Proceedings of Machine Learning Research 58 (2016) i-iii

NIPS Workshop on Imperfect Decision Makers 2016: Preface

The volume contains 12 papers accepted to and presented at the 4th Workshop on Imperfect Decision Makes held in conjunction with the 30th Annual Conference on Neural Information Processing Systems (NIPS 2016) in Barcelona on December 9, 2016. The workshop continues the series of international meetings on imperfect decision making that are devoted to the attractive interdisciplinary field covering theory and algorithms. The focus of this workshop was on real-world rationality.

The assumption on decision-maker's rationality is central to the prescriptive Bayesian theory of decision making (DM). However, empirical research indicates that this assumption is often violated by real decision makers. This limits the ability of the prescriptive Bayesian theory to describe the real world. One of the reasons that have been proposed for why the assumption of rationality might be violated by real decision makers is the limited cognitive and computational resources of such decision makers. The workshop intended to inspect this core assumption and to consider possible ways to modify or complement it in order to understand and diminish the discrepancy between the established prescriptive theory and real-world DM.

We were very pleased to have welcomed invited talks from the leading experts in the field, Tom Griffiths (University Berkeley), Itzhak Gilboa (HEC Paris), Daniel Braun (Max Planck Institute for Biological Cybernetics), Naftali Tishby (Hebrew University), Timothy Pleskac (Max Planck Institute for Human Development), David H. Wolpert (Santa Fe Institute), Pedro Ortega (DeepMind). The panel discussions on existing research challenges have indicated that scientific discussions will be going far beyond the limited time of the workshop and encourage further collaboration.

We would like to express our sincere gratitude to all the authors, invited speakers, moderators, discussants, programme committee and reviewers for their invaluable contributions to the success of the workshop. We also would like to thank the NIPS 2016 organisers for their help and assistance. The partial support of the project GA16-09848S is acknowledged.

December 2016

Tatiana V.Guy
guy@utia.cas.cz

Organising Committee:

Miroslav Kárný, Institute of Information Theory and Automation, Czech Republic

David Ríos Insua, Universidad Rey Juan Carlos, Spain

David H. Wolpert, Santa Fe Institute, USA

Tatiana V. Guy, Institute of Information Theory and Automation, Czech Republic

Programme Committee:

Nils Bertschinger, Frankfurt Institute for Advanced Studies, Germany

Anne Collins, UC Berkeley, USA

Michael Harre, University of Sydney, Australia

Gaël Le Mens, Universitat Pompeu Fabra, Spain

David Leslie, Lancaster University, UK

Anthony Quinn, Trinity College Dublin, Ireland

Adrian Raftery, University of Washington, USA

David Ríos Insua, Instituto de Ciencias Matemáticas, Spain

Alessandro E.P. Villa, Université de Lausanne, Swizerland

Invited Talks

Bounded Optimality and Rational Metareasoning in Human Cognition
Tom Griffiths, University Berkeley, USA

Rationality and the Bayesian Paradigm
Itzhak Gilboa, HEC Paris, France

Information-Theoretic Bounded Rationality for Learning and Decision-Making
Daniel Braun, Max Planck Institute for Biological Cybernetics, Germany

Principles and Algorithms for Self-Motivated Behaviour
(joint work with S. Tiomkin and D. Polani)
Naftali Tishby, Hebrew University, Israel

The Rational Status of Quantum Probability Theory Applied to Human Decision Making
(joint work with J. R. Busemeyer)
Timothy J. Pleskac, Max Planck Institute for Human Development, Germany

What the Recent Revolution in Network Coding Tells Us About the Organization of Social Groups (joint work with J. Grana)
David H. Wolpert, Santa Fe Institute, USA

Agency and Causality in Decision Making
Pedro Ortega, DeepMind, UK

Panel discussions

(Ir-)rationality of humans?
Moderator: Peter Grünwald, CWI, Netherlands

Modelling of Rational Decision Making
Moderator: David H. Wolpert, Santa Fe Institute, USA

Proceedings of Machine Learning Research 58 (2016) 1-11

Decision Heuristics for Comparison: How Good Are They?

Marcus Buckmann BUCKMANN@MPIB-BERLIN.MPG.DE
Özgür Şimşek OZGUR@MPIB-BERLIN.MPG.DE
Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, Germany

Editor: T.V. Guy, M. Kárný, D. Rios-Insua, D.H. Wolpert

Abstract

Simple decision heuristics are cognitive models of human and animal decision making. They examine few pieces of information and combine the pieces in simple ways, for example, by considering them sequentially or giving them equal weight. They have been studied most extensively for the problem of *comparison,* where the objective is to identify which of a given number of alternatives has the highest value on a specified (unobserved) criterion. We present the most comprehensive empirical evaluation of decision heuristics to date on the comparison problem. In a diverse collection of 56 real-world data sets, we compared heuristics to powerful statistical learning methods, including support vector machines and random forests. Heuristics performed surprisingly well. On average, they were only a few percentage points behind the best-performing algorithm. In many data sets, they yielded the highest accuracy in all or parts of the learning curve.

1. Introduction

People and animals spend much of their time choosing among alternative options. For example, a venture capitalist chooses among companies to invest in, a writer among potential publishers, and a bee colony among suitable nest sites. The true values of the alternatives are typically not known in advance. A choice is made by examining other pieces of relevant information. For instance, venture capitalists do not know how much they will earn from investing in a particular company but can examine the track record of the founders.

How do people and animals make such decisions? One theory is that they use simple heuristics (Gigerenzer et al., 1999). These simple decision rules examine only a few pieces of information, perhaps only a single piece of information, and combine the pieces in simple ways. For example, lexicographic heuristics consider the various pieces of information sequentially, one at a time, while tallying heuristics give different pieces of information equal weight. There is evidence supporting the use of such simple models in a wide range of decisions made by people and animals (Gigerenzer et al., 2011; Hutchinson and Gigerenzer, 2005).

Compared to standard statistical decision methods, heuristics are very frugal in their use of information and have very low computational requirements. And yet earlier studies have shown that heuristics can be surprisingly effective when compared to statistical

learning methods, including logistic regression, decision trees, naive Bayes, and nearest neighbor methods (Brighton and Gigerenzer, 2008; Martignon and Laskey, 1999; Şimşek and Buckmann, 2015).

Intrigued by these earlier results, we systematically analyzed the performance of heuristics compared to the very best statistical models, including random forests and support vector machines (SVMs), the two models that performed best in a recent large-scale comparison of 179 classification algorithms across 121 data sets (Fernández-Delgado et al., 2014). In this study, of the five highest ranked classifiers, three were random forest implementations (first, second, and fifth position) while the other two were SVM implementations.

We analyzed a diverse collection of 56 real-world data sets that included two well-known heuristics, take-the-best (Gigerenzer and Goldstein, 1996) and tallying (Czerlinski et al., 1999). Heuristics performed remarkably well. On average, they were only a few percentage points behind the best performing algorithm. In many data sets, they yielded the highest accuracy in all or parts of the learning curve.

In the following sections, we first formally define the decision problem we address and describe decision models that are based on heuristics, classification, and regression. We then provide an overview of earlier results on how well heuristics perform. We continue with a description of our methodology and results. We conclude with a discussion of our findings.

2. Background

The decision problem we address is *comparison,* where the objective is to identify the alternative with the highest criterion value, given m alternatives and k attributes on each alternative. We focus on problems with exactly two alternatives. An example is to determine which of two stocks will have a higher return on investment in 5 years, given attributes such as the name recognition of the company.

Let A and B denote the first and the second alternative, respectively. Let $\mathbf{x}_A$ denote the vector of attribute values of alternative A, and y_A its criterion value. The outcome variable of interest is $o_{AB} = sgn(y_A - y_B) \in \{-1, 0, 1\}$, where sgn is the mathematical sign function. The objective is to construct a decision rule for selecting one or the other alternative using the available attributes, in other words, to learn a decision rule $f(\mathbf{x}_A, \mathbf{x}_B) \in \{-1, 0, 1\}$ from training data $T = \{\mathbf{x}_A^i, \mathbf{x}_B^i, o_{AB}^i\}_{i=1}^N$.

The comparison problem is intrinsically symmetrical. Comparing A to B should return the same decision as comparing B to A. That is, $f(\mathbf{x}_A, \mathbf{x}_B)$ should equal $-f(\mathbf{x}_B, \mathbf{x}_A)$. One can expect better accuracy when imposing this symmetry constraint on the learning method.

Notice that an outcome value of 0 may be useful for training but the learned model need not identify this outcome at all. From the decision maker's perspective, when alternatives have equal value, either alternative would qualify as a correct decision.

In the heuristics literature, attributes are called *cues*; we follow this custom when discussing heuristics.

Below, we describe three approaches to comparison and comment on the informational needs of the various approaches.

2.1 Decision heuristics

We consider three heuristics: single-cue (Hogarth and Karelaia, 2005; Şimşek and Buckmann, 2015), tallying (Czerlinski et al., 1999), and take-the-best (Gigerenzer and Goldstein, 1996). These heuristics associate each cue with a *direction* to determine how the cue decides on its own. Cue direction can be positive or negative, favoring the object with the higher or lower cue value, respectively. It can also be neutral, favoring neither object. Cue directions can be learned in a number of ways, including social learning. In our analysis, they are learned from training examples.

Single-cue is perhaps the simplest decision method one can imagine. It compares the alternatives on a single cue, breaking ties randomly. A model for this heuristic specifies the identity of the cue and its direction. We learn both from training examples. Specifically, among the $2k$ possible models, where k is the number of cues, the single-cue model is the one that has the highest accuracy in the training examples.

Tallying is a voting model. It determines how each cue votes on its own—for alternative A, for alternative B, or for neither—and selects the object with the highest number of votes, breaking ties randomly. A tally model needs only to specify cue directions.

Take-the-best is a lexicographic model. It considers the cues one at a time, in a specified order, until it finds a cue that *discriminates* between the alternatives, that is, a cue whose value differs on the two alternatives. It then decides based on that cue alone. A take-the-best model specifies cue directions and cue order.

Cue directions are learned in the same manner for all heuristics. The direction of each cue is learned independently of the directions of other cues. The information required from each training example is simply the direction d of the cue in that example: $d = sgn((y_A - y_B) \times (x_A - x_B)) \in \{-1, 0, +1\}$. Let p and n respectively denote the number of positive and negative samples in the training set. Specifically, $p = \sum_{i=1}^{N} I(d_i = 1)$ and $n = \sum_{i=1}^{N} I(d_i = -1)$, where I is the indicator function. Cue direction is set to positive, negative, or neutral, respectively, if $p > n$, $p < n$, or $p = n$.

Cue order in take-the-best is set to the order of decreasing validity of the cues in the training sample, where *validity* is $max\{p/(p+n), n/(p+n)\}$. That is, cues are ordered by how often they decide correctly when they are able to discriminate between the alternatives.

2.2 Classification

Because our outcome variable is discrete, any classification algorithm is directly applicable. In principle, learning can be done as a function of attribute values of individual objects, but in practice the training data required will be prohibitive. Thus we explore learning a decision rule as a function of attribute differences, $\mathbf{\Delta x}_{AB} = \mathbf{x}_A - \mathbf{x}_B$. Our objective then is to learn decision rule $f(\mathbf{\Delta x}_{AB}) \in \{1, -1, 0\}$ from training data $T = \{\mathbf{\Delta x}^i_{AB}, o^i_{AB}\}_1^N$.

