PREFLIB and Empirical Testing in Computational Social Choice

Nicholas Mattei NICTA and UNSW



Australian Government

NICTA

Department of Broadband, Communications and the Digital Economy

Australian Research Council















Queensland Government







From imagination to impact

PrefLib: A Library for Preferences



- Many research communities have libraries, datasets, and tool chains that are standard and widely used.
- Preference handling and computational social choice have largely centered around theoretical results.
- We have collected datasets and tools to establish PREFLIB, a library of preference data, as a service to the wider community.



- 1. PrefLib: What it is and why we made it.
- 2. A detailed look at the data within PrefLib.
- 3. An overview of PrefLib:Tools and what they can do for you.

Preference Aggregation





- Problems arise when groups of agents (humans and/or computers) need to make a collective decision.
- How do we aggregate individual (possibly conflicting) preference profiles into a collective preference profile?



Welcome to the UC Irvine Machine Learning Repository!

We currently maintain 239 data sets as a service to the machine learning community. You may view all data sets through our searchable interface. Our old web site is still available, for those who prefer the old format. For a general overview of the Repository, please visit our <u>About page</u>. For information about citing data sets in publications, please read our <u>citation policy</u>. If you wish to donate a data set, please consult our <u>donation policy</u>. For any other questions, feel free to <u>contact the Repository librarians</u>. We have also set up a <u>mirror site</u> for the Repository.









Software

Publications

People Related

Weka 3: Data Mining Software in Java

Book

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.

Found only on the islands of New Zealand, the Weka is a flightless bird with an inquisitive nature. The name is pronounced like **this**, and the bird sounds like **this**.

Weka is open source software issued under the GNU General Public License.

Project

Getting started

Further information

- Requirements
- Download
- Documentation
- FAQ
- Getting Help

- Citing Weka
- Datasets
- Related Projects
- Miscellaneous Code
- Other Literature

Developers

- Development
- History
- Subversion
- Contributors

Nicholas Mattei NICTA and UNSW

Challenges

- Variety
 - We need examples from a large suite of domains where preference handling and decision making research happens.
- Elicitation
 - How do we collect this data how do we ensure the quality of the data we receive?
- Modeling
 - What are the correct formalisms to represent preferences?





Challenges

- Over-fitting
 - How do we prevent becoming too focused on metrics that can be measured from our datasets?
- Privacy and Information Silos
 - Some data cannot be shared, some data is not shared because it may create a competitive advantage – how do we convince people to share?





{PrefLib}: A Library for Preferences

Main About Papers Data Formats Data By Domain Data By Type Tools

A reference library of preference data and links assembled by <u>Nicholas Mattei</u> and <u>Toby</u> <u>Walsh</u>. This site and library is proudly supported by the <u>Optimization Research Group</u> at <u>NICTA</u>. We currently house over 3,000 datasets for use by the community.

We want to provide a comprehensive resource for the multuple research communities that deal with preferences, including computational social choice, recommender systems, data mining, machine learning, and combinatorial optimization, to name just a few.

Please see the <u>about</u> page for information about the site, contacting us, and our citation policy. We rely on the support of the community in order to grow the usefulness of this site. To contribute, please contact <u>Nicholas Mattei</u> at: nicholas**{dot}**mattei@nicta.com.au



Supported By:



/ Dec. 10, 2013:

We are hosting a workshop at <u>AAMAS 2014</u> on <u>Exploring Beyond</u> the Worst Case in Computational Social Choice. Nick will give a talk about PrefLib! Please consider joining us in Paris in the coming year.

Nov. 6, 2013:

The first release of the tool suite is now available on the <u>Tools</u> page. Python3 scripts to read, write, and generate preference data in our formats!

Sept. 3, 2013:

A big update today brings us over 3000 datasets hosted on the site with a full data archive over 7 GB!

We have also added a <u>Thanks!!</u> section to recognize those individuals who have helped make PrefLib possible.

/ July 1, 2013:

Our paper has been accepted to <u>2013 Conference on Algorithmic</u> <u>Decision Theory</u>. We have also had several new donated datasets which have been parsed and posted.

Links

- <u>UC Irvine Machine Learning Repository</u>
- <u>University of Minnesota GroupLens Data</u> Sets
- CSPLib: A Problem Library for Constraints
- Microsoft Learning to Rank Datasets
- SATLib: The Satisfiability Library
- Preference-Learning.org
- <u>Toshihiro Kamishima's Sushi Preference</u> <u>Dataset</u>
- MAX-SAT Evaluations and Datasets





Visits by Country





Types of Data



- We divide our data into 4 broad categories:
 - Election Data
 - Matching Data
 - Combinatorial Data
 - Optimization Data
- Simple data formats, mostly derived from comma separated lists.
 - As SATLib showed, this format is easy to use across software and disciplines.

