PrefLib: A Library for Preferences
HTTP://WWW.PREFLIB.ORG

Nicholas Mattei and Toby Walsh

NICTA and UNSW
Sydney, Australia
{nicholas.mattei,toby.walsh}@nicta.com.au

Abstract. We introduce PrefLib: A Library for Preferences; an online resource located at http://www.preflib.org. With the emergence of computational social choice and an increased awareness of the applicability of preference reasoning techniques to areas ranging from recommendation systems to kidney exchanges, the interest in preferences has never been higher. We hope to encourage the growth of all facets of preference reasoning by establishing a centralized repository of high quality data based around simple, delimited data formats. We detail the challenges of constructing such a repository, provide a survey of the initial release of the library, and invite the community to use and help expand PrefLib.

Keywords: Preferences, Computational Social Choice, Empirical Analysis.

1 Introduction

To date, research in computational social choice has been largely theoretical. There is, however, a growing realization of the limitations of a purely theoretical approach to computational questions in social choice. For example, worst-case results about the hardness of manipulation may not reflect the cost in practice to compute manipulations [22–24]. On the other hand, average-case analysis (e.g. [6, 16, 25]) may need to make additional assumptions, which can be rather simplistic and unrepresentative of preferences met in practice. Many such analyses assume that all preferences are equally likely; which is not supported by studies in behavioral social choice [15, 17] or studies on real data [12, 18, 21].

While our main interest is in computational voting there are other fields which fall in the area of preference handling and computational social choice including recommender systems [19], matching [8], and fair division problems [14]. These areas have found new and exciting application areas in modern life including matching kidney donors [5] and allocating students to seats in classrooms. There is a growing movement in the computational social choice community to identify and use **real** preference data to test algorithms and assumptions about voting systems (e.g. [10, 11, 13, 20]). To encourage and facilitate more empirical studies, we discuss building a library of preferences, PrefLib guided by our experiences building CSPLib [7].

Initiatives such as the UCI Machine Learning Repository [1] have fostered a greater sense of sharing and collaboration in the machine learning and data mining communities. The contents of the UCI database focus on a broad range of problems. We hope to

P. Perny, M. Pirlot, and A. Tsoukiàs (Eds.): ADT 2013, LNAI 8176, pp. 259–270, 2013.

provide a similar level of community, exposure, and sharing, with PREFLIB while taking to heart the lessons learned by other communities that have created and maintained shared tools and data.

2 Motivation

Whilst one of the prime motivations for building a preference library is to encourage and facilitate more empirical studies in computational social choice, there are some other related motivations.

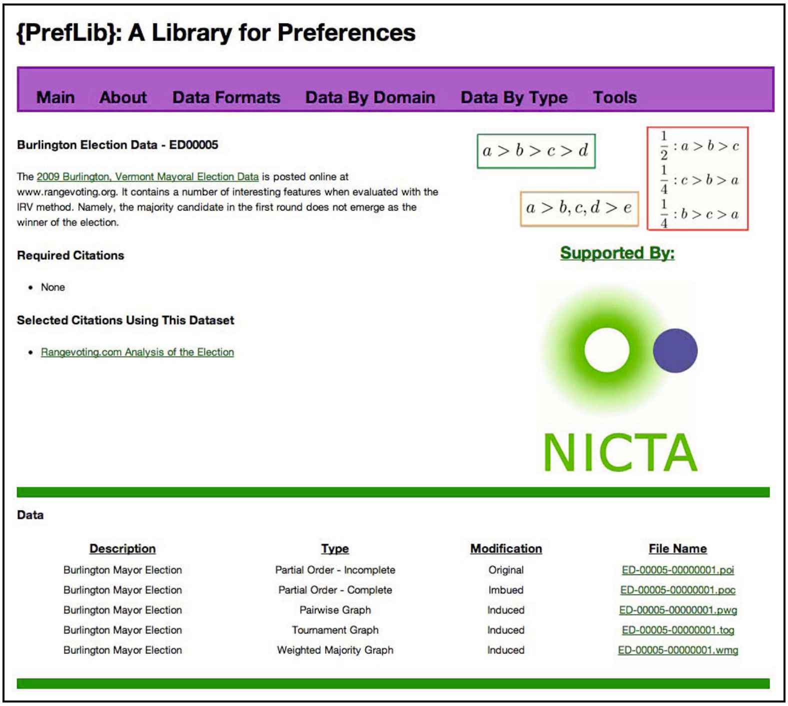


Fig. 1. An example page from `http://www.preflib.org` with data from the 2009 Burlington, Vermont mayoral elections

Benchmarking: A library can provide a common set of problems on which different research groups can quickly compare their algorithms. For example, we can compare the ability of different heuristics to compute solutions to NP-hard problems like finding a Borda manipulation [4].