2.3 Regression

Regression can be used to estimate the difference in criterion values of the two alternatives and deduce which alternative has the higher criterion value based on these estimates. Specifically, we train a regressor $h(\mathbf{\Delta x}_{AB})$ with training data $T = \{\mathbf{\Delta x}^i_{AB}, (y_A - y_B)^i\}_1^N$ to use in decision rule $f(\mathbf{x}_A, \mathbf{x}_B) = sgn(h(\mathbf{\Delta x}_{AB}))$.

2.4 Data requirements

Regression requires the highest level of information. It requires that the difference in the criterion values of the alternatives be known in training data, which may not always be possible. In contrast, the training data required for classification is more easily available because all that is required is the identity of the alternative with the higher criterion value—the criterion values of either alternative or the difference in their criterion values are not needed.

Informational needs of heuristics are substantially less than those of classifiers. They do not even require the differences in cue values to be quantified; they need only the sign of cue differences. For example, if *height of a person* is a cue, heuristics need to know which of two people is taller but not the height of either person or the magnitude of the difference.

3. Earlier work

Czerlinski et al. (1999) compared take-the-best and tallying to multiple linear regression in 20 real-world data sets. In each data set, the authors trained the models on all pairwise comparisons among 50% of the objects and tested them on all pairwise comparisons among the remaining objects. They dichotomized the numerical attributes around the median, converting the attribute to binary to mimic people's typically limited knowledge about attribute values and the potential unreliability of precise values. Take-the-best performed best, with a mean accuracy of 0.72 across data sets, compared to 0.69 for tallying, and 0.68 for multiple linear regression. When the authors repeated their analysis without dichotomizing the attributes but using their exact numerical values, mean accuracies of take-the-best and multiple linear regression were identical at 0.76. The authors did not test tallying with numerical attributes.

Brighton (2006) presented learning curves on eight data sets, where attributes were again dichotomized around the median, comparing take-the-best to neural networks, decision trees, and nearest neighbor methods. In four data sets, take-the-best had the highest accuracy on almost the entire learning curve. In the other four data sets, take-the-best had the highest accuracy on at least some parts of the learning curve.

Katsikopoulos et al. (2010) showed mean accuracy in 19 data sets for training samples of 2 to 10 objects. They compared multiple models, including take-the-best, tallying, multiple linear regression, and naive Bayes, implementing most models with and without dichotomizing the attributes. Take-the-best (with undichotomized cues) had the highest accuracy for all but the smallest training-sample size of two objects, in which case tallying (with dichotomized cues) had the highest accuracy. Again, tallying was not tested with exact numerical values.

Brighton and Gigerenzer (2012) compared take-the-best to SVM in a single data set, where attributes were naturally binary, and found that the accuracy levels of the two models were comparable throughout the learning curve.

Şimşek and Buckmann (2015) presented learning curves for heuristics, logistic regression, and decision trees on 63 natural data sets. On average, tallying was the most accurate method on very small sample sizes. When models were trained on 50% of instances in the data set, mean accuracy across data sets was 0.725 for tallying, 0.743 for take-the-best, 0.746 for CART, and 0.747 for logistic regression.

Table 1: Data sets used in the analysis.

ID	Name	Objects	Cues	ID	Name	Objects	Cues
1	Manpower	17	5	29	Mortality	60	15
2	Waste	20	5	30	Movie	62	12
3	Jet	22	5	31	Dropout	63	18
4	Sperm	24	8	32	Land	67	4
5	Cigarette	25	3	33	Lakes	69	10
6	Galápagos	29	5	34	City	76	9
7	Agriculture	29	6	35	Car	93	21
8	Ice	30	3	36	Basketball	96	4
9	Oxidant	30	4	37	Infant	101	3
10	Recycling	31	7	38	Obesity	136	11
11	Reactor	32	10	39	Contraception	152	6
12	Rebellion	32	6	40	Votes	159	5
13	Excavator	33	5	41	Pitcher	176	15
14	Occupation	36	3	42	Birthweight	189	8
15	Pinot	38	6	43	Athletes	202	8
16	Highway	39	11	44	CPU	209	6
17	AFL	41	4	45	Tip	244	6
18	Air	41	6	46	Bodyfat	252	13
19	Bones	42	6	47	Hitter	263	19
20	Mussels	44	8	48	Diamond	308	4
21	Mines	44	4	49	Algae	340	11
22	Prefecture	45	5	50	Faculty	397	5
23	Crime	47	15	51	Mileage	398	7
24	Homeless	50	7	52	Monet	430	4
25	SAT	50	4	53	Affair	601	8
26	Fuel	51	5	54	Lung	654	4
27	Salary	52	5	55	Rent	2053	10
28	Sleep	58	7	56	Home	3281	4

4. Analysis

We used 56 data sets gathered from a wide variety of sources, including online data repositories, statistics and data-mining competitions, packages for R statistical software, textbooks, and research publications. The subjects were diverse, including biology, business, computer science, ecology, economics, education, engineering, environmental science, medicine, political science, psychology, sociology, sports, and transportation. The data sets varied in size, ranging from 17 to 3,281 objects. They also varied in the amount of information available, ranging from 3 to 21 attributes. Many of the smaller data sets contained the entirety of the population of objects, for example, all 29 islands in the Galápagos archipelago. The data sets are listed in Table 1 and described in the supplementary material. All data sets are publicly available.

Missing attribute values were imputed by the mean, median, and mode value in the data set for interval, ordinal, and categorical attributes, respectively. Objects with missing criterion values were discarded. For ordinal attributes, attribute difference between two objects was recoded into a new ordinal attribute with three values, indicating if the differ-

ence is positive, negative, or zero. This is because the exact value of the difference is not meaningful for ordinal attributes, only its sign. A categorical attribute with c categories was recoded into c binary attributes, each indicating membership in one category.

Almost all earlier studies have dichotomized the numerical attributes as described in Section 3. We did not. All models, including heuristics, generally yield higher accuracy when attributes are not dichotomized.

We examine two performance metrics on decision quality. Our primary metric is accuracy, where a decision is considered to be accurate if it selects the object with the higher criterion value or if the objects are equal in criterion value. We examine, in addition, a linear loss metric that equals 0 if the decision is accurate, and $\frac{|y_A - y_B|}{z}$ otherwise, where z is a normalizing constant for the data set, computed as follows:

$$z = \frac{\sum_{\forall \langle A,B \rangle} |y_A - y_B|}{\sum_{\forall \langle A,B \rangle} 1}. \tag{1}$$

We present results with the following classification algorithms: random forests, SVMs, and naive Bayes; the following regression algorithms: multiple linear regression (MLR) with elastic net penalty and random forest regression; and the following heuristics: single-cue, take-the-best, and tallying. We trained SVMs using their implementation in the R package e1071 (Meyer et al., 2014). We tried both a linear and a Gaussian kernel, with a 10-fold cross-validated grid search for parameter values. We trained random forests and random forest regression using the implementation in the R package randomForest (Liaw and Wiener, 2002), with 10-fold cross-validated search for the best value of the parameter *mtry*. We trained linear regression with elastic net regularization (Zou and Hastie, 2005), using the R package glmnet (Friedman et al., 2009). We selected the parameter values of α and λ using 10-fold cross-validation. We used the naive Bayes implementation in R package e1071 (Meyer et al., 2014), with Laplace smoothing. Additional implementation details are described in the supplementary material.

5. Results

We present results on the performance of each algorithm as training-set size increases, starting from a size of one. Recall that a training instance requires information on two objects. Consequently, a single training instance uses two objects from the data set. These two objects are then discarded, not to be used again in any train or test instance.

To generate learning curves, we randomly sampled m test instances from each data set ($m = n/10$, where n is the number of objects in the data set). We then progressively sampled training sets of increasing size using the remaining objects in the data set. We replicated this procedure 4,000 times for SVM, and 10,000 times for all other algorithms. The smaller numbers for SVM are due to the substantially higher training time required for this model.

Figure 1 shows the mean learning curve across 56 data sets for various algorithms. Table 2 reports mean accuracy and linear loss across 56 data sets for training sets of size 10, 20, 50, and 90 instances. Figure 2 displays individual learning curves on 20 of the data sets.

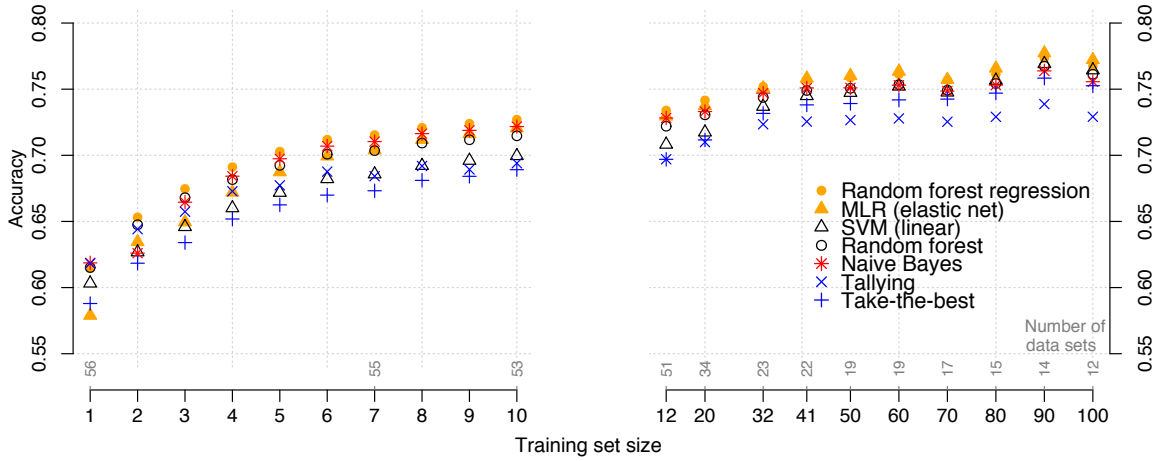

Figure 1: Mean accuracy in 56 data sets as a function of the training-set size. On the horizontal axis, the figure also shows the number of data sets contributing to the mean performance.

Table 2: Mean performance across 56 data sets

	Accuracy				1 − Linear loss			
Training set size	10	20	50	90	10	20	50	90
Data sets	$n = 53$	$n = 34$	$n = 19$	$n = 14$	$n = 53$	$n = 34$	$n = 19$	$n = 14$
Take-the-best	0.689	0.712	0.739	0.758	0.741	0.772	0.797	0.819
Tallying	0.694	0.710	0.727	0.739	0.750	0.772	0.784	0.798
Single-cue	0.673	0.695	0.723	0.746	0.721	0.749	0.775	0.804
Naive Bayes	0.722	0.733	0.751	0.764	0.792	0.807	0.817	0.832
Random forest	0.715	0.731	0.751	0.767	0.781	0.798	0.811	0.829
SVM (linear)	0.700	0.717	0.747	0.769	0.765	0.784	0.811	0.835
SVM (radial)	0.687	0.710	0.741	0.762	0.745	0.776	0.801	0.824
Random forest regression	0.727	0.742	0.759	0.774	0.800	0.814	0.823	0.839
MLR (elastic net)	0.721	0.737	0.760	0.777	0.788	0.809	0.827	0.846

Along the mean learning curves, the differences between the heuristics and the statistical learning algorithms are relatively small. The maximum difference in accuracy between the best heuristic and the best algorithm at a given sample size is at most 0.037. On the early parts of the curve, tallying performed better than take-the-best but is roughly 0.025 percentage points behind the best performing algorithms: random forest, random forest regression, and naive Bayes. With larger training-set sizes, take-the-best performed remarkably well. With training-set sizes of 44 or larger, it trailed the top algorithm by 0.019 on average. We should also note that a relatively simple method, naive Bayes, performed remarkably well. Along the mean learning curve, it closely trailed random forest regression on most training-set sizes.