Elections, rank-ordered preferences partial-ord

Election Data

preferences, partial-orders, sports competitions, tournaments, majority graphs.

a > b > c > d

 Datasets ranging from NASA satellite path selection to international skating competitions and real elections



Sample Data Format



11 Number of Candidates 1.Australia 2.Braille **3**,Brush Strokes 4, Exponential 5,College Candidate List 6, Graph Coloring 7.Red 8,Simple Number of Voters, 9,Star Trek 10,TSP Number of Orders, 11,VRP Number of Unique Orders 30,30,30 1,10,6,7,8,11,5,3,2,1,9,4 1,1,10,11,9,6,7,3,5,8,2,4 1,11,10,1,3,6,8,5,7,4,9,2 **Comma Separated Orders** 1,10,3,5,8,11,1,6,2,4,7,9



- 14 Unique Data Sources
 - Irish Elections, Scottish Elections, Course Selection, Skate Rankings, Movies, Sushi
- 470+ different election instances with strict rankings
 - 229 Contain only complete rankings.
 - 243 Contain some incomplete rankings.
- 381 different election instances where rankings have indifference

Voting Datasets





Voting Datasets



Voting Datasets





Models – Strict (SOC)





Models – Strict (SOI)







JAWS



Models – Partial (POC)















NICTA Copyright 2014

From imagination to impact

Models – Incomplete (POI)















- Irish Election Data:
 - 5% submitted complete ballots for Dublin North.
 - 12% for Dublin West.
- APA Election Data
 - Chamberlin's original data had over 65% incomplete ballots over 5 candidates.
- ANES Thermometer Rankings
 - Takes ratings and turns them into rankings, breaking ties randomly.
- Sushi Dataset
 - Incomplete survey's are discarded (sample bias, not incentivized).

Problems with Extensions



- Behavioral aspects of individuals can have substantial impacts on the resulting computational problems
 - Youtube's dropping of the star ratings system...
 - Single-peakedness...
- As we move forward pay particular attention to the domain in which we wish to deploy our results.
- We must be acutely aware of model dependence.

Models - Agnostic











Models - Pessimistic















Models – Anchor and Adjust





NICTA Copyright 2014

From imagination to impact

Some Extensions We Provide



- All have been turned into pairwise majority graphs, weighted pairwise majority, and tournament graphs.
- We have not posted all the different types of model extensions possible – we want these extensions to be explicit.
- Our goal is to present the data and preserve the context for future researchers avoid over-fitting!

Combinatorial Data

- CP-nets, UAInets, single and multi-attribute rating data.
- Currently we have a large collection of data scraped from TripAdvisor.





0 BARREL LUNDERSTANDING THE SKU BOOM LJERSEY'S FLYING FISI **Beeradvocate**

> Nicholas Mattei NICTA and UNSW

Audience Question

- New Category?
 - Rating / Multi-Dimensional **Preference Data?**

🦂 Online Reviews

amazon

Name	Туре	Number of items	Descr
BeerAdvocate	Beer reviews	1,586,259 beer reviews	Beer re BeerAc
RateBeer	Beer reviews	2,924,127 beer reviews	Beer re
CellarTracker	Wine reviews	2,025,995 wine reviews	Wine n CellarT
Amazon reviews	Amazon reviews (all categories)	34,686,770 product reviews	Review
Fine Foods	Food reviews	568,454 food reviews	Food n
Movies	Movie reviews	7,911,684 movie reviews	Movie











Nicholas Mattei NICTA and UNSW

Matching Data

- Kidney donor matching, runway slots, student seats in classrooms, student housing.
- Currently we have synthetic kidney matching data and student preferences over courses.









NICTA Copyright 2014

Nicholas Mattei NICTA and UNSW

Optimization Data

- Max-SAT, Max-CSP, weighted-CSP, TSP-games, etc.
- We lack good data here!!





Optimization Data



Max-SAT, Max-CSP. Max-SAT 2014

Ninth Max-SAT Evaluation

Introduction

The Ninth Evaluation of Max-SAT Solvers (Max-SAT-2014) is organized as an affiliated event of the 17th International Conference on Theory and Applications of Satisfiability Testing (SAT-2014).