Competitions: The Netflix Prize [2] demonstrate the benefits to research that a common set of preference data can have. While the large cash prize undoubtably had a strong impact, the Netflix data continues to be used extensively today despite the prize having been awarded.

Realism: As argued before, real world preference data could help direct the research community onto more practical computational issues in social choice. Representation and learning of complex preference models can happen more readily with a large corpus of preference data that is easily available.

Challenges: A benchmark library can be a forum for challenges that can help push the technology onto new heights. For instance, it can include open problems that may help drive research.

Insularity: The research community looking into computational social choice is rather insular: most people work on their own problems and their own data. A common problem library can encourage people to tackle a common set of problems, and help break down many barriers.

The construction of a benchmark library appears to be a common rite of passage for many research communities [1, 7]. For many reasons, it appears a good time for the computational social choice community and, generally, the preference handling community, to take this step.

3 Challenges

There are a number of challenges in building a preference library.

Variety: Preferences come in many shapes and forms. There are qualitative and quantitative preferences. There are voting preferences which might be simple plurality ballots, lists of approved candidates, lists of vetoes or complete rankings of the candidates. There are preferences for matching problems like hospital-resident problems and kidney exchanges. There are preferences over products in recommender system. There are temporal preferences for scheduling problems. When domains are large, there are combinatorial preferences which might be expressed using CP-nets or one of the many compact preference formalisms. While we wish to include all these in PrefLib it will be hard to find and post high-quality datasets spanning the entire range of preferences formalisms and domains.

Elicitation: Preferences are difficult to elicit. Users will only answer a limited number of questions about their preferences before they expect a system to start making good recommendations. In addition, users often have difficulty in articulating their true preferences and may not reply truthfully. These problems are well known and somewhat understood by other disciplines such as psychology. While we can learn lessons from these other disciplines, elicitation remains a key challenge in the preference handling community.

Modeling: Part of the challenge in social choice is modeling users' preferences. It is unlikely that your or my preferences are actually a CP-net [3] or even a linear order. These are just formalisms to approximate the complete preference functions we actually have. The existence of a preference library may distract attention from such important modeling issues.

Over-fitting: If the library is small, we run the risk of over-fitting. On the other hand, collecting a lot of preference data to avoid over-fitting may require considerable time and effort.

Privacy and Data Silos: Sharing of many datasets may be precluded or difficult for a variety of reasons. Medical and admissions data may need to be cleaned or anonymized in a suitable way before it can be shared. Additionally, some researchers may not want to share data in order to maintain exclusivity of a research topic. While we hope that all researchers understand the benefit of posting data openly and sharing, there are and will be bumps in the road.

Fortunately, none of these problems are insurmountable. In fact, the answer to many of these problems can be found in the community rallying around a common, open standard and library. With enough contributors we can ensure a large and rich cross section of problem instances, models, and elicitation procedures. With enough contributors the individual time investment will be minimal. And with a focus on sharing researchers will, hopefully, take careful consideration of their methods in order to create datasets.

Taking a page from the UCI Machine Learning Repository [1] we are maintaining a list of research publications that use individual datasets as well as research publications that should be cited along with the use of any particular dataset. By providing credit and exposure to researchers who give back to the community and creating a common resource for research *within* the community we hope that more groups will fully and publicly share their data.

4 Structure of the Library

PrefLib is currently divided into a four large sections according to overarching data type. The following list is not exhaustive, and is just a starting point for the library. For instance, it does not include preferences in fair division problems or preferences in facility location problems However, we expect that PrefLib will eventually grow to include such preferences.

Currently PrefLib holds over 2,000 datasets describing elections, ratings, and matchings. We have 100's of preferences from the Netflix Challenge, 100's of examples of matching data from kidney matching markets. We have several real elections including the 2007 Glasgow City Council elections, Mayoral Elections from the United States, and elections held in Dublin. We are still expanding PrefLib and we hope to bring more datasets online in the coming weeks and months.