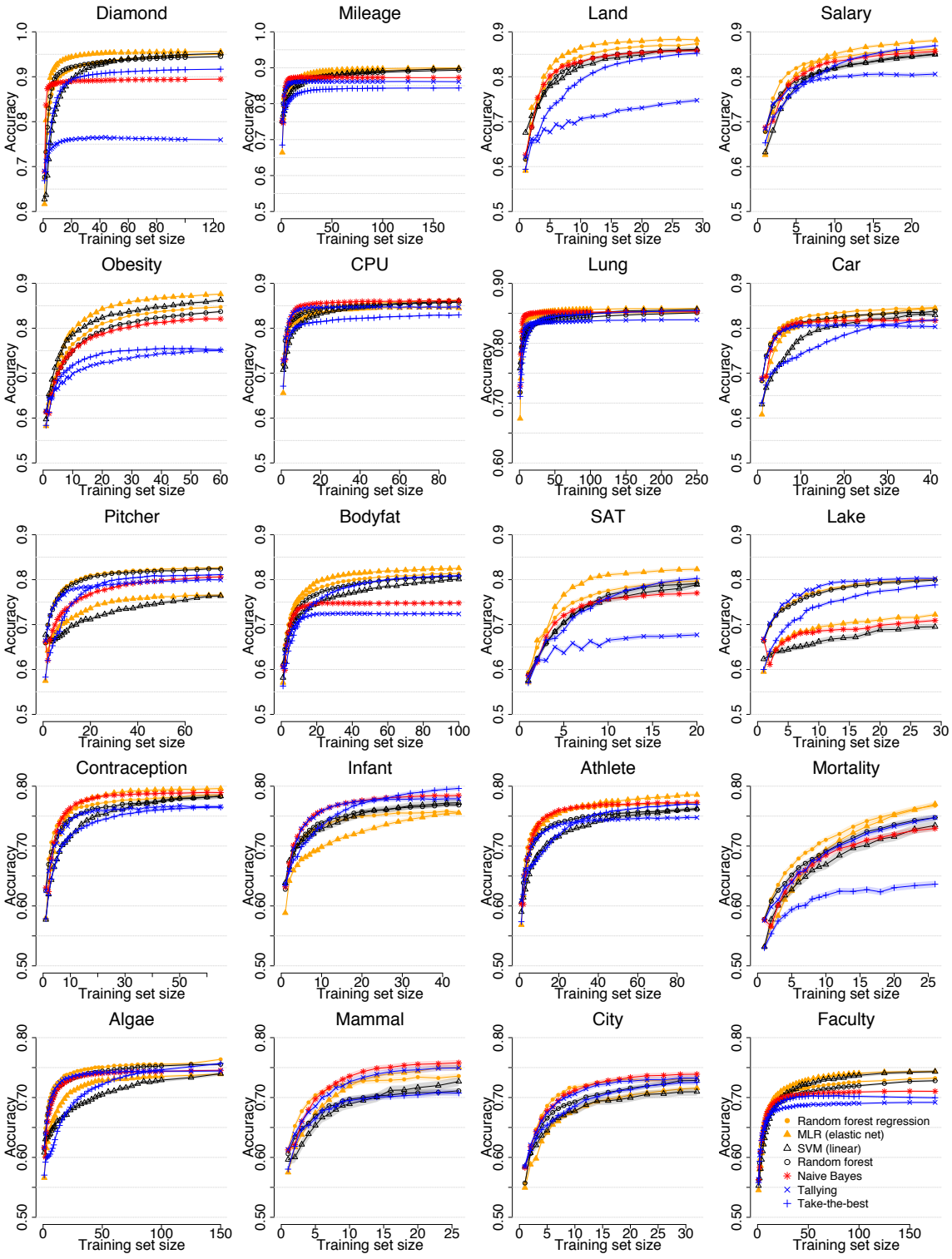

Figure 2: Learning curves on individual data sets.

On individual data sets, there were substantial differences in the performance of the various algorithms. In most of the data sets, at least one heuristic performed as well as or better than the best performing classification algorithm in all or parts of the learning curve. The results on linear loss are very similar to those on accuracy.

We measured the computation time required for training and testing the various models midway along the learning curve, where the size of the training set was roughly 90% of its maximum value. The experiment was run on a single kernel of a cluster (Intel Xeon CPU E5-2670, 4GB memory). For one pass over the data sets, SVM with the radial kernel required on average 283 minutes, SVM with the linear kernel required 16 minutes, random forest required two minutes, and MLR (elnet) required one minute. In contrast, take-the-best and tallying each required only 0.40 seconds. Note that the R packages we used to train SVM and random forests both call highly efficient Fortran and C code, while our implementation of the heuristics is programmed entirely in the much slower R language.

6. Discussion

Among the decision methods tested earlier in the literature, including logistic regression and decision trees, tallying stood out as the best method for sample sizes with 1–10 instances (Şimşek and Buckmann, 2015). We found that tallying (on average) falls short of more powerful statistical algorithms—random forest, random forest regression, and naive Bayes—even when training sets are small.

One surprising result is that with larger training-set sizes, multiple linear regression (with elastic net penalty) performed better on average than any other algorithm. This result has important implications because decision heuristics are often treated as an approximation of a linear decision rule. Several properties of the decision environment are known to allow heuristics to approximate a linear algorithm (Hogarth and Karelaia, 2006; Baucells et al., 2008; Martignon and Hoffrage, 2002; Katsikopoulos, 2011). Furthermore, these properties are prevalent in natural decision problems (Şimşek, 2013; Şimşek et al., 2016).

It is fair to conclude that in a diverse collection of natural environments, heuristics fared remarkably well when compared to powerful statistical learning algorithms. To put this result into context, it is useful to remember that the computational, informational, and memory requirements of heuristics, both at training and decision time, are extremely low.

One possible reason for the success of heuristics is that comparison is an easy problem, at least when compared to regression or classification. Given the fundamental importance of comparison for intelligent behavior, it would be fruitful to examine this problem theoretically and to develop statistical learning algorithms that address it directly, taking advantage of its special properties.

We hope these results will encourage further study of decision heuristics. In particular, we hope they will motivate further mathematical analysis as well as development of additional heuristic models.

References

Manel Baucells, Juan A. Carrasco, and Robin M. Hogarth. Cumulative dominance and heuristic performance in binary multiattribute choice. *Operations research*, 56(5):1289–1304, 2008.

Henry Brighton. Robust Inference with Simple Cognitive Models. In C. Lebiere and R. Wray, editors, *AAAI Spring Symposium: Cognitive Science Principles Meet AI Hard Problems*, pages 17–22. American Association for Artificial Intelligence, 2006.

Henry Brighton and Gerd Gigerenzer. Bayesian brains and cognitive mechanisms: Harmony or dissonance? In Nick Chater and Mike Oaksford, editors, *The probabilistic mind: Prospects for Bayesian cognitive science*, pages 189–208. Oxford University Press, New York, 2008.

Henry Brighton and Gerd Gigerenzer. Are rational actor models "rational"outside small worlds. *Evolution and Rationality: Decisions, Co-operation, and Strategic Behavior*, pages 84–109, 2012.

Jean Czerlinski, Gerd Gigerenzer, and Daniel G. Goldstein. How good are simple heuristics? In Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, editors, *Simple heuristics that make us smart*, pages 97–118. Oxford University Press, New York, 1999.

Manuel Fernández-Delgado, Eva Cernadas, Senén Barro, and Dinani Amorim. Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research*, 15(1):3133–3181, 2014.

Jerome Friedman, Trevor Hastie, and Rob Tibshirani. *glmnet: Lasso and elastic-net regularized generalized linear models*, 2009. URL `http://CRAN.R-project.org/package=glmnet`.

Gerd Gigerenzer and Daniel G. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669, 1996.

Gerd Gigerenzer, Peter M. Todd, ABC Research Group, et al. *Simple heuristics that make us smart*. Oxford University Press, New York, 1999.

Gerd Gigerenzer, Ralph Hertwig, and Thosten Pachur, editors. *Heuristics: The Foundations of Adaptive Behavior*. Oxford University Press, New York, 2011. ISBN 9780199744282.

Robin M. Hogarth and Natalia Karelaia. Ignoring information in binary choice with continuous variables: When is less "more"? *Journal of Mathematical Psychology*, 49(2):115–124, 2005.

Robin M. Hogarth and Natalia Karelaia. "take-the-best" and other simple strategies: Why and when they work "well" with binary cues. *Theory and Decision*, 61(3):205–249, 2006.

John Hutchinson and Gerd Gigerenzer. Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69(2):97–124, 2005.

Konstantinos V. Katsikopoulos. Psychological heuristics for making inferences: Definition, performance, and the emerging theory and practice. *Decision Analysis*, 8(1):10–29, 2011.

Konstantinos V. Katsikopoulos, Lael J. Schooler, and Ralph Hertwig. The robust beauty of ordinary information. *Psychological Review*, 117(4):1259–1266, 2010.

Andy Liaw and Matthew Wiener. Classification and regression by randomforest. *R News*, 2(3):18–22, 2002. URL `http://CRAN.R-project.org/doc/Rnews/`.

Laura Martignon and Ulrich Hoffrage. Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52(1):29–71, 2002.

Laura Martignon and Kathryn B. Laskey. Bayesian benchmarks for fast and frugal heuristics. In Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, editors, *Simple heuristics that make us smart*, pages 169–188. Oxford University Press, New York, 1999.

David Meyer, Evgenia Dimitriadou, Kurt Hornik, Andreas Weingessel, and Friedrich Leisch. *e1071: Misc Functions of the Department of Statistics (e1071), TU Wien*, 2014. URL `http://CRAN.R-project.org/package=e1071`. R package version 1.6-4.

Özgür Şimşek. Linear decision rule as aspiration for simple decision heuristics. In Chris J. C. Burges, Leon Bottou, Max Welling, Zoubin Ghahramani, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 2904–2912. Curran Associates, Inc., Red Hook, New York, 2013.

Özgür Şimşek and Marcus Buckmann. Learning from small samples: An analysis of simple decision heuristics. In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 3159–3167. Curran Associates, Inc., Red Hook, New York, 2015.

Özgür Şimşek, Simón Algorta, and Amit Kothiyal. Why most decisions are easy in tetris—and perhaps in other sequential decision problems, as well. In *Proceedings of The 33rd International Conference on Machine Learning*, pages 1757–1765, 2016.

Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.

Proceedings of Machine Learning Research 58 (2016) 13-25

Simple Regression Models

Jan Malte Lichtenberg LICHTENBERG@MPIB-BERLIN.MPG.DE
Özgür Şimşek OZGUR@MPIB-BERLIN.MPG.DE
Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, Germany

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

Developing theories of when and why simple predictive models perform well is a key step in understanding decisions of cognitively bounded humans and intelligent machines. We are interested in how well simple models predict in regression. We list and review existing simple regression models and define new ones. We identify the lack of a large-scale empirical comparison of these models with state-of-the-art regression models in a predictive regression context. We report the results of such an empirical analysis on 60 real-world data sets. Simple regression models such as equal-weights regression routinely outperformed state-of-the-art regression models, especially on small training-set sizes. There was no simple model that predicted well in all data sets, but in nearly all data sets, there was at least one simple model that predicted well.