The objective of the evaluation is assessing the state of the art in the field of Max-SAT solvers, as well as creating a collection of publicly available Max-SAT benchmark instances.



Х

Tools

 We have released version 0.1 of our tool suite – read, write, and convert between all the election data instances and matching data instances.

• I'll show some features of 0.2 which will accompany a new data update around 1 June.

- PreflibUtils.py: Main library of with functions for reading, writing, printing, and manipulating the data files.
- GenProfiles.py: Main file for generating preferences. Thanks to Andreas Pfandler for help with updates!
 - Impartial Culture Impartial Anonymous Culture -Polya-Eggenberger Urn Models - Mallows Mixtures -Single Peaked Impartial Culture
- DomainRestrictions.py: Test for various Domain restrictions including Single Peakedness.



The Impartial Culture assumes that the probability of observing any order is uniform.

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
1/6	1/6	1/6	1/6	1/6	1/6



- The Impartial Anonymous Culture assumes that the probability of observing any distribution over orders is equally likely.
- That is, any vector where X1+X2+X3+X4+X5+X6 = 1.0

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
X1	X2	X3	X4	X5	X6



• The **Urn Model** starts with a bag with N! elements in it.

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
1/6	1/6	1/6	1/6	1/6	1/6

• We draw a vote from it, and replace it with some fixed number of copies of the vote.



 The Urn Model starts with a bag with N! elements in it.

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
2/7	1/7	1/7	1/7	1/7	1/7

• We draw a vote from it, and replace it with some fixed number of copies of the vote.

- Replace with 1 extra copy.



- Mallows Models are parameterized by a reference ranking and a dispersion parameter.
- The probability of observing the non-reference ranking is proportional to the normalized value of the dispersion parameter (\$\phi\$) raised to the swap distance.

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
0	1	1	2	2	3

Statistical Cultures

- $P(C > B > A) = \frac{1}{X} \phi^{KT(A > B > C, C > B > A)}$ With $X = 1 \cdot (1 + \phi) \cdot (1 + \phi + \phi^2) \dots (1 + \dots + \phi^{m-1})$
- So this means that as $\phi \to 1$ we get IC
- And as $\phi \to 0$ we get a unanimous distribution.

• So X = 2.625,
$$\phi$$
 = 0.5

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
0.381	0.1905	0.1905	0.095	0.095	0.04762

NICTA



- Single Peaked Impartial Culture draws uniformly at random from the set of orders that are consistent with a given social axis.
- If the axis is A > B > C then we have positive probability only on the $2^{m-1} = 2^2 = 4$ consistent orders.

A > B > C	A > C > B	B > A > C	B > C > A	C > A > B	C > B > A
1/4	0	1/4	1/4	0	1/4

MatPlotLib and PyLab

- Available as a combined distribution with Numpy and other libraries at <u>http://scipy.org/</u>
- Easy to use and intuitive way to make graphs and do end to end experiments.
- Python is magic...





Demo Time!







- Using the IC model and IC generalized to partial rankings we asked how much PrefLib looks like generative models.
 - Should we define a discrete probability distribution called the PrefLib Culture?
- For comparison sake we randomly create additional datasets:
 - 50% Off: Each observed probability is off from IC by 50%. For 3 candidates: P(A > B > C) = 0.25 instead of 0.16667.
 - 50% Missing: There is probability 0 of predicting 50% of the observations. P(A > B > C) = 0, P(C > B > A) = 0.333.

Comparing IC – Complete Orders



IC – Complete Orders – Log Scale



Comparing IC – Incomplete Orders



IC – Incomplete Orders – Log Scale



All Sets – Error by Candidate Count





- Now that we have more data, what can we do in decision theory that we couldn't before?
 - Learning the preferences of one or multiple users in ordinal environments?
 - Learning and modeling preference drift over time.
 - Extending preferences from incomplete to complete orders.

Thanks



- Everyone who has participated and donated data or their time to the project so far:
 - Robert Bredereck
 - John Dickerson
 - Carleton Coffrin
 - Piotr Faliszewski
 - Toshihiro Kamishima
 - Jeffrey O'Neill
 - Andreas Pfandler
 - Florenz Plassmann
 - Nicolaus Tideman
 - Hongning Wang



- Continue to grown the community and types of data. <u>Please tell us if you have data to share!!</u>
- Other key tools or pieces that we should consider adding?
- Any issues with the data-formats we are proposing?
- What kind of data would this community like to see?

NICTA Copyright 2014

Thanks!



Questions





• Comments