Election Data: Election data includes data from real elections and other instances where rank orders are elicited from individuals. Currently we host a variety of rank order preference information with sources as varied as NASA spacecraft path selection to real ballots from mayoral elections in the United States.

Matching Data: Matching problems include two-sided markets (specifically stable marriage, hospital/resident and hospital/resident with couples problems), and one-sided markets (specifically room-mate and kidney exchange problems). We currently host synthetic data about kidney matching problems and real data related to university course selections.

Combinatorial Data: Large, combinatorial domains introduce interesting issues regarding representation. Combinatorial preferences subdivide into CP-nets, GAI-nets, and lexicographical preferences. Additionally we host single and multi-attribute rating data such as TripAdvisor data. Quantitative and qualitative multi-attribute rating data is of interest to researchers in the recommendation systems area and we hope to bring more datasets in this area online in the near future.

Optimization Data: Optimization problems subdivide into max-SAT, max-CSP, weighted-CSP and fuzzy-CSP problems.

Each dataset is posted in its original format as well as several easily induced (derived from) or imbued (data added) formats. For instance, for each set of rank ordered preference data we also include an imbued instance where each unranked candidate is placed tied, at the end, of each ranking. We also include an induced tournament graph for each set of rank ordered preference data. We have clearly marked each dataset as original, induced, or imbed, respectively. We encourage caution when drawing broad conclusions from studies on imbued or induced data, (see, e.g., [15, 17] for a discussion of potential pitfalls). However the data is interesting for testing of algorithmic results.

5 Preference Data

An important aspect of building a preference library is ensuring the preference data is in an easily accessible and computer readable form. Here we can learn from other domains. For instance, the propositional satisfiability community has a very successful library, SATLib which grew out of the Second DIMACS challenge in 1992 to 1993. The DIMACS format is widely accepted as the standard for Boolean formulae in CNF. Indeed, every satisfiability solver that we know about will read problems in the DIMACS format.

Why did DIMACS format become a standard? First, the format was in the right place at the right time. The format was proposed at the time that there was a lot of interest in developing new SAT solvers. There was therefore a very immediate need to compare solvers on a set of common benchmarks. Second, the format was quickly adopted by SATLib and by the semi-annual SAT competition. Third, the format is very simple. Each clause is a line in the input, made up of a sequence of positive and negative numbers terminated by a zero. It doesn't matter what computer language you write in, it takes just a few minutes to write a parser to read such problems.

Another successful format is the TPTP dataset for first order theorem proving. This is a slightly more complex format (but that is perhaps inevitable as first order problems are more complex to specify than purely propositional problems). The TPTP library includes a very useful tool, TPTP2X that converts TPTP problems into all the different formats used by the main first order theorem provers. This helps compensate for the greater complexity of the TPTP format. It means that users can quickly read problems into whatever theorem prover they might want to try out.

Based on these experiences, we use very simple formats for expressing preference data. We have attempted, as much as possible, to preserve a basic comma separated (CSV) format as is possible. This has several advantages including human readability and interoperability with outside data handling programs such as Excel, R, and Matlab. For example, candidates in an election are represented by the numbers 1 to m, and each preference ballot is represented by a permutation of these numbers in a comma separated format.

6 Datasets, Numbering and Tools

Our data comes from a variety of sources and locations. In general, we want data that is honest and comes from real decision makers regarding things that they care about and are incentivized to answer honestly. For example, while an anonymous surveys are good, there is no guarantee that respondents will respond truthfully (or something not completely random). Datasets that are derived from real elections, or real preference data (such as Netflix), or judging on real competitions, can have far more value than random surveys.