Keywords: Regression, Simple Heuristics, Improper Models

1. Introduction

The study of simple predictive models is an important topic in decision making and machine learning. High predictive accuracy seems to be the main focus of most current research. Yet low time complexity, robustness to small sample sizes, and interpretability are also desirable properties of a useful model. In this article, we study simple models which are much faster to estimate and easier to understand than their current state-of-the-art peer algorithms. Their advantage in computation time and interpretability will be obvious. We are mainly interested in how much predictive power is lost when using simple models, if any.

Predictive models can be simple in many different ways. One seemingly extreme approach is to take only one predictor into account. Another approach is to take all predictors into account but to combine them in simple ways, for example, by giving them equal or random weights. Such simple models have been shown to predict remarkably well in tasks such as classification (Holte, 1993), paired comparison (Czerlinski et al., 1999; Brighton, 2006; Şimşek and Buckmann, 2015), and portfolio optimization (DeMiguel et al., 2009). Less attention has been given to simple models in a predictive regression context, that is, when the problem under consideration is to estimate the value of a continuous response variable on previously unseen data.

Simple regression models such as equal weights and random weights regression have been examined empirically and theoretically in the psychological literature. Collectively,

such models were termed *improper* linear models to distinguish them from *proper* linear models whose weights are obtained by optimizing some objective function. For example, ordinary least squares are obtained by minimizing the residual sum of squares.

Wainer (1976) argued that "it don't make no nevermind" if optimal weights are replaced by equal weights, as the loss in explained variance is small if predictors are directed properly (see also Laughlin, 1978; Wainer, 1978). Similar findings, commonly known as the *flat maximum effect*, show that the space of nearly optimal weights for linear models is large (Ehrenberg, 1982; von Winterfeldt and Edwards, 1982; Lovie and Lovie, 1986). Dawes and Corrigan (1974) and Dawes (1979) concluded that optimal weighting of predictors would therefore be less important than choosing the right predictors and knowing their directional relationship with the response.

Moreover, Einhorn and Hogarth (1975) argued that improper models suffer from smaller estimation error compared to proper models (or no estimation error) because the weights of improper models do not have to be estimated from the data. Because they combine a small loss in accuracy with increased robustness due to smaller estimation error, the common message of these studies was that improper models would be superior to multiple linear regression in some situations and not greatly inferior in others when the aim is out-of-sample prediction.

These surprising results showed that improper models can match the performance of more complex models or even outperform them. However, existing work on simple regression models did not study these models in a regression context. Some of them used loss functions that are not regression adequate (Wilks, 1938; Dawes and Corrigan, 1974; Dawes, 1979; Einhorn and Hogarth, 1975; Dana and Dawes, 2004). For example, Dawes (1979) used the correlation coefficient between predicted and true response values to assess the prediction performance of different models. Other studies evaluated improper regression models in a fitting rather than in a prediction context (Wainer, 1976; Waller and Jones, 2009). Furthermore, many results were presented relative to the performance of multiple linear regression (Einhorn and Hogarth, 1975; Wainer, 1976; Graefe, 2015), which is known to have severe estimation issues under a large variety of conditions. It is unclear whether the results still hold relative to more recent proper models, such as the elastic net (Zou and Hastie, 2005) or other regularized linear models.

In this article, we examine how well simple regression models predict in a regression context when compared with state-of-the-art statistical models in a large, diverse collection of real-world data sets. In doing so, we complement and contrast findings from other domains such as classification and paired comparison, building toward a more general theory of when and why simple models perform well.

Our results show that simple regression models such as equal weights regression routinely outperformed not only multiple linear regression but also state-of-the-art regression models, especially on small training-set sizes. There was no simple model that predicted well in all data sets, but for nearly all data sets, there was at least one simple model that predicted well.

In Sections 2 and 3 we review existing simple regression models, define new ones, and describe how to estimate their parameters. In Section 4 we report the results of a large empirical study that compared simple models with state-of-the-art regression algorithms on 60 data sets using a regression-adequate loss function.

2. Simple regression models

Let us assume that we have some data $(y_i, \boldsymbol{x}_i),\ i = 1, \ldots, n$, where $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})$ is a p-dimensional vector of predictors and y_i a real-valued response for the ith observation. A *regression model* f is a model that makes a prediction $\hat{y}$ of y for a potentially new input vector $\boldsymbol{x}$, that is,

$$\hat{y} = f(\boldsymbol{x}).$$

The simple models we consider are special instances of the linear regression model[1]

$$\hat{y} = \beta_0 + \gamma \sum_{j=1}^{p} x_j \alpha_j,$$

and share the following properties: (a) weights α_j are chosen heuristically (for example, equal weights), and (b) weights α_j can be estimated or determined independently of the location parameter β_0 and the scale parameter γ. Intuitively, the weighted sum determines the nature of how the predictors are combined or selected. The two parameters β_0 and γ then determine the location and scale of this weighted sum, respectively. Different simple models correspond to different ways of determining α_j. Estimation of β_0 and γ is the same for all considered simple models and is explained in the following section.

We assume that predictors and responses are *centered*, that is,

$$\bar{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i = 0 \quad \text{and} \quad \bar{\boldsymbol{x}}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} = 0 \text{ for all } j = 1, \ldots, p,$$

and *scaled*, that is,

$$s_{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i^2 = 1 \quad \text{and} \quad s_{\boldsymbol{x}_j} = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^2 = 1 \text{ for all } j = 1, \ldots, p.$$

A centered and scaled variable is called *standardized*. Scaling the response is not necessary for the simple models to function well but simplifies the analysis across different data sets. Furthermore, predictors are said to be *directed* if they correlate non-negatively with the response. We now define each of the considered simple models. Table 1 points to existing literature for each model.

Mean prediction. This is the simplest available model and always predicts the mean value of the response calculated on the training data. The corresponding model can be written as

$$\hat{y} = \beta_0.$$

Mean prediction is appropriate when no predictor is available. Typically, data sets that do not contain predictive predictors are not considered. Therefore, mean prediction does not play a role in supervised learning in general. Yet the model can still serve as a baseline.

Random weights. This is perhaps the most improper model one could imagine. Once standardized and directed, each predictor is assigned a random weight stemming from a

1. Setting $\beta_j = \gamma\alpha_j$, we obtain the classic formulation of linear regression: $\hat{y} = \beta_0 + \sum_{j=1}^{p} x_j\beta_j$.

Model	Literature
Random weights	Wilks (1938); Dawes and Corrigan (1974)
Equal weights	Wesman and Bennett (1959); Schmidt (1971); Dawes and Corrigan (1974); Einhorn and Hogarth (1975); Wainer (1976); Dawes (1979); Dana and Dawes (2004); Davis-Stober et al. (2010); Graefe (2015)
Correlation weights	Dana and Dawes (2004); Waller and Jones (2009)
Single-cue regression	Dana and Dawes (2004); Davis-Stober et al. (2010)
Correlation ranks	Wesman and Bennett (1959)

Table 1: Literature on simple regression models.

uniform distribution, that is,

$$\hat{y} = \beta_0 + \gamma \sum_{j=1}^{p} \omega_j x_j,$$

where $\omega_j \sim \mathcal{U}(a, b)$. Different authors used different values for a and b. We used $a = 0$ and $b = 1$. Nearly 80 years ago, Wilks (1938) showed that the correlation of predictions of two independent random-weights models tends to 1 with an increasing number of positively intercorrelated variables. Random weights should be outperformed by equal weights because of the smaller variance of the latter (Dawes, 1979). We include random weights as a lower benchmark model in our empirical analysis.

Equal weights. This model takes all standardized predictors into account and weights them equally, that is,

$$\hat{y} = \beta_0 + \gamma \sum_{j=1}^{p} x_j. \tag{1}$$

Under the assumption that all predictors are directed, equal weights has only two free parameters, location and scale.

Equal-weighting has been discussed in a large variety of settings resulting in slightly different models: If $\beta_0 = 0$ and $\gamma = 1$ in Equation (1), the resulting model is called *unit weights* (Einhorn and Hogarth, 1975). In a paired comparison context, where equal-weights models have been called *tallying* or *Dawes's rule*, these models have been shown to outperform more complex models, especially on small sample sizes (Gigerenzer et al., 1999; Şimşek and Buckmann, 2015). An equal-weights model has been found to compete well with state-of-the-art portfolio theory models in a portfolio allocation problem, where it is called the *1/*N *rule* (DeMiguel et al., 2009).

Correlation weights. This model weights all predictors by their correlation with the response, that is,

$$\hat{y} = \beta_0 + \gamma \sum_{j=1}^{p} r_{yx_j} x_j,$$

where r_{yx_j} is the sample correlation coefficient between the response and predictor $\boldsymbol{x}_j$. Correlation weights has to estimate $p + 2$ parameters. However, these coefficients are easier to calculate than ordinary least squares (OLS) weights in terms of both computational complexity and numerical stability issues. Whereas the OLS model suffers, for example, from the multicollinearity problem, the correlation coefficients are calculated independently

of each other and independently of β_0 and γ. The correlation-weights model thus scales favorably with the number of predictors when compared to the OLS model and its matrix inversions.

Single-cue regression. This model considers only the predictor that has the highest correlation with the response among all available predictors. The corresponding model can be written as

$$\hat{y} = \beta_0 + \gamma x_1,$$

where x_1 is the cue that is most correlated with the criterion y. To determine *the* single cue, the correlations between all predictors and the response are estimated and the one with the highest absolute value is chosen. Estimation of single-cue regression is not less complex than estimation of correlation weights as it is necessary to calculate all p predictor–response correlations in order to determine the single cue. Yet there may be simpler ways to (approximately) determine the single cue, and single-cue regression is simpler at decision time, where it requires only the information of one predictor.

Correlation ranks. This model does not need the exact values of the correlation weights but only their ranks, that is, their relative order. The corresponding model can be written as

$$\hat{y} = \beta_0 + \gamma \sum_{j=1}^{p} \rho_j \; x_j, \quad \text{where} \quad \rho_j = rank(r_{yx_j}).$$

The lowest correlated cue has rank 1 and the highest correlated cue has rank p. Ties are assigned the average rank.[2] Correlation ranks might be easier to estimate and thus more robust than correlation or OLS weights but still allow for differential weighting of multiple predictors, as opposed to equal-weighting or single-cue strategies. Our implementation of correlation ranks actually first estimates all correlations and then assigns ranks. However, there may be simpler ways to (approximately) determine the ranking of correlations. Models similar to correlation ranks have been compared to true-weights and equal-weights models in the context of *multiattribute decision making*[3] in Barron and Barrett (1996). We know of no study that compares the prediction accuracy of correlation-ranks models to other regression models in a regression context.

3. Parameter estimation from training data

Unless specifically stated otherwise, we assume that β_0 and γ are calculated using simple linear regression (SLR). Estimation of the weights α_j depends on the respective algorithm and is done before the estimation of β_0 and γ.

