\alpha})^k}{1 - (\alpha e^{1-\alpha})^k}\right]$.

Pr[i appears in top-k rankings]; *correction (Chernoff)*; *1 / Pr[i ranked first] (ignoring consideration)*; *Pr[i ranked first]*

Thm. 3 (absolute lower bound): i must have been considered whenever it's ranked + correction for conditioning.

Thm. 4 (absolute upper bound): if i was ranked first less than expected, it can't have been considered + correction for conditioning.

Efficient bound propagation algorithm

Given: utilities u_i, top-ℓ ranking probabilities $\Pr_{PL+C}(\mathscr{R}_{i \le \ell})$, $\alpha > 1$ s.t. $\sum_{i \in \mathscr{U}} p_i \ge \alpha k$ ($\ge \alpha k$ items considered on average).

1. Initialize upper/lower bounds according to Theorem 3/4
2. Construct DAG G of all item reversals (utility vs top-ℓ ranking probability)
3. Propagate bounds per Thm. 2 along topological order of G

Experiments

Survey of Americans' perceptions of U.S. History [3]:

~3K participants.

Top-3 Q: What three states contributed most to U.S. history? ***ranking with consideration***

Random-10 Q: What percentage of U.S. history did *[10 random states]* contribute? ***ranking without consideration***

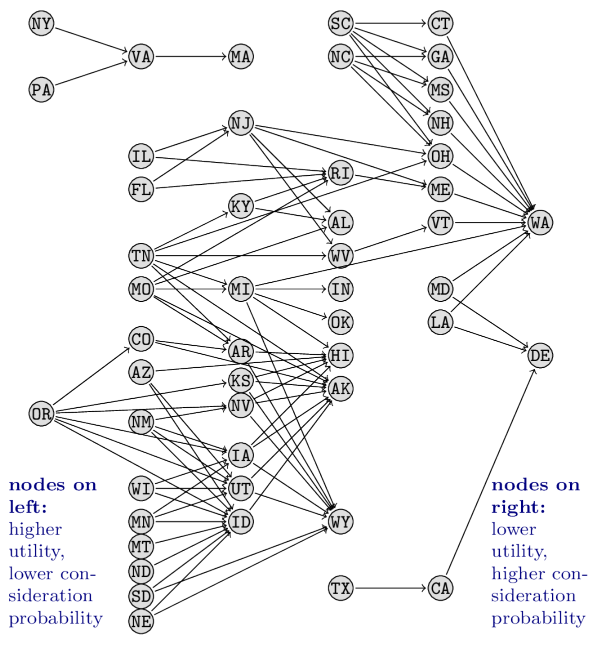
learn utilities from Random-10

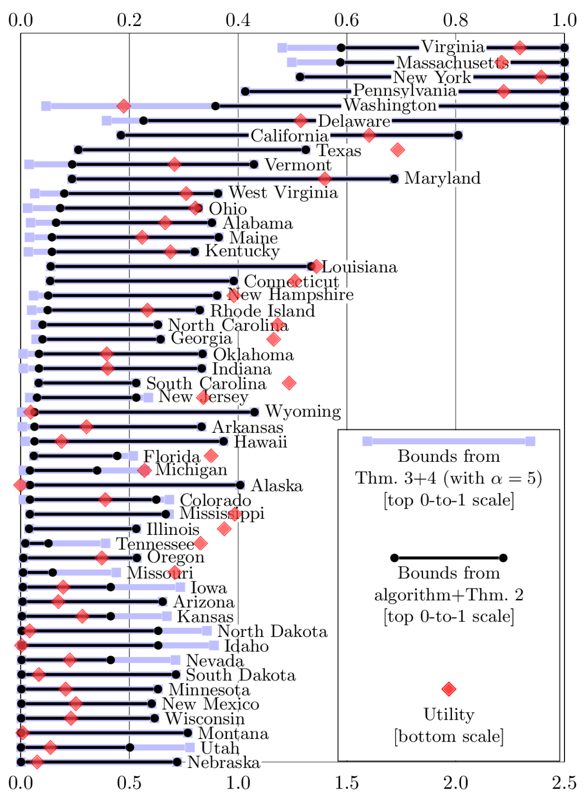
↓

apply propagation algorithm to Top-3 with $\alpha = 5$ (assume 15+ states considered on average)

↓

bounds on how often each state was considered





References

[1] Plackett. The analysis of permutations. *J R Stat Soc Ser C Appl Stat*, 1975.
[2] Luce. *Individual Choice behavior: A theoretical analysis*, Wiley 1959.
[3] Putnam, Ross, Soter, and Roediger. Collective Narcissism: Americans Exaggerate the Role of Their Home State in Appraising U.S. History *Psychological Science*, 2018.

Fruit icons from Icons8.