To understand the formatting and presentation of the data we present a full element of one of our datasets. This particular dataset provides a partial order over the 20 skaters in the women's 1998 world championships B group qualifier according to their ratings by 9 individual judges.

```
20
1,Maria Butyrskaya
2,Silvia Fontana
3,Vanessa Gusmeroli
4,Yankun Du
5,Anna Wenzel
6,Anna Rechnio
7,Olga Vassiljeva
8,Elena Liashenko
9,Rocia Salas
10,Tanja Szewczenko
11,Valeria Trifancova
12,Marta Andrade
13,Tatyana Malinina
14,Lucinda Ruh
15,Diana Poth
16,Mojca Kopac
17,Zuzana Paurova
18,Roxana Luca
19,Helena Pajovic
20,Yulia Lavrenchuk
9,9,9
1,1,6,20,8,13,10,3,15,12,2,17,14,16,5,4,11,7,19,9,18
1,1,6,8,10,20,3,13,14,2,15,16,17,7,5,12,11,4,18,19,9
1,1,6,3,13,8,10,20,15,2,17,12,5,7,14,16,11,4,19,18,9
1,1,6,10,13,20,3,8,2,14,15,16,17,{11,19},5,7,12,4,18,9
1,1,8,6,3,20,13,10,15,14,12,16,2,17,5,4,18,7,11,19,9
1,1,6,8,10,20,3,13,15,2,14,17,{12,16},4,5,7,11,19,18,9
1,1,6,13,8,20,10,15,3,17,5,2,16,12,7,4,14,11,18,19,9
1,1,6,13,8,20,10,3,16,15,2,17,4,14,5,12,7,11,9,19,18
1,1,6,10,8,13,3,20,15,2,14,12,17,5,16,7,4,11,19,18,9
```

The first line contains the number of candidates or items in this instances. The next set of lines are a number for each of the candidates and the real name or label for the candidate.

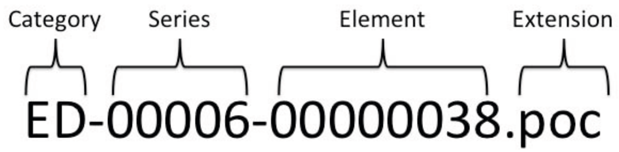
The first line under the list of candidates contains information about the number of voters. the first number is the number of actual ballots cast in the instance. The second number is the sum of the preference count (the number of preferences expressed). In most cases the number of ballots is the same as the sum of the vote preference count, except where for example, we have induced a relation like generating a pairwise graph from a set of linear orders. In this case we would have some number n of voters over m alternatives but we would have $\binom{n \cdot m}{2}$ as the sum of preferences since each voter expresses a relation between each pair of elements. The final number of this line is the number of unique preferences expressed.

The remaining lines in the file are the all of the format: count, preference list. The first element is the number of voters expressing the preference list. In the preference list each element is separated by a comma, and we close indifferent alternatives in { }.

In the example, each voter has selected skater 1, Maria Butyrskaya, as the best skater in the pool. Each unique order is indicated by a single line that is comma separated. This allows our data to be easily ported between different applications as it is delimited in a very simple manner.

6.1 Numbering

In order to make navigating particular datasets in PrefLib easier every individual datafile has a unique identifier which has a common numbering format. Below is the number for the woman's ice-skating world championship dataset shown above along with an explanation of the fields.



Category: Is a 2 letter category code; ED for election data, MD for matching data, CD for combinatorial data, and OD for optimization data.

Series: Is a 5 digit Series Code which specifies the source of the data. The Skate data shown above is number ED-00006.

Element: Is an 8 digit Element Number for each individual file of a particular extension. The example from the Skate data is 00000038, signifying that it is the 38th dataset from the Skate set with the same file extension.

Extension: Is a unique file extension to described the type of data in the file. The list of extensions is updated every time we obtain data in a new domain. For example, we use **soc** for datasets that are complete strict orders (all candidates are ranked with no ties between candidates) and **poc** for complete partial orders (all candidates are ranked but there are some ties between some candidates).

