Learning Cooperative Solution Concepts From Voting Behavior

A Case Study on the Israeli Parliament

Lu Wei

Department of Computer Science
National University of Singapore

Jan 2020

Outline

Background

- Coalition formation games:
  - Mostly theoretical analysis
  - Most models require full information on player preference
  - Lack of large dataset with ground truth
- Clustering & community detection:
  - Data-driven analysis
  - Missing strategic behavior modelling

Data

- Israeli parliament (the Knesset) voting data
- Available since March 2017
- Contains 147 parliament members' votes on over 7500 bills in ~ 4 years
- Ground truth clusters: political party affiliations
  - 10 parties aligned along a left-right axis
  - Ideological agreement among the government (right) and the opposition (left) parties respectively

Probably Approximately Correct (PAC) Stability

Research Questions

- Can we use hedonic games to model real-world collaborative activities?
- How well does the outcome compare to ground truth?
- How well does the outcome compare to that of canonical clustering and community detection models?

Methodology

Game Theoretic Models

- Assumes hedonic preferences
  - Top responsive
  - Bottom responsive
  - Boolean

Methodology

Game Theoretic Models

- Assumes hedonic preferences
  - Top responsive
  - Bottom responsive
  - Boolean
- Two approaches:
  - Full-information models: construct complete preference profile, then derive core stable solution
  - PAC models: learn a probably stable solution directly from partial preference relations

Methodology

Game Theoretic Models

- Assumes hedonic preferences
  - Top responsive
  - Bottom responsive
  - Boolean
- Two approaches:
  - Full-information models: construct complete preference profile, then derive core stable solution
  - PAC models: learn a probably stable solution directly from partial preference relations
    - Simulate i.i.d.: sampling with replacement 3/4 of all bills
    - Repeat 50 times to check consistency of solution between runs
    - Select the "centroid" to represent model output: partition with minimum sum of information distance from other 49 partitions

Methodology

Comparisons

- Comparison machine learning models:
  - K-Means
  - Stochastic Block Model
- Comparing every hedonic & ML model output partition to ground truth party affiliations
  - Quantitatively: information theoretic measures
  - Qualitatively: political analysis

How different are two partitions, quantitatively?

Objectives

We want a measure that...

- has strong mathematical foundation: information theoretic measures
- is intuitive: satisfing metric property
  - non-negativity
  - symmetry
  - triangle inequality

How different are two partitions, quantitatively?

Definitions

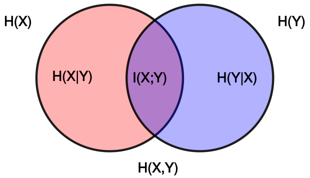

Figure: Venn diagram showing additive and subtractive relationships various information measures associated with correlated variables X and Y

- Entropy:
  $H(\pi) = -\sum_{j=1}^{J} \frac{|S_j|}{|N|} \log \frac{|S_j|}{|N|}$
- Conditional entropy: $H(\pi|\pi') = -\sum_{j=1}^{J} \sum_{k=1}^{K} \frac{|S_j \cap S'_k|}{|N|} \log \frac{|S_j \cap S'_k|/|N|}{|S_k|/|N|}$
- Mutual Information (MI):
  $I(\pi, \pi') = H(\pi) - H(\pi|\pi')$
- Variation of Information (VI):
  $VI(\pi, \pi') = H(\pi) + H(\pi') - 2I(\pi, \pi')$, metric!

How different are two partitions, quantitatively?

Baseline Values

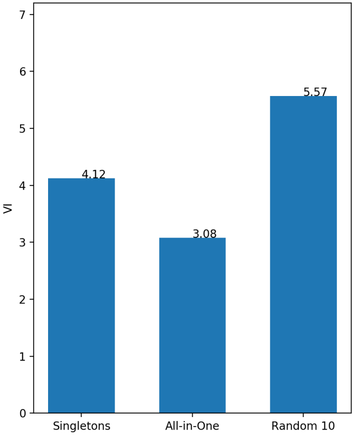

Figure: The Knesset partition baseline VI values

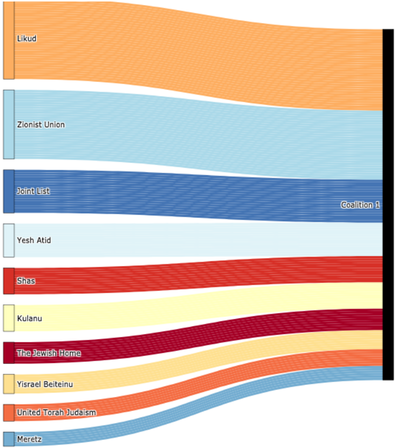

Figure: All-in-One partition of the Knesset

How different are two partitions, quantitatively?

Additional Measure: AMI

Adjusted Mutual Information (AMI):

$$AMI(\pi, \pi') = \frac{I(\pi, \pi') - E(I(\pi, \pi'))}{\max(H(\pi), H(\pi')) - E(I(\pi, \pi'))}$$

- Adjusted for chance
- Normalized: $[0, 1]$
- Not metric

Good for detecting "bad" (very different) partitions:

	Ajusted Mutual Information
Singletons	3e-14
All-in-One	-5e-16
Randome 10	0.007

Table: The Knesset partition baseline AMIs

Second Slide Title

- First item.

Second Slide Title

- First item.

Second Slide Title

- First item.
- Third item.

Second Slide Title

- First item.
- Third item.

Second Slide Title

- First item.
- Third item.
- Fifth item.

Second Slide Title

- First item.
- Third item.
- Fifth item. Extra text in the fifth item.

Main Theorem

Theorem 1

Theorem Statements. Example for citation [1].

Proof.

Proof of the theorem goes here. $\square$

Summary

- The **first main message** of your talk in one or two lines.
- The **second main message** of your talk in one or two lines.
- Perhaps a **third message**, but not more than that.
- Outlook
  - Something you haven't solved.
  - Something else you haven't solved.

Bibliography

[1] A. Author.
Handbook of Everything.
Some Press, 1990.

[2] S. Someone.
On this and that.
Journal of This and That, 2(1):50–100, 2000.