Bayesian Estimators As Voting Rules

Lirong Xia Rensselaer

UAI 2016

Top 250 movies

- “Complex voter weighting system”
  - Claimed to be accurate
    - a “true Bayesian estimate”
  - Claimed to be fair

1. The Shawshank Redemption (1994)	9.2
2. The Godfather (1972)	9.2
3. The Godfather: Part II (1974)	9.0
4. The Dark Knight (2008)	8.9

Different Voice

- Q: "This is *unfair* ! "
  - "That film / show has received awards, great reviews, commendations and deserves a much higher vote!"
  - My read: obviously strong candidates should win
- IMDB: "…only votes cast by IMDb users are counted. We do not delete or alter individual votes"

IMDb Votes/Ratings Top Frequently Asked Questions

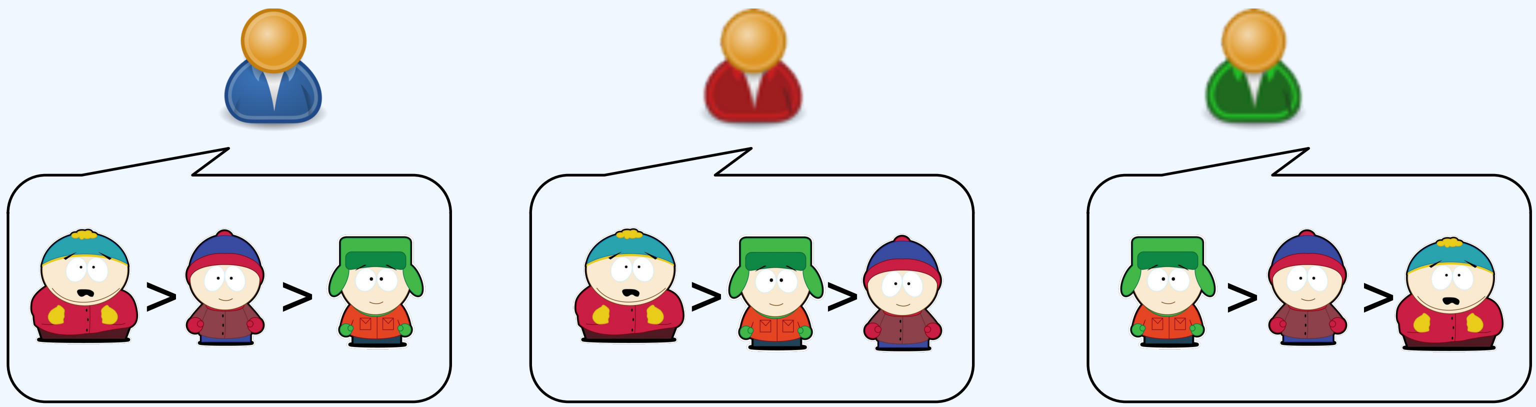
http://www.imdb.com/help/show_leaf?votestopfaq

This paper

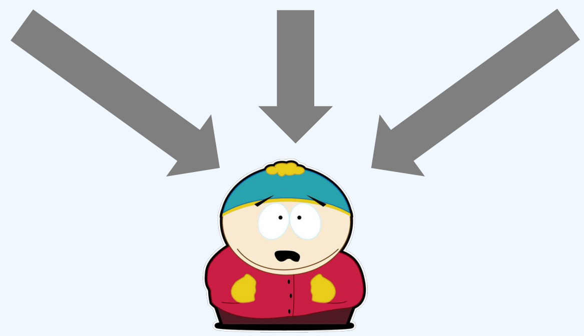
- Q1: How to measure fairness?
  - A: View them as voting rules
    - Evaluate by fairness axioms in social choice
- Q2: How can we design fair Bayesian estimators?
  - A: model + loss function [APX NIPS-14]

Who cares about both truth and fairness?