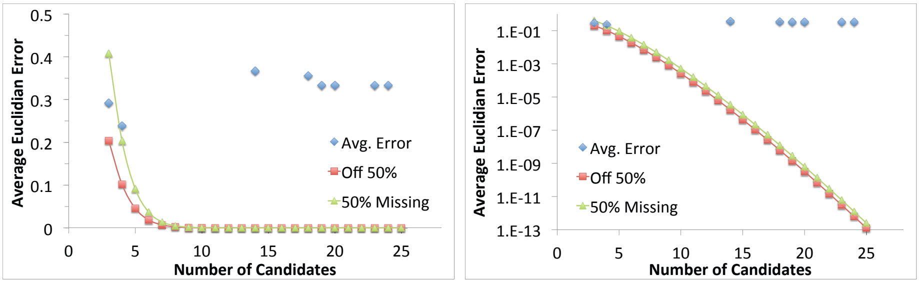


Fig. 2. Comparison of IC and the PREFLIB dataset. Both graphs show the average Euclidian Distance between the empirical distribution given some number of candidates in PREFLIB and IC. Additionally, we have plotted two hypothetical distributions: 50% MISSING assumes probability 0 for 50% of the possible strict orders and OFF 50% assumes each probability of observing a strict order is 50% different than the prediction made by IC. The left plot is linear while the right plot is a log scale.

6.2 Tools

In addition to the preference data on the site we plan to have a small set of tools available to the community. At this time we have no plans to create a monolithic tool chain like Weka [9]. All of the existing toolchain is written in Python3 and includes the ability to read, write, and process all of the data formats present on the site. Additionally, the toolchain includes functions to generate synthetic preferences according to a number of well studied preference cultures including the Impartial Culture, the Impartial Anonymous Culture, the Urn Model, and others [12, 21]. We look forward to adding support for other program languages and models in the future as we receive feedback on the requirements of the research community.

7 Distributions of Preferences

In collecting such a large and diverse set of preference data we hope researchers can begin to ask questions that were not possible before due to lack of data. A central question to the social choice community is testing the validity of generative models of preferences. In particular, we can start to look at the Impartial Culture (IC) assumption with more rigor than previously possible [12, 17, 21].

When looking at the data available in PREFLIB, one of the first observations that we can make is that research in computational social choice may need expand to generalizations of many current results for strict orders to strict orders that are not complete rankings. The current version of PREFLIB contains 220 instances of complete strict order ranking data. However, it also contains 118 incomplete strict order (SOI) ranking data. In fact, the SOI data contains many instances of actual elections from Dublin, Glasgow, and trade unions in the EU. These instances contain, on average, 80% incomplete preference relations, with many votes only expressing a top alternative. There are a number of hazards associated with dropping the incomplete votes or randomly

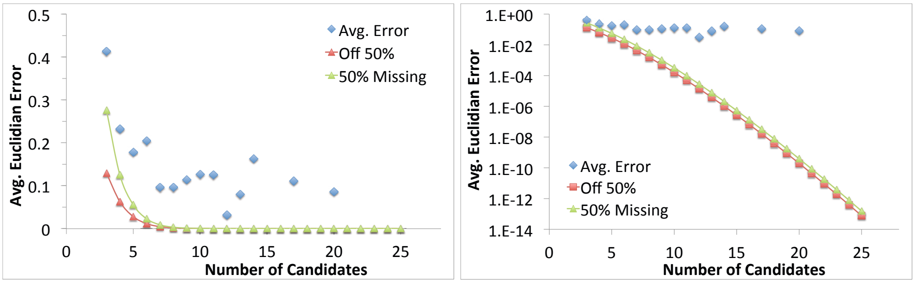


Fig. 3. Comparison of IIC and the PREFLIB dataset. Both graphs show the average Euclidian Distance between the empirical distribution given some number of candidates in PREFLIB and IIC. Additionally, we have plotted two hypothetical distributions: 50% MISSING assumes probability 0 for 50% of the possible strict orders and OFF 50% assumes each probability of observing a strict order is 50% different than the prediction made by IIC. The left plot is linear while the right plot is a log scale.

extending them as these procedures introduce many assumptions about the underlying data [15]. Without generalizing our thinking to include SOI data we may leave real-world behaviors unstudied in computational social choice.

Figures 2, 3, and 4 we compare the Impartial Culture with the SOI and SOC data currently in the PREFLIB data base. While IC is well defined for complete strict orders, we needed a suitable generalization to incomplete strict orders. For this, we make as few assumptions as possible to create the Incomplete-Impartial Culture (IIC): every ordering (including truncated orderings) has an equal probability of occurring. The probability of observing a given ranking r for n candidates is:

$$Pr(r) = \left(\sum_{i=1}^{n} i! \binom{n}{i} \right)^{-1}.$$

We follow the same procedures as Tideman and Plassmann [21] and Mattei et al. [11, 12] in our study. In order to compare an empirical distribution to a generative one we reorder the empirical distribution such that the preference order of the most frequent vote is the labeling for the candidates (we use the most frequent complete order for relabeling in IIC). This procedure ensures distributions from different empirical scenarios are comparable by giving them a uniform shape. Once we have done this we compute the Euclidian Distance between the empirical distribution and IC or IIC, respectively. We call this number the Euclidian Error and it gives us an idea of how near or far two distributions are from each other.