SLR is much easier to estimate than OLS regression in general as it involves no inversion of matrices but only simple estimates of scale and covariation. Defining $c(\boldsymbol{x}_i) = \sum_{j=1}^{p} x_{ij}\alpha_j$ and $\boldsymbol{c} = (c(\boldsymbol{x}_1), \ldots, c(\boldsymbol{x}_n))^T$, the SLR estimates for model (2) are given by

$$\gamma = r_{\boldsymbol{yc}} \frac{s_{\boldsymbol{y}}}{s_{\boldsymbol{c}}} \quad \text{and} \quad \beta_0 = \bar{\boldsymbol{y}} - \gamma \bar{\boldsymbol{c}},$$

where $r_{\boldsymbol{yc}}$ is the sample correlation coefficient between $\boldsymbol{y}$ and $\boldsymbol{c}$, and $s_{\boldsymbol{y}}$ and $s_{\boldsymbol{c}}$ are the standard deviations of $\boldsymbol{y}$ and $\boldsymbol{c}$, respectively.

2. For example, the vector $(7, 4, 4, 2)$ has ranks $(4, 2.5, 2.5, 1)$.
3. Find the alternative with the highest response value among a set of $n \geq 2$ alternatives.

ID	Name	Obs.	Predictors
1	Abalone	4,177	8
2	AFL	41	5
3	Air	41	6
4	Airfoil	1,503	5
5	Algae	340	11
6	Athlete	202	8
7	Basketball	96	4
8	Birth weight	189	8
9	Body fat	252	13
10	Bone	42	6
11	Car	93	21
12	Cigarette	528	7
13	Concrete	1,030	8
14	Contraception	152	6
15	CPU	209	6
16	Crime	47	15
17	Diabetes	442	10
18	Diamond	308	4
19	Dropout	63	17
20	Excavator	33	4
21	Fish	413	3
22	Fuel	51	5
23	Gambling	47	4
24	Highway	39	11
25	Hitter	263	19
26	Home	3,281	4
27	Homeless	50	7
28	Infant	101	3
29	Labor supply	5,320	5
30	Lake	69	10

Id	Name	Obs.	Predictors
31	Land	67	4
32	Lung	654	4
33	Mammal	58	7
34	Medical expenditure	5,574	14
35	Men	34	3
36	Mileage	398	7
37	Mine	44	4
38	Monet	430	4
39	Mortality	60	15
40	Movie	62	12
41	Mussel	44	8
42	News	39,644	52
43	Obesity	136	11
44	Occupations	36	3
45	Pinot	38	6
46	Pitcher	176	15
47	Plasma	315	12
48	Prefecture	45	5
49	Prostate	97	8
50	Reactor	32	10
51	Rebellion	32	6
52	Recycle	31	7
53	Rent	2,053	10
54	Salary	52	5
55	SAT	50	4
56	Schooling	3,010	22
57	Tip	244	6
58	Vote	159	5
59	Wage	4,360	10
60	White wine	4,898	11

Table 2: Data sets used in the empirical comparison.

Note that β_0 can be omitted (set to 0) for all models when predictors and responses are standardized. Some authors do not include the scale parameter γ when the loss function is invariant under scaling. In this article, we are interested in regression under squared error loss, which is not invariant under scaling. Inclusion of γ is therefore crucial.

4. Empirical analysis

We examined the predictive accuracy of simple regression models on a large collection of real-world data sets. We assessed the predictive accuracy of different algorithms using root mean squared prediction error (RMSE).

Data sets. We used 60 publicly available data sets from varying domains. Sources included online data repositories, statistics and data-mining competitions, packages for R statistical software, textbooks, and research publications. The number of observations ranged from 31 to 39,644, the number of predictors from 3 to 52. Table 2 shows the number of observations and the number of predictors in each data set. Many data sets are from

earlier studies (Czerlinski et al., 1999; Şimşek, 2013). Detailed information about the data sets can be found in the Online Appendix.

We had difficulties finding data sets for regression where $p > n$. Although regularized linear models have been developed primarily for problems where $p > n$, such data naturally mostly seem to occur in (binary) classification contexts in biology and genomics. Nonetheless, $p > n$ situations occurred in our learning curve analysis for small training-set sizes.

For each data set, the response was standardized and all predictors were standardized and directed. Missing predictor values were mean imputed and observations with missing response values were removed from the data set.

Benchmark models. We chose to include three benchmark models, described below. Two are state-of-the-art regression models. We included the OLS model for historic reasons.

(1) The OLS model minimizes the mean squared error between predicted and true values on the training data. We used the R (R Core Team, 2015) built-in function *lm* for estimating OLS. Whenever there were p predictors, n observations, and $p > n$, we used only the $n - 1$ predictors that were most correlated with the response.

(2) The elastic net (Zou and Hastie, 2005) is a state-of-the-art regularized linear regression model. Regularized linear models were originally developed to overcome the estimation difficulties of OLS (Hoerl and Kennard, 1970). They attempt to optimize prediction accuracy by finding the happy medium between simplicity and complexity. The elastic net has two main parameters. Parameter $\lambda \geq 0$ controls the overall strength of regularization. The elastic net reduces to OLS for $\lambda = 0$. Parameter $0 \leq \alpha \leq 1$ controls the amount of *ridge* versus *Lasso* characteristics. The elastic net reduces to two other regularized linear models, ridge regression (Hoerl and Kennard, 1970) when $\alpha = 0$ and the Lasso (Tibshirani, 1996) when $\alpha = 1$. We used the R package *glmnet* (Friedman et al., 2015) to estimate the elastic net. The parameters α and λ were jointly optimized using 10-fold cross-validation on a two-dimensional grid with $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$ and λ on the built-in search path of glmnet, which uses a log-spaced grid with a maximum of 100 candidate values. We also tested ridge regression and the Lasso in our empirical study and found their results to be very similar to those of the elastic net.

(3) Random forest regression (Breiman, 2001) is a non-parametric and non-linear regression model. It optimizes prediction accuracy by fitting an ensemble of regression trees. We used the R package *randomForest* (Liaw and Wiener, 2002) with `ntree = 500` trees per forest.

Results. We show three sets of results. Figure 1 shows the mean RMSE of each algorithm across 60 data sets, computed using 10-fold cross-validation. These estimates of the prediction error correspond to large training-set sizes relative to the total size of the data set (90% of available observations).

Figure 2 shows learning curves averaged across 60 data sets, as the training set varied in size from 4 to 100. The figure shows learning curves for all models except mean prediction, random weights, and OLS.[4] The test set consisted of 10% of the total number of observations in the data set and did not overlap with the corresponding training set. The estimation procedure was repeated 100 times for each training-set size and algorithm. There is a slight

4. OLS overfits for small ratios of n/p. Resulting average RMSEs were outside the figure boundaries because some data sets had a large number of predictors p.

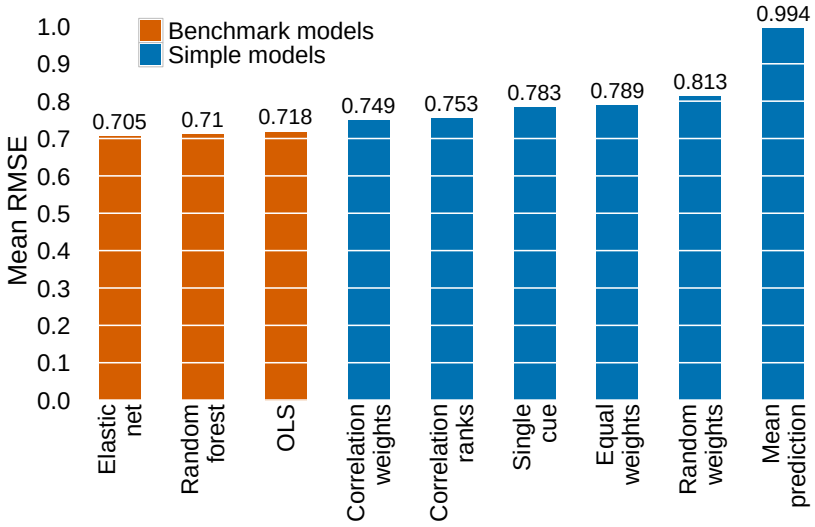

Figure 1: 100 × 10-fold cross validated root mean squared error (RMSE) across 60 data sets. OLS = ordinary least squares.

increase toward the end of the learning curve because the means were calculated on fewer data sets for higher training-set sizes. The number of data sets large enough to be eligible for a given training-set size is indicated at the top of the figure.

Finally, we present learning curves of various algorithms in individual data sets. Figure 3 shows learning curves in the data sets *Diabetes, Prostate*, and *SAT*. Figures A.1 to A.3 in the Online Appendix present the learning curves in all remaining data sets.

We first compare simple regression models to benchmark models collectively. We then comment on results within the groups of simple and benchmark models, respectively.

Averaged across 60 data sets, simple models were collectively outperformed by all benchmark models for larger training-set sizes.[5] However, equal weights and correlation ranks outperformed all competing models for training-set sizes below 15 on the mean learning curve. In addition, the learning curves in individual data sets show that for many data sets, there is at least one simple model that performed well across large parts of the learning curve. Let us define the minimum error curve as the algorithm with minimum error among all algorithms as a function of training-set size. Then, in 22 of 60 data sets, simple models occupied the entire minimum error curve except for possibly one training-set size. In another 21 data sets, simple models occupied at least half of the minimum error curve. The 17 remaining data sets were dominated by benchmark models.

On many data sets some simple models performed very well while others performed very poorly, rather than all simple models performing equally well. A good example is the

5. The end of the learning curve shows the average across all 30 data sets that were large enough for training-set sizes of 100. The cross-validation-based analysis of Figure 1 shows the average across all 60 data sets for training-set sizes ranging from 27 to 35,679 observations, corresponding to 90% of the total size of the respective data sets. The two analyses show similar results.

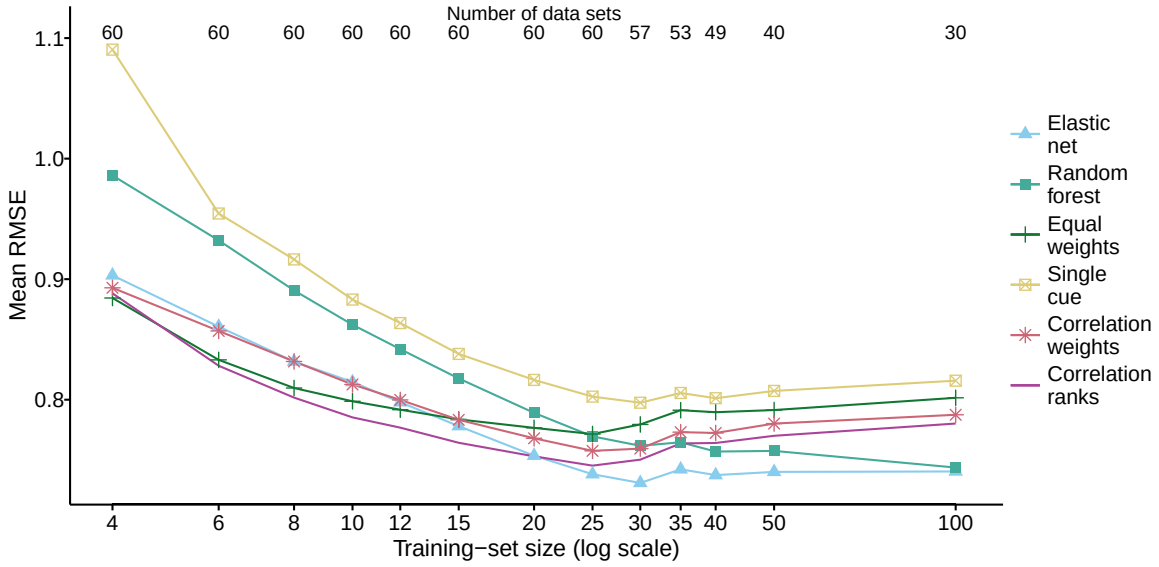

Figure 2: Learning curves across 60 data sets. The number of large-enough data sets per training-set size is indicated on top of the graph. Mean RMSEs for OLS are beyond the plot range and have not been plotted.