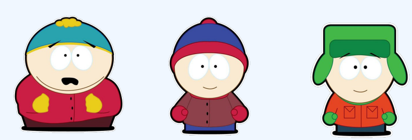
Social choice (rank aggregation)



mechanism: plurality rule

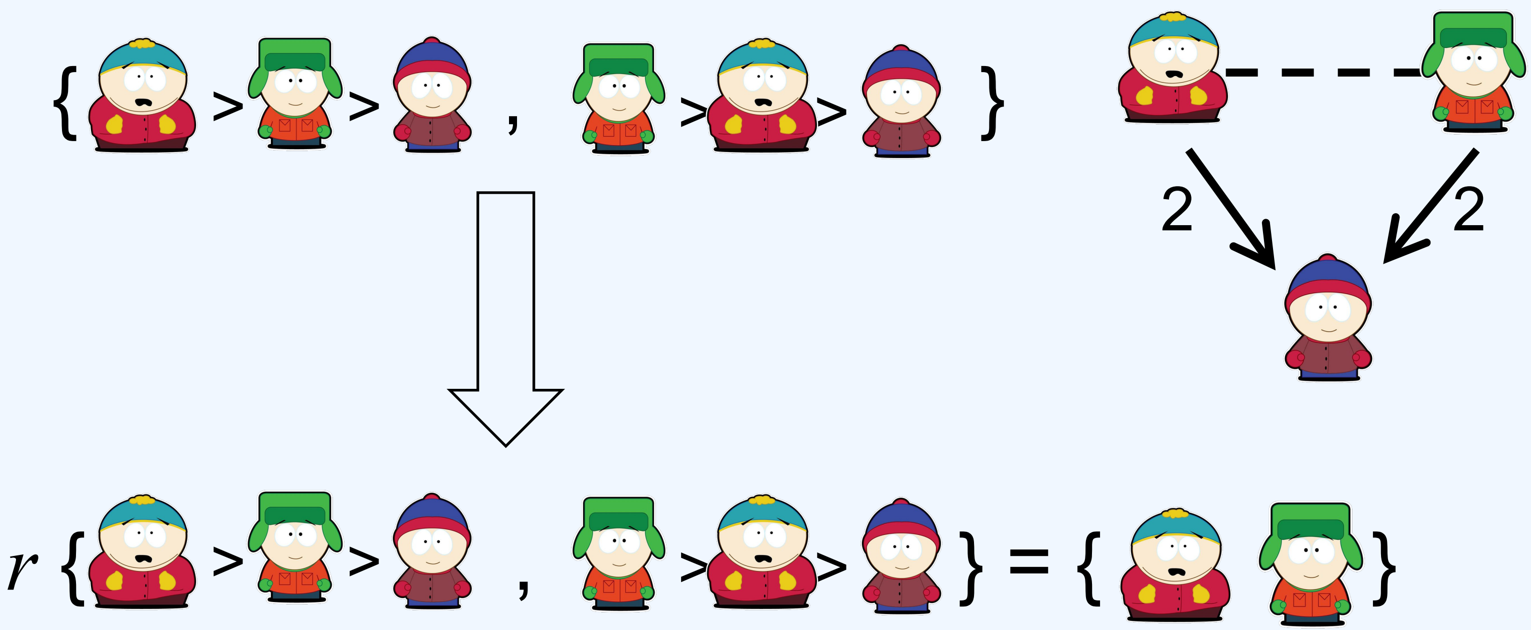


Decision

Candidates = {  } = Decision space

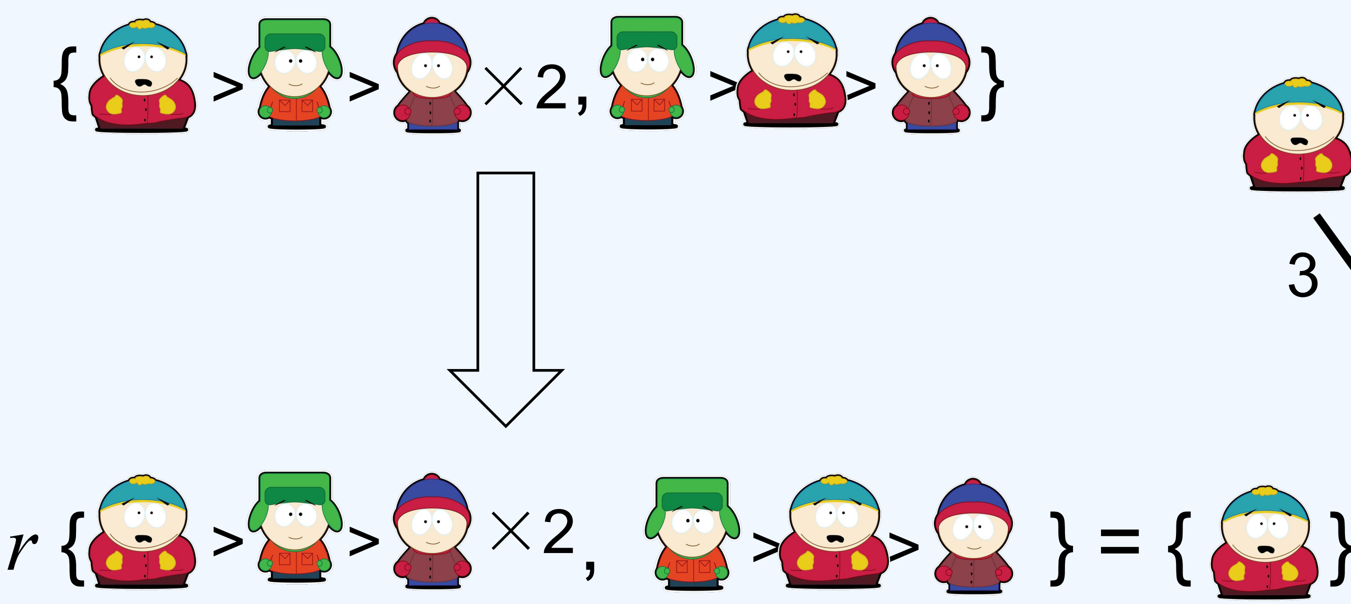
Measuring Fairness of Voting Rules with Ties

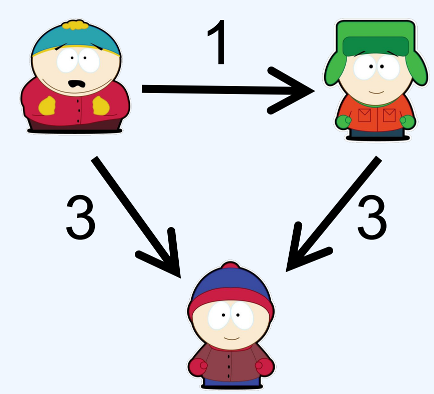