Figure 2 and 3 shows the error of the empirical distribution on linear and log plots for a given number of candidates. We only plot points where PREFLIB has 2 or more unique datasets. Additionally, we have plotted two hypothetical distributions: 50% MISSING assumes probability 0 for 50% of the possible strict orders and OFF 50% assumes each

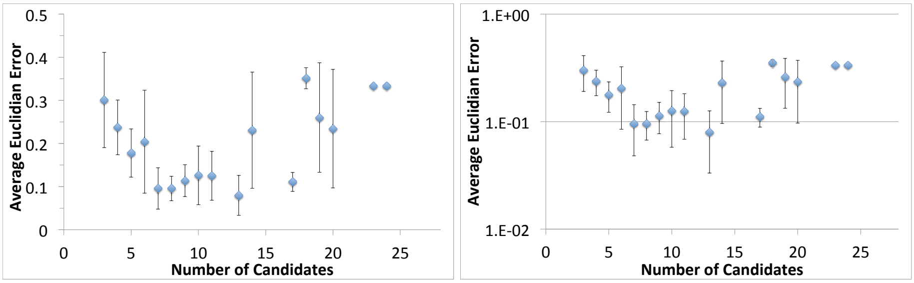


Fig. 4. Comparison of IC/IIC and the PrefLib dataset. Both graphs show the average Euclidian Distance between the empirical distribution given some number of candidates in PrefLib and IC/IIC respectively, including standard error bars. The left plot is linear while the right plot is a log scale.

probability of observing a strict order is 50% different than the prediction made by IC/IIC.

We have plotted these additional distributions to give some perspective on just how different IC/IIC is from our empirical distribution. Most would agree that a distribution that is *always* off by 50% to be a fairly poor estimate. When we look at the SOI and SOC data we see that this *bad* distribution is significantly closer to IC/IIC than the empirical distribution found in PrefLib. These distributions show us just how much IC/IIC diverges with SOI: falling completely outside of the projected curve for either of the "bad" distributions once we have more than 4 or 5 candidates.

Figure 4 shows the SOC and SOI combined average difference and standard error on linear and log plots. Here we can see that, in general, for a given number of candidates, the empirical distribution in PrefLib is (1) extremely variable and (2) a reasonable distance from IC/IIC. Even with all the data we have collected so far, we are under-sampling. Thus making it unwise to draw too many conclusions from this data. While we are still collecting data we see that what we have currently does not support the simplistic IC/IIC assumptions.

8 How to Contribute

One way to increase the usefulness of PrefLib is to build a community around the datasets. What we have presented here is only the beginning; we hope that interested researchers will contact us with donations of data or pointers to datasets that we may have missed while constructing the first version of the site.

In order to contribute data please contact Nicholas Mattei; we host all the data so that it is available in a central location. We work collaboratively with all our data donors to convert the data into a simple, CSV-like format. We post links and citations to any donor suggested papers or external websites as well as citations that are requested to accompany the use of particular datasets. We want to make sure that donors who take the time and effort to work with us on posting datasets receive the recognition they deserve for taking the time to support the community.

We make no claims on ownership of data on the website. While we have worked hard to only include high quality, accurate datasets we make no explicit warranties or guarantees about the data and distribute the data "as is."

9 Conclusion

We have introduced the first version of PrefLib and an associated toolchain for working with preference data. We hope to provide a ongoing and valuable service to not only the computational social choice community, but the preference handling and reasoning community writ large. To support this mission we must have the support and donations of the research community. We encourage anyone with interesting datasets to contact us; we will work with you on encoding and hosting interesting data. Please help us to grow the empirical side of of preference handling.

Acknowledgements. The authors are supported by the Australian Government's Department of Broadband, Communications and the Digital Economy, The Australian Research Council and the Asian Office of Aerospace Research and Development through grant AOARD-124056.