SAT data set shown in Figure 3, which is one of the few data sets for which both equal weights and correlation ranks perform poorly, but for which single cue is the best-performing algorithm across the entire learning curve.

The data sets *Prostate* and *Diabetes* have been used to illustrate the favorable prediction performance of the elastic net and other sophisticated regression models in the past (Tibshirani, 1996; Efron et al., 2004; Zou and Hastie, 2005). Figure 3 shows that correlation weights outperformed the elastic net in both data sets in training-set sizes smaller than 30.

On the mean learning curve, correlation ranks outperformed all other simple models across the entire curve. However, on individual data sets, correlation ranks was often outperformed by one or more of the other simple models. In fact, in almost all data sets, the learning curve of correlation ranks lay in between those of equal weights and correlation weights, independent of which of the two latter models performed better. This confirms the intuition that correlation ranks is an intermediately complex model that is able to perform well in situations of scarce information (similar to equal weights) but can also exploit the benefits of weighting predictors differently when there is enough information to reliably estimate the ranking of predictors.

Both equal-weights and single-cue regression share the property of performing either very well or very poorly on many data sets. Single-cue regression was the second-worst or worst model over the entire range of training-set sizes in 23 of the 60 data sets. Yet it was also the best-performing model across the entire learning curve in *SAT* and across large parts of the learning curve in *Body fat*. Equal weights outperformed all other models across the entire learning curve in data sets *Bone, Fuel, Pinot, Reactor, Rent*, and *Wage*. But it

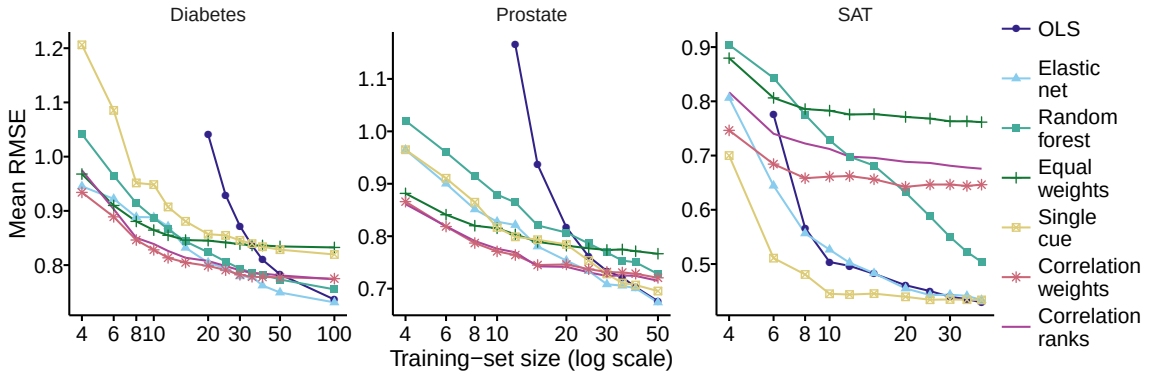

Figure 3: Individual learning curves for data sets *Diabetes*, *Prostate*, and *SAT*. Mean RMSEs above 1.2 have not been plotted. OLS = oridnary least squares.

was by far the worst model on large parts of the learning curve in *Diamond, Excavator, Fish*, and *SAT*.

Among benchmark models, OLS was outperformed by random forest regression and the elastic net, both on average and individually on most of the data sets. Elastic net generally outperformed random forest regression, especially on small training sets.

5. Discussion

Our analysis shows that simple regression models, for example, equal-weights regression, routinely outperform not only multiple linear regression but also state-of-the-art regression models, especially on small training sets.

None of the simple models we examined predicted well in all data sets. But in nearly all data sets, there was at least one simple model that predicted well.

Because OLS has severe estimation difficulties with small training sets, it would be reasonable to expect simple regression models to perform better than OLS on small training sets. However, we did not expect the simple models to be able to compete with state-of-the-art regularized linear models such as the elastic net.

Regularized linear models attempt to optimize prediction accuracy by searching through a possibly infinite-dimensional hypothesis space of linear models, ranging from a sparse linear model to the full, complex OLS solution. All simple models considered here are special cases of the linear regression model. Even though we tested only four of them, these simple models could sometimes outperform the carefully-optimized elastic net. These results indicate that it may be possible to substantially reduce the size of the hypothesis space of linear models needed to make good inferences. In other words, it is possible to make good inferences based on simple models if one only knows which simple model to choose.

Future work could examine models that adaptively choose between a few but maximally different simple models. For example, a model that chooses between single cue, equal weights, and correlation ranks using only information in the training data could be a fast

and robust alternative to current state-of-the-art models, while being computationally less challenging. The main question will be whether this algorithm can choose an appropriate simple model on the basis of only a small number of training examples.

An important research direction is to examine whether people can intuitively pick an appropriate simple model for a given problem. Such a finding may explain how people often make good decisions despite their bounded cognitive capacities.

Acknowledgments

We would like to thank Marcus Buckmann and three anonymous reviewers for comments on earlier drafts.

References

F. Hutton Barron and Bruce E. Barrett. Decision Quality Using Ranked Attribute Weights. *Management Science*, 42(11):1515–1523, 1996.

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

Henry Brighton. Robust Inference with Simple Cognitive Models. In *AAAI Spring Symposium: Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems*, pages 17–22, 2006.

Jean Czerlinski, Gerd Gigerenzer, and Daniel G. Goldstein. How good are simple heuristics? In *Simple heuristics that make us smart*, pages 97–118. Oxford University Press, 1999.

Jason Dana and Robyn M. Dawes. The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, 29(3): 317–331, 2004.

Clintin P. Davis-Stober, Jason Dana, and David V. Budescu. A Constrained Linear Estimator for Multiple Regression. *Psychometrika*, 75(3):521–541, 2010.

Robyn M. Dawes. The robust beauty of improper linear models in decision making. *American psychologist*, 34(7):571, 1979.

Robyn M. Dawes and Bernard Corrigan. Linear models in decision making. *Psychological Bulletin*, 81(2):95–106, 1974.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5): 1915–1953, 2009.

Bradley Efron, Trevor Hastie, Iain Johnstone, Robert Tibshirani, and others. Least angle regression. *The Annals of statistics*, 32(2):407–499, 2004.

Andrew S. C. Ehrenberg. How Good Is Best? *Journal of the Royal Statistical Society. Series A (General)*, 145(3):364, 1982.

Hillel J. Einhorn and Robin M. Hogarth. Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*, 13(2):171–192, 1975.

Jerome Friedman, Trevor Hastie, Noah Simon, and Rob Tibshirani. glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models, April 2015.

Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group. *Simple heuristics that make us smart*. Oxford University Press, USA, 1999.

Andreas Graefe. Improving forecasts using equally weighted predictors. *Journal of Business Research*, 68(8):1792–1799, 2015.

Arthur E. Hoerl and Robert W. Kennard. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1):55, 1970.

Robert C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine learning*, 11(1):63–90, 1993.

James E. Laughlin. Comment on" Estimating coefficients in linear models: It don't make no nevermind.". *Psychological Bulletin*, 85(2):247–253, 1978.

Andy Liaw and Matthew Wiener. Classification and regression by randomForest. *R news*, 2(3):18–22, 2002.

Alexander D. Lovie and Patricia Lovie. The flat maximum effect and linear scoring models for prediction. *Journal of Forecasting*, 5(3):159–168, July 1986.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

Frank L. Schmidt. The Relative Efficiency of Regression and Simple Unit Predictor Weights in Applied Differential Psychology. *Educational and Psychological Measurement*, 31(3): 699–714, 1971.

Özgür Şimşek. Linear decision rule as aspiration for simple decision heuristics. In *Advances in Neural Information Processing Systems*, pages 2904–2912, 2013.

Özgür Şimşek and Marcus Buckmann. Learning From Small Samples: An Analysis of Simple Decision Heuristics. In *Advances in Neural Information Processing Systems*, pages 3141–3149, 2015.

Robert Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

Detlof von Winterfeldt and Ward Edwards. Costs and payoffs in perceptual research. *Psychological Bulletin*, 91(3):609, 1982.

Howard Wainer. Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin*, 83(2):213–217, 1976.

Howard Wainer. On the sensitivity of regression and regressors. *Psychological Bulletin*, 85 (2):267–273, 1978.

Niels G. Waller and Jeff A. Jones. Correlation Weights in Multiple Regression. *Psychometrika*, 75(1):58–69, 2009.

Alexander G. Wesman and George K. Bennett. Multiple regression vs. simple addition of scores in prediction of college grades. *Educational and Psychological Measurement*, 1959.

Samuel S. Wilks. Weighting systems for linear functions of correlated variables when there is no dependent variable. *Psychometrika*, 3(1):23–40, 1938.

Hui Zou and Trevor Hastie. Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B*, 67:301–320, 2005.

Proceedings of Machine Learning Research 58 (2016) 19-30

Towards Implementable Prescriptive Decision Making

Miroslav Kárný SCHOOL@UTIA.CAS.CZ
Institute of Information Theory and Automation, the Czech Academy of Sciences, POB 18, 182 08 Prague, Czech Republic

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

The need for inspecting (ir)rationality in decision making (DM) — the observed discrepancy between real and prescriptive DMs — stems from omnipresence of DM in individuals' and society life. Active approaches try to diminish this discrepancy either by changing behaviour of participants (DM subjects) or modifying prescriptive theories as done in this text. It provides a core of unified merging methodology of probabilities serving for knowledge fusion and information sharing exploited in cooperative DM. Specifically, it unifies merging methodologies supporting a flat cooperation of interacting self-interested DM participants. They act without a facilitator and they are unwilling to spare a non-negligible deliberation effort on merging. They are supposed to solve their DM tasks via the fully probabilistic design (FPD) of decision strategies. This option is motivated by the fact that FPD is axiomatically justified and extends standard Bayesian DM.

Merging is a supporting DM task and is also solved via FPD. The proposed merging formulation tries to be as general as possible without entering into technicalities of measure theory. The results generalise and unify earlier work and open a pathway to systematic solutions of specific, less general, problems.

Keywords: fully probabilistic design, distributed decision making, cooperation

1. Introduction

Decision making (DM), seen as a purposeful choice among available options, covers a substantial portion of human activities as well as of institutions and devices created by people. DM almost always runs under uncertainty and with incomplete knowledge. This makes DM, by its nature optimising, extremely demanding on cognitive resources of any participant — the DM subject. Among a range of DM methodologies those based on Savage's DM concept, Savage (1954), seem to be the most promising and they are permanently generalised and refined. The fully probabilistic design of decision strategies (FPD), Kárný (1996); Kárný and Guy (2006); Kárný and Kroupa (2012), on which this paper relies, is of a Savage's type. The space limitations prevent us to discuss very rich related work even immediate predecessors as Kárný et al. (2009); Sečkárová (2013). The next characterisation of the used FPD is an exception enforced by its weak penetration to DM communities.