- Strict Condorcet criterion
  - Weak Condorcet winners (if exist) must win
  - Fairness for obviously strong candidates



Fairness Axiom: Condorcet Criterion

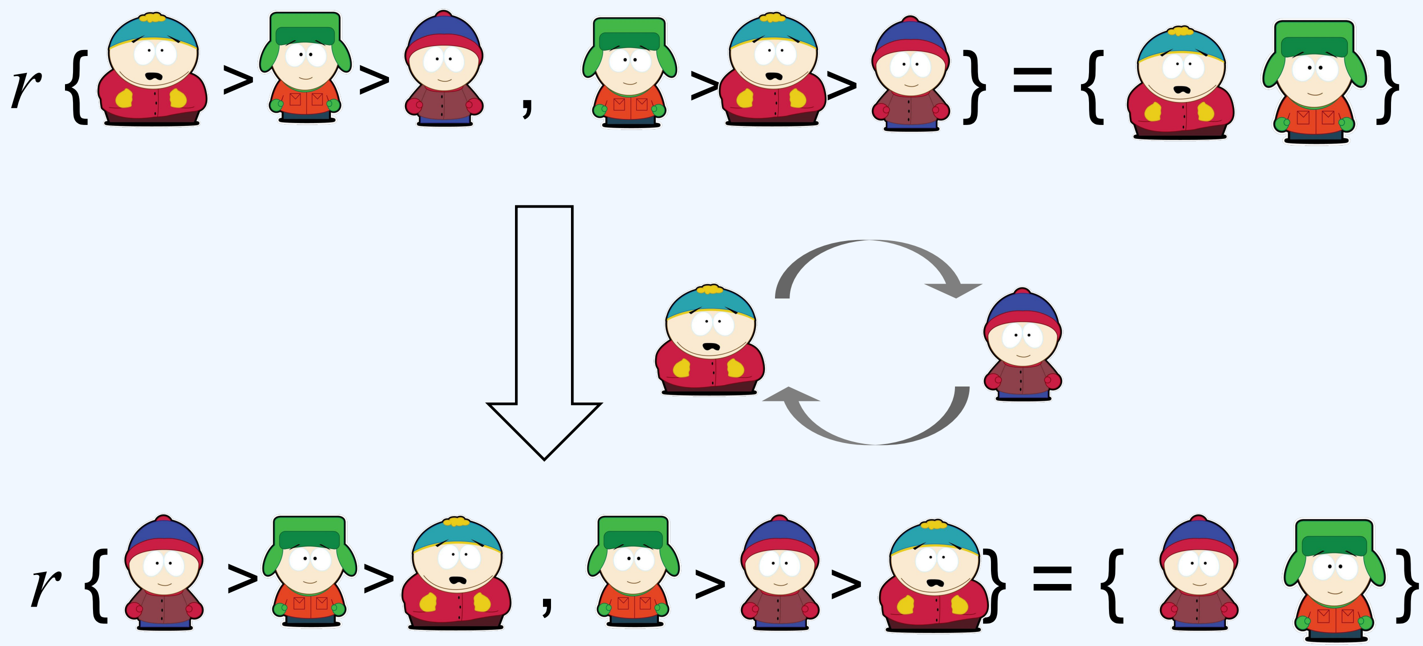
- (non-strict) Condorcet criterion
  - Condorcet winner (if exist) must win
  - Fairness for obviously strong candidates