References

1. Bache, K., Lichman, M.: UCI machine learning repository, University of California, Irvine, School of Information and Computer Sciences (2013), `http://archive.ics.uci.edu/ml`
2. Bennett, J., Lanning, S.: The Netflix prize. In: Proceedings of the KDD Cup and Workshop (2007)
3. Boutilier, C., Brafman, R.I., Domshlak, C., Hoos, H.H., Poole, D.: CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. Journal of Artificial Intelligence Research (JAIR) 21, 135–191 (2004)
4. Davies, J., Katsirelos, G., Narodytska, N., Walsh, T.: Complexity of and algorithms for Borda manipulation. In: Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI 2011), pp. 657–662. AAAI Press (2011)
5. Dickerson, J.P., Procaccia, A.D., Sandholm, T.: Optimizing kidney exchange with transplant chains: Theory and reality. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012), pp. 711–718 (2012)
6. Friedgut, E., Kalai, G., Nisan, N.: Elections can be manipulated often. In: Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2008), pp. 243–249. IEEE Computer Society Press (2008)
7. Gent, I.P., Walsh, T.: CSPlib: A benchmark library for constraints. In: Jaffar, J. (ed.) CP 1999. LNCS, vol. 1713, pp. 480–481. Springer, Heidelberg (1999)
8. Gusfield, D., Irving, R.W.: The Stable Marriage Problem: Structure and Algorithms. MIT Press, Cambridge (1989)
9. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1), 10–18 (2009)
10. Lu, T., Boutilier, C.: Robust approximation and incremental elicitation in voting protocols. In: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, pp. 287–293. AAAI Press (2011)

11. Mattei, N.: Empirical evaluation of voting rules with strictly ordered preference data. In: Brafman, R. (ed.) ADT 2011. LNCS, vol. 6992, pp. 165–177. Springer, Heidelberg (2011)
12. Mattei, N.: Decision Making Under Uncertainty: Theoretical and Empirical Results on Social Choice, Manipulation, and Bribery. Ph.D. thesis, University of Kentucky (2012)
13. Mattei, N., Forshee, J., Goldsmith, J.: An empirical study of voting rules and manipulation with large datasets. In: Fourth International Workshop on Computational Social Choice (COMSOC 2012). Springer (2012)
14. Moulin, H.: Fair Division and Collective Welfare. The MIT Press (2004)
15. Popova, A., Regenwetter, M., Mattei, N.: A behavioral perspective on social choice. Annals of Mathematics and Artificial Intelligence (to appear, 2013)
16. Procaccia, A.D., Rosenschein, J.S.: Average-case tractability of manipulation in voting via the fraction of manipulators. In: Proceedings of 6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007), pp. 718–720. IFAAMAS (2007)
17. Regenwetter, M., Grogman, B., Marley, A.A.J., Testlin, I.M.: Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications. Cambridge Univ. Press (2006)
18. Regenwetter, M., Kim, A., Kantor, A., Ho, M.H.R.: The unexpected empirical consensus among consensus methods. Psychological Science 18(7), 629–635 (2007)
19. Ricci, F., Rokach, L., Shapira, B.: Introduction to recommender systems handbook. In: Recommender Systems Handbook, pp. 1–35. Springer (2011)
20. Skowron, P., Faliszewski, P., Slinko, A.: Achieving fully proportional representation is easy in practice. In: Proceedings of 12th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), pp. 399–406 (2013)
21. Tideman, T.N., Plassmann, F.: Modeling the outcomes of vote-casting in actual elections. In: Electoral Systems, pp. 217–251. Springer (2012)
22. Walsh, T.: Where are the really hard manipulation problems? The phase transition in manipulating the veto rule. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009), pp. 324–329 (2009)
23. Walsh, T.: An empirical study of the manipulability of single transferable voting. In: Proceedings of the 19th European Conference on Artificial Intelligence (ECAI 2010). Frontiers in Artificial Intelligence and Applications, vol. 215, pp. 257–262. IOS Press (2010)
24. Walsh, T.: Where are the hard manipulation problems? Journal of Artificial Intelligence Research 42, 1–39 (2011)
25. Xia, L., Conitzer, V.: Generalized scoring rules and the frequency of coalitional manipulability. In: EC 2008: Proceedings of the 9th ACM conference on Electronic Commerce, pp. 109–118. ACM (2008)