FPD, as all Bayesian solution, models probabilistically the closed-loop behaviour — the collection of all considered and opted variables within the DM task. Unlike its predecessors, FPD probabilistically describes the DM preferences through an ideal closed-loop model. FPD selects the optimal strategy — the sequence of randomised decision rules mapping the

knowledge on actions — minimising the KLD divergence, Kullback and Leibler (1951), of the closed-loop model to its ideal counterpart.

DM is always performed with limited resources — the time-span devoted to the solved DM task, the extent of the manageable knowledge, physical and computational resources. This naturally led to a division of DM tasks. In it, the involved participants (individuals or groups formed both by people, technical and organisational tools they use) solve smaller DM sub-tasks and select actions — irreversibly implemented decisions. This allows groups of participants to function but at substantial costs expended on cooperation — goods and knowledge exchanging, sharing, making concessions with respect to participant's aims, etc.

Governance of the cooperation process by a participant at a higher hierarchical level (facilitator, coordinator, institutional or legal authority, etc.) is a DM on its own. It can make the multi-participants' DM efficient but with increasing complexity of the addressed DM tasks demands on governance resources quickly increase and its efficiency strongly drops. Due to this, the need arises for distributed DM (almost) without a facilitator — a sort of democratic scenario arises, Kárný and Guy (2004).

The democratic scenario essentially lets individual DM subjects to act in a selfish way. The selfish (self-centered) participant cares about her own DM tasks only. She possibly cooperates with her neighbours — the participants with which she interacts and about which she is able and willing to care. The DM aims, however, persist: each participant tries to optimally reach her "personal" DM objectives under the given circumstances.

Note that the term selfish lacks here a moral dimension. For instance, it can be perceived as a quite positive if the care about societal welfare is adopted by the participant as her personal aim. Also note that real limited cognitive resources of any participant imply that the number of her neighbours will remain relatively small.

This paper aims to equip any selfish participant having limited cognitive resources with a tool supporting her DM by taking into account influence of neighbours. In the quest for an applicable prescriptive DM methodology, the constructed "advisor": i) is to be impersonal and implementable as a feasible algorithm, which adds as little as possible (ideally none) additional options (parameters, tuning knobs); ii) must not require from the supported participants more than what they handle anyway; iii) must offer its support in a way understandable to individual participants: as a by-product this guarantees privacy of the respective participants; iv) must let the supported participants decide whether they accept the advise or not.

The proposed solution assumes that all involved participants use FPD as the tool for solving their DM tasks. This is the most general prescriptive DM methodology, which is feasible for realistic but sufficiently small DM tasks. It also covers Bayesian participants.

Section 2 prepares the problem formulation and solution presented in Section 3. The solution depends on unwanted options (parameters), which are unambiguously specified in Section 4. The solution operates on joint probabilities acting on the same collection of variables. Section 5 describes the way how to cope with this restrictive condition. Section 6 adds remarks on conversions of non-probabilistic elements into the merged probabilities, recommends when to accept the gained advices and outlines open problems.

2. Preliminaries

We use the following notions and conventions.

DM subjects: A participant $\mathcal{P}$ belongs to a group of her neighbours $(\mathcal{P}_p)_{p\in p^\star}$, $|p^\star|<\infty$. The term advisor $\mathcal{A}$ refers to the cooperation-supporting algorithm serving to this group.

Sets, mappings, finite collections: Sets of entities $a, X, \mathsf{R}, \ldots$ are denoted $a^\star, X^\star, \mathsf{R}^\star, \ldots$ $|X^\star|$ means cardinality of the set $X^\star$. Sets are subsets of separable spaces, typically, finite sets of integers, finite-dimensional real spaces or sets of probability densities (pd, Radon-Nikodým derivatives with respect to counting or Lebesgue's measure — both denoted $\mathrm{d}\bullet$). Pds and other mappings are distinguished by $\mathsf{san\ serif\ fonts}$. A finite collection of entities, say real scalars $(\lambda_p)_{p\in p^\star}$, is often referred as $\lambda=(\lambda_p)_{p\in p^\star}$. Inequalities like $\lambda>0$ are then understood component-wise.

Behaviour, ignorance, action, knowledge: Each supported participant $\mathcal{P}_p$, $p\in p^\star$, operates on a specific (closed-loop) *behaviour* $b_p\in b_p^\star$, which is adopted name for all variables that $\mathcal{P}_p$ considers, opts or knows. The opted action splits the behaviour as follows[1]

$$\begin{aligned} b_p &= (g_p,a_p,k_p)=(\text{ignorance, action, knowledge})\text{ behaviour parts} \\ &=(\text{considered but inaccessible, opted by }\mathcal{P}_p,\text{ used for the action choice})\text{ parts.}\end{aligned} \tag{1}$$

The performed minimisation of any *expected* loss over randomised decision rules reduces to the minimisation of the expected loss conditioned on the knowledge available for the action choice, Berger (1985). This allows us to simplify presentation by (mostly) not spelling explicitly the used knowledge in conditions of the involved pds. This formally reduces behaviours to

$$\begin{aligned} b_p &= (g_p,a_p)=(\text{ignorance, action})\text{ behaviour parts} \\ &=(\text{considered but inaccessible, opted by }\mathcal{P}_p)\text{ parts, }\ p\in p^\star.\end{aligned} \tag{2}$$

Simplifying assumptions: The following concessions from full generality will be made.

- The cardinalities of the behaviour sets of all participants are finite

$$b_p\in b_p^\star=\{1,\ldots,|b_p^\star|\},\ 2\le|b_p^\star|<\infty,\quad p\in p^\star. \tag{3}$$

  The theory of the numerical representation of DM preferences, Debreu (1954), implies that the considered, numerically representable, behaviour spaces have to be separable. Then, the considered finite sets of behaviours can be seen as images of finite projections (discretisations) of the underlying separable spaces of infinite cardinalities.

- Static DMs are considered. Each participant selects a single action, i.e. she selects and applies a single randomised decision rule described by a pd $\mathsf{r}(a_p)=\mathsf{r}(a_p|k_p),\ p\in p^\star$.

1. While terms "knowledge" and "action" (an irreversibly adopted decision) are common and well-accepted, the unusual term "ignorance" (linguistically opposite to knowledge) is often felt as inappropriate. We ask patient reader to take it as technical term describing the part of the closed-loop behaviour differing from the action and knowledge.

FPD: Each participant $\mathcal{P}_p$, $p \in p^\star$, within a group of neighbours deals with her closed-loop model, see e.g. Kárný and Guy (2006), which is a joint pd,

$$\underbrace{\mathsf{c}_{\mathsf{r}_p}(b_p)}_{closed-loop\,model} = \underbrace{\mathsf{m}(g_p|a_p)}_{environment\,model} \times \underbrace{\mathsf{r}(a_p)}_{decision\,rule}, \quad b_p \in b_p^\star. \tag{4}$$

The factorisation (4) is implied by the chain rule for pds. The first factor on the right-hand side relates action to ignorance, i.e. to the considered but (yet) unknown reaction of the participant's environment. This motivates its interpretation. The second factor already has been recognised as the model of the decision rule.

$\mathcal{P}_p$ applying FPD possesses a preference-expressing ideal counterpart $\mathsf{c}_{Ip}(b_p)$ of $\mathsf{c}_{\mathsf{r}_p}(b_p)$, which determines her FPD-optimal decision rule $\mathsf{r}_{\mathrm{O}p} = \mathsf{r}_{\mathrm{O}p}(a_p)$

$$\mathsf{r}_{\mathrm{O}p} \in \mathrm{Arg}\min_{\mathsf{r}_p \in \mathsf{r}_p^\star} \int_{b_p^\star} \mathsf{c}_{\mathsf{r}_p}(b_p) \ln\left(\frac{\mathsf{c}_{\mathsf{r}_p}(b_p)}{\mathsf{c}_{Ip}(b_p)}\right) \mathrm{d}b_p = \mathrm{Arg}\min_{\mathsf{r}_p \in \mathsf{r}_p^\star} \mathsf{KL}(\mathsf{c}_{\mathsf{r}_p}||\mathsf{c}_{Ip}),\ p \in p^\star, \tag{5}$$

i.e. the optimal decision rule $\mathsf{r}_{\mathrm{O}p}$ minimizes the KLD $\mathsf{KL}(\mathsf{c}_{\mathsf{r}_p}||\mathsf{c}_{Ip})$ of $\mathsf{c}_{\mathsf{r}_p}$ from c_{Ip}.

3. Merging Problem Formulation and Solution

We assume that a participant seeks for support and her abilities delimit a group of neighbours. This defines a group of supported participants. Any participant can be a member of many groups, each with its advisor. Groups act in an asynchronous way and advices are offered for exploitation when created. The following design concerns a fixed group with a fixed knowledge processed by the group members having a single fixed advisor $\mathcal{A}$.

The required support: Each group member $\mathcal{P}_p$, $p \in p^\star$, provides the advisor $\mathcal{A}$ her closed-loop model with the aim obtaining an advice about a non-void factor of her closed-loop model (4), which she is willing to change according to the $\mathcal{A}$ advice.

The group behaviour, $\mathcal{A}$ action: The group behaviour[2] $B \in B^\star$ has the structure

$$\begin{aligned}
B &= (G, A) = (\text{ignorance}, \text{action}) \text{ of the group behaviour parts} \\
G &= \text{the group ignorance consists of behaviours (2) of all participants} \\
G^\star &= \cup_{p \in p^\star} b_p^\star \\
A &= \mathsf{A}(G) = \text{the } \mathcal{A} \text{ action is a pd on } G^\star \text{ merging pds } \mathsf{c}_{\mathsf{r}} = (\mathsf{c}_{\mathsf{r}_p})_{p \in p^\star} \\
\mathsf{A}^\star &= \text{all pds with the domain } G^\star \text{ of a finite cardinality} \\
K &= \text{the } \mathcal{A} \text{ knowledge consists of the pds } \mathsf{c}_{\mathsf{r}} \text{ and possibly } \mathsf{c}_I.
\end{aligned} \tag{6}$$

Verbally, the advisor action A is the pd $\mathsf{A}(G) = \mathsf{A}(G|K)$ modelling the group ignorance G formed by $(g_p, a_p)_{p \in p^\star}$ — the advice is offered to any participant $\mathcal{P}_p$ before she makes her action a_p. Knowledge K is insufficient for a unique specification of the pd $\mathsf{A} \in \mathsf{A}^\star$, and a randomised decision rule $\mathsf{R}(\mathsf{A}) = \mathsf{R}(\mathsf{A}|K)$ is to be designed. This is the factor of the group closed-loop model C_R

$$\mathsf{C}_\mathsf{R}(B) = \mathsf{C}_\mathsf{R}(G|A)\mathsf{C}_\mathsf{R}(A) = \mathsf{A}(G)\mathsf{R}(\mathsf{A}). \tag{7}$$

2. The mathematical entities related to the supported participants are marked by small letters while the group quantities, handled by $\mathcal{A}$, are denoted by their capital counterparts.

The first equality in (7) uses the definition of the group behaviour $B = (G, A) \in (G^\star, A^\star)$ (6) and the chain rule for pds. The second equality in (7) is implied by the definition of the advisor action $A = \mathsf{A}$ and of the decision rule R. The knowledge implicit in conditions of processed pds (cf. simplified version (2) of (1) is uninfluenced by the optimised rule R.