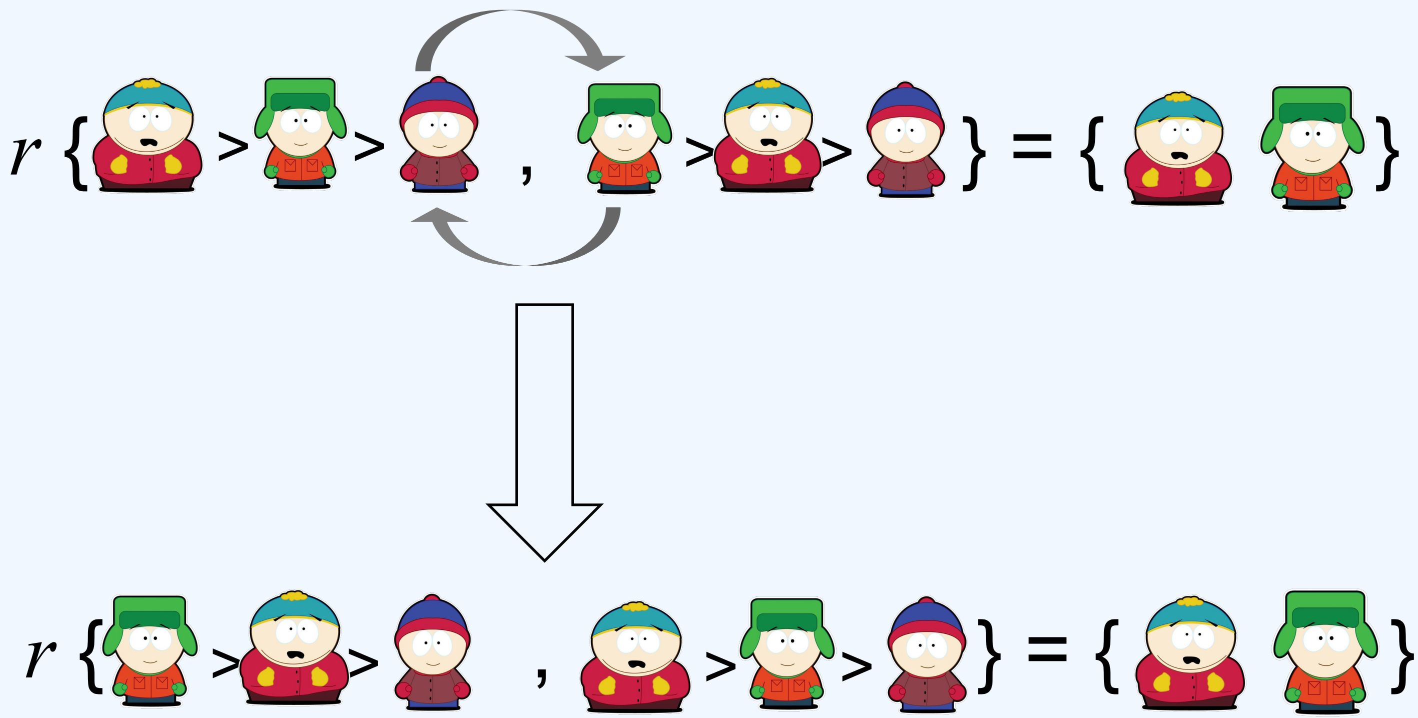
Fairness Axiom: Neutrality

- Neutrality
  - Fairness for candidates



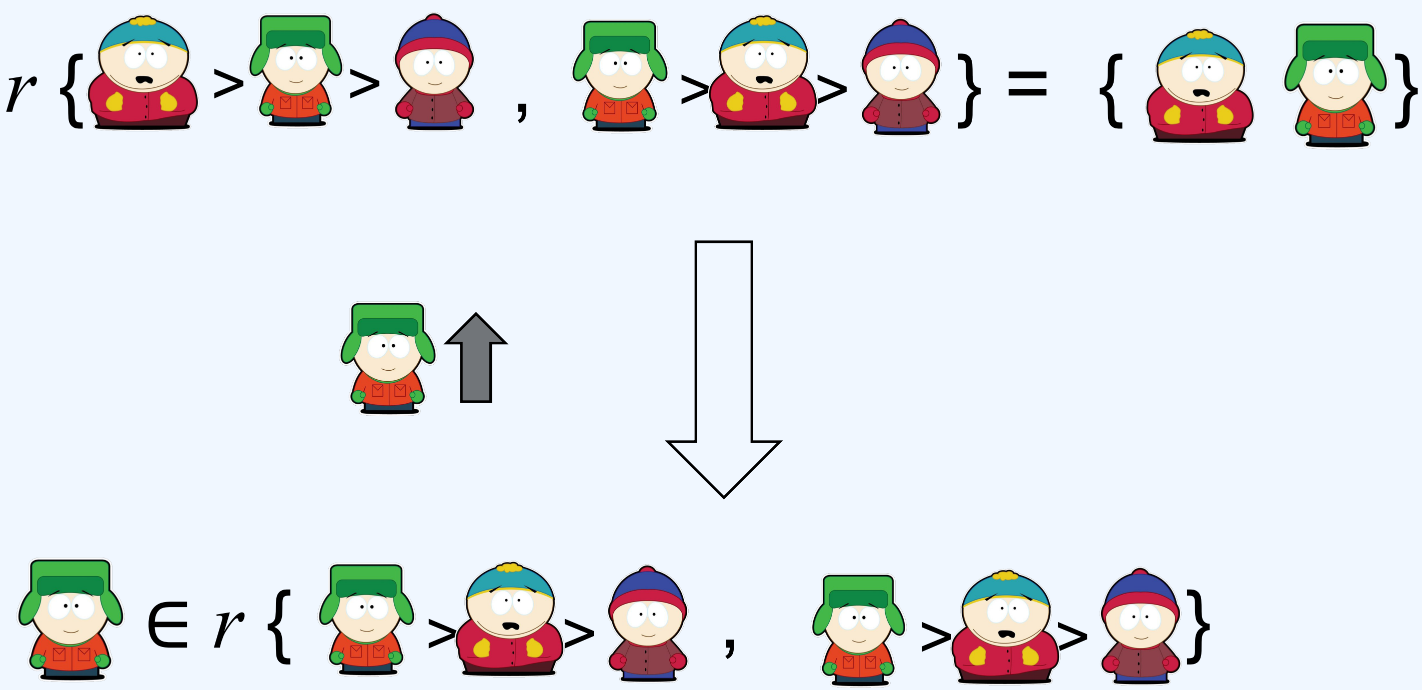
Fairness Axiom: Anonymity

- Anonymity
  - Fairness for voters



Fairness Axiom: Monotonicity

- Monotonicity
  - Weak form of strategy-proofness
  - Fairness for non-sophisticated voters



Bayesian estimators

Inputs

- statistical model + prior
- decision space
- loss function: $L(\theta, d) \in \mathbb{R}$

unknown ground truth

decision to make

Bayesian estimator

r : Data$\longrightarrow D$ with minimum Bayesian expected lost:

- $r(P)$ = argmin$_d$ $\mathrm{E}_{\theta|P} L(\theta, d)$

General results

- **Theorem**: Strict Condorcet

  No Bayesian estimator satisfies strict Condorcet criterion

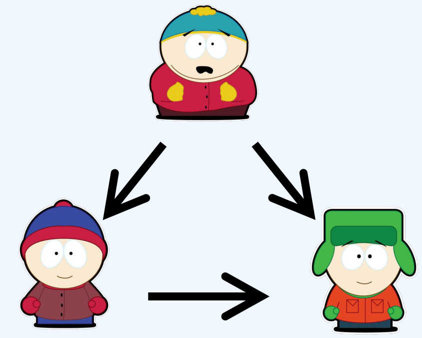
- **Theorem**: Neutrality

  Neutral Bayesian estimators

  = Bayesian estimators of "neutral" models

Other fairness axioms?

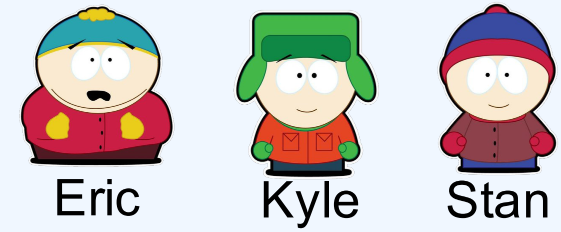
Mallows' model [Mallows-1957]



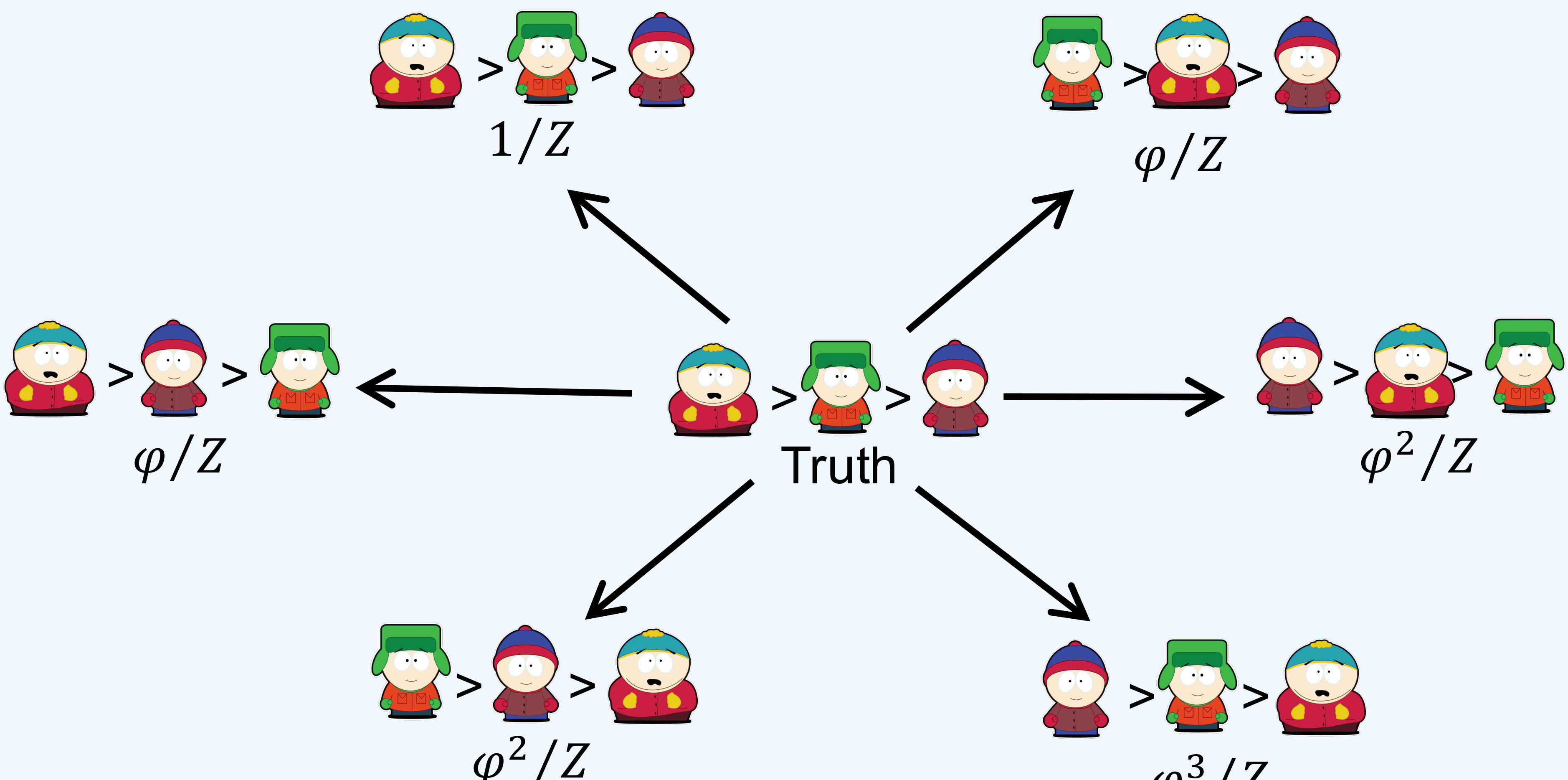
- Fixed dispersion φ <1
- Parameter space
  - all full rankings over candidates
- Sample space
  - i.i.d. generated full rankings
- Probabilities:

$$\Pr_W(V) \propto \varphi^{\mathrm{Kendall}(V,W)}$$

Example: Mallows for



➢ Probabilities: $Z = 1 + 2\varphi + 2\varphi^2 + \varphi^3$



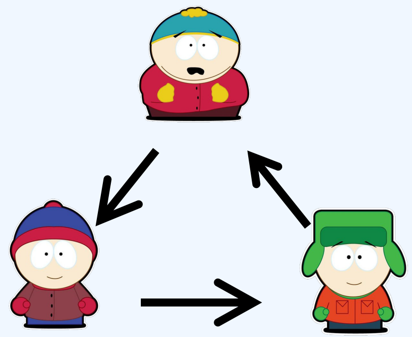
A Bayesian estimator

- $f_{\mathrm{Ma}}^{\mathrm{Top}}$ (Mallows with the top loss) [Young 1988]
  - Mallows' model
  - Decision: a set of winners
  - Loss: the top loss function
    - $L(W, a)$ =0 if a is top-ranked in W, otherwise it is 1
  - Uniform prior

Condorcet's model

[Condorcet-1785, Young-1988, ES UAI-14, APX NIPS-14]

- Fixed dispersion φ <1
- Parameter space
  - all binary relations over candidates



- Sample space
  - i.i.d. generated binary relations
- Probabilities:

$$\Pr_W(V) \propto \varphi^{\mathrm{Kendall}(V,W)}$$

A New Mechanism

- $f_{\mathrm{Co}}^{\mathrm{Borda}}$ (Condorcet with Borda loss)
  - Condorcet's model
  - Decision: a set of winners
  - Loss: the Borda loss function
    - $L(W, a)$ = # alternatives who beats a in W
  - Uniform prior

Our Results

<table>
<tr><th>Bayesian estimator</th><th>Anonymity, neutrality, monotonicity</th><th>Strict Condorcet</th><th>Condorcet</th><th>Complexity</th></tr>
<tr><td>$f_{\mathrm{Ma}}^{\mathrm{Top}}$</td><td rowspan="4">✔</td><td rowspan="4">✗</td><td>✔ iff $\frac{\varphi(1-\varphi^{m-1})}{1-\varphi} \leq 1$</td><td>NPH
[PRS UAI-12]</td></tr>
<tr><td>$f_{\mathrm{Co}}^{\mathrm{Borda}}$</td><td>✔ iff $\varphi \leq \frac{1}{m-1}$</td><td rowspan="3">P</td></tr>
<tr><td>f_{Pair}^{1}</td><td>✔ iff $\varphi \leq \frac{1}{m-1}$</td></tr>
<tr><td>f_{Pair}^{2}</td><td>✗</td></tr>
</table>

m: number of alternatives

f_{Pair}^{1} and f_{Pair}^{2} are BEs of a new model

Answering the Questions

- Q1: How to measure fairness?
  - A: View them as voting rules
    - Evaluate by fairness axioms in social choice
    - Impossibility theorem about strict Condorcet criterion
- Q2: How can we design fair Bayesian estimators?
  - A: model + loss function [APX NIPS-14]
    - New BEs that satisfy many desirable axioms

Future Work

- Other axioms
- Other types preferences
  - Partial orders, range voting (IMDb), probabilistic preferences…
- Other types of mechanisms
  - Probabilistic mechanisms

Thank you!