The desired merger: Let $\mathsf{e}_p(G)$ be a model of the group behaviour $G \in G^\star$ from the $\mathcal{P}_p$ point of view. Section 5 discusses the construction of these pds from the processed pds $\mathsf{c}_\mathsf{r}(G) = (\mathsf{c}_{\mathsf{r}_p}(b_p))_{p \in p^\star}$. It contains factors, which the cooperating participants are ready to replace according to the $\mathcal{A}$ advice. Of course, $\mathcal{P}_p$ may be inclined to accept the replacement if the *merger* — the advisor action — $A = \mathsf{A}$ is not too far from e_p. The approximation quality *from the participant view point* should be measured by the KLD, Bernardo (1979); Kárný and Guy (2012),

$$\mathsf{KL}(\mathsf{e}_p||\mathsf{A}) = \int_{G^\star} \mathsf{e}_p(G) \ln\left(\frac{\mathsf{e}_p(G)}{\mathsf{A}(G)}\right) \mathrm{d}G,\ p \in p^\star. \tag{8}$$

The decision rule R *is to provide mergers* $\mathsf{A} \in \mathsf{A}^\star$*, which make the KLD (8) small* $\forall p \in p^\star$*.*

The supporting DM task: The choice of R is the DM task, which is here formulated as FPD of the advisor decision rule entering the group closed-loop model (7). The ideal group closed-loop model described by a pd $\mathsf{C}_I(B)$ quantifies the $\mathcal{A}$ DM aim

$$\mathsf{C}_I(B) = \mathsf{C}_I(G|A)\mathsf{C}_I(A) = \mathsf{A}(G)\mathsf{R}_I(\mathsf{A}),\ G \in G^\star,\ \mathsf{A} \in \mathsf{A}^\star. \tag{9}$$

The first equality in (9) follows the definition of the group behaviour (6) and the chain rule for pds. The second one expresses the wish to get such a merger A, which describes well the group ignorance $G \in G^\star$. The ideal decision rule $\mathsf{R}_I(\mathsf{A})$ is one of the design options (tuning knob) through which the merging aim is fed into the solved group DM task.

For options (7), (9), the optimal decision rule $\mathsf{R}_\mathrm{O} = \mathsf{R}_\mathrm{O}(\mathsf{A}),\ \mathsf{A} \in \mathsf{A}^\star$, is

$$\mathsf{R}_\mathrm{O} \in \mathrm{Arg}\min_{\mathsf{R} \in \mathsf{R}^\star} \mathsf{KL}(\mathsf{C}_\mathsf{R}||\mathsf{C}_I) = \mathrm{Arg}\min_{\mathsf{R} \in \mathsf{R}^\star} \mathsf{KL}(\mathsf{R}||\mathsf{R}_I), \tag{10}$$

where the equality follows from cancelling the common factor A and integrating out the group ignorance $G \in G^\star$.

The set of admissible DM rules: The optimisation task (10) is determined by the set $\mathsf{R}^\star$ of admissible DM rules R and by the ideal decision rule R_I. They are gradually selected with the selection finalised in Section 4.

The KLD $\mathsf{KL}(\mathsf{R}||\mathsf{R}_I)$ is finite iff the support of R

$$\mathrm{supp}[\mathsf{R}] = \{\mathsf{A} \in \mathsf{A}^\star :\ \mathsf{R}(\mathsf{A}) > 0\} \tag{11}$$

is included in the support of R_I. Thus, meaningful choices for $\mathsf{R}^\star$ are

$$\emptyset \neq \mathsf{R}^\star \subset \{\mathsf{R} :\ \mathrm{supp}[\mathsf{R}] = \mathrm{supp}[\mathsf{R}_I]\}. \tag{12}$$

A specific choice of the set $\mathsf{R}^\star$ should primarily reflect the already formulated wish to deal with such Rs, which generate mergers $\mathsf{A} \in \mathsf{A}^\star$, which make KLDs (8) small. This singles out the rules R, for which the expectations (with respect to $\mathsf{R}(\mathsf{A})$) of divergences

$\mathsf{KL}(\mathsf{e}_p||\mathsf{A})$, $p \in p^\star$, (8) are finite and small. This wish and non-negativity of the inspected KLDs delimit

$$\begin{aligned} \mathsf{R}^\star &= \left\{\mathsf{R} : \int_{\mathsf{A}^\star} \mathsf{R}(\mathsf{A})\mathsf{KL}(\mathsf{e}_p||\mathsf{A})\mathrm{d}\mathsf{A} \right. \\ &= \left. \int_{\mathsf{A}^\star} \mathsf{R}(\mathsf{A}) \left[\int_{G^\star} \mathsf{e}_p(G) \ln\left(\frac{\mathsf{e}_p(G)}{\mathsf{A}(G)}\right) \mathrm{d}G\right] \mathrm{d}\mathsf{A} \le \phi_p < \infty,\ p \in p^\star \right\}, \end{aligned} \tag{13}$$

where the positive scalars $\phi = (\phi_p)_{p\in p^\star}$ should be chosen as small as possible. This intuitive wish is given the precise meaning in Section 4. The limits ϕ serve us a technical tool for determining the structure of the FPD-optimal decision rule R_O (10). Section 4 removes these undesirable tuning knobs of the advisor $\mathcal{A}$.

The next proposition determines the FPD-optimal merger selected from (13).

Proposition 1 (The Form of the FPD-Optimal Merger) *Let us consider a fixed ideal decision rule R_I and a non-empty set $\mathsf{R}^\star$ (13) of optional $\mathcal{A}$ decision rules. Then, the FPD-optimal decision rule $\mathsf{R}_O(\mathsf{A})$ solving (10) on this set is proportional to*

$$\begin{aligned} &\mathsf{R}_I(\mathsf{A}) \exp\left[\sum_{p\in p^\star} \lambda_p \int_{G^\star} \mathsf{e}_p(G) \ln\left(\mathsf{A}(G)\right) \mathrm{d}G\right] \\ = \ &\mathsf{R}_I(\mathsf{A}) \prod_{G\in G^\star} \left(\mathsf{A}(G)\right)^{\nu_\lambda(G)}, \quad \nu_\lambda(G) = \sum_{p\in p^\star} \lambda_p \mathsf{e}_p(G), \end{aligned} \tag{14}$$

where the non-negative Kuhn-Tucker multipliers $\lambda = (\lambda_p)_{p\in p^\star}$ are chosen so that inequalities in (13) are satisfied.

Proof We use the Kuhn-Tucker functional, Kuhn and Tucker (1951), respecting the constraints (13) on the optimised rules $\mathsf{R} \in \mathsf{R}^\star$ and rearrange it into the KLD of the optimised R to R_O (10), which is minimised by equating these arguments. The factor independent of A cancels and the final form of (14) uses simple operations relying on $|G^\star| < \infty$. □

4. Choice of the Tuning Knobs

The usefulness of the solution described by Proposition 1 is strongly influenced by the optional tuning knobs in them. They are gradually and unambiguously specified here.

The replacement of ϕ by the choice λ: A good advisor $\mathcal{A}$ should make the approximation of all processed pds $\mathsf{e} = (\mathsf{e}_p)_{p\in p^\star}$ in (13) as tight as possible. It primarily means that all inequalities are to be active and thus $\lambda > 0$. In this case, the choice of $\phi > 0$ becomes formally equivalent to the choice of $\lambda > 0$.

A desired *impartial* $\mathcal{A}$ must not prefer any $\mathcal{P}_p$. This implies the *basic requirement* on possible $\lambda > 0$, cf. Sečkárová (2015),

$$\int_{\mathsf{A}^\star} \mathsf{R}_\mathsf{O}(\mathsf{A})\mathsf{KL}(\mathsf{e}_p||\mathsf{A})\mathrm{d}\mathsf{A} = \Phi,\ \forall p \in p^\star, \tag{15}$$

where the finite constant Φ is common for all $p \in p^\star$.

This specify $|p^\star|-1$ conditions on $|p^\star|$ Kuhn-Tucker multipliers $\lambda>0$. The quest for *tightness* of the approximation (8) implies that the advisor $\mathcal{A}$, which uses the ideal decision rule R_I leading to a Φ in (15), is preferable against the advisor $\tilde{\mathcal{A}}$ with $\tilde{\mathsf{R}}_I$ leading to a $\tilde{\Phi}\geq\Phi$. This provides the needed $|p^\star|$-th condition for an unambiguous choice of $\lambda>0$.

The ideal decision rule R_I: The ideal decision rule $\mathsf{R}_I(\mathsf{A})$ is chosen as a finite mixture of Dirichlet pds $\mathsf{D}(\cdot|\cdot)$, which can arbitrarily-well approximate any $\mathsf{R}_I(\mathsf{A})$, Antoniak (1974),

$$\begin{aligned}\mathsf{R}_I(\mathsf{A}) &= \sum_{k\in k^\star}\alpha_k\mathsf{D}(\mathsf{A}|\nu_{Ik}),\ k^\star=\{1,\dots,|k^\star|\},\ |k^\star|<\infty,\quad\text{where}\\ \alpha\in\alpha^\star &= \Big\{\alpha_k\geq 0,\ \sum_{k\in k^\star}\alpha_k=1\Big\},\ \nu_{Ik}=(\nu_{Ik}(G))_{G\in G^\star}>0,\ k\in k^\star,\\ \mathsf{D}(\mathsf{A}|\nu) &= \frac{\prod_{G\in G^\star}\mathsf{A}(G)^{\nu(G)-1}}{\mathsf{B}(\nu)},\ \mathsf{B}(\nu)=\frac{\prod_{G\in G^\star}\Gamma(\nu(G))}{\Gamma\left(\sum_{G\in G^\star}\nu(G)\right)},\ \nu=(\nu(G))_{G\in G^\star},\end{aligned}\tag{16}$$

where the gamma function $\Gamma(x)=\int_0^\infty t^{x-1}\exp(-t)\mathrm{d}t,$ for $x>0$.

The ideal decision rule (16) leads to an optimal rule (14) with the Dirichlet mixture form

$$\mathsf{R}_{\mathrm{O}}(\mathsf{A})=\sum_{k\in k^\star}\alpha_k\mathsf{D}(\mathsf{A}|\nu_{Ik}+\nu_\lambda).\tag{17}$$

The specific choice of α, $|k^\star|$ and ν_I follows from the required impartiality of the advisor $\mathcal{A}$ and from the following simple uncontroversial requirement: An impartial $\mathcal{A}$ chooses its ideal decision rule beforehand for all possible $\mathsf{e}=(\mathsf{e}_p)_{p\in p^\star}$. It means that it has to process well even the special case $\mathsf{e}_1=\ldots=\mathsf{e}_{|p^\star|}\neq\nu_{Ik}$, $k\in k^\star$. In this case, the expected values

$$\int_{\mathsf{A}(G)^\star}\mathsf{A}(G)\mathsf{R}_{\mathrm{O}}(\mathsf{A}(G))\mathrm{d}\mathsf{A}(G)=\mathsf{e}_1(G)=\ldots=\mathsf{e}_{|p^\star|}(G),\quad\forall G\in G^\star,\tag{18}$$

represent the only meaningful option. The optimal rule (17) has the expected values

$$\int_{\mathsf{A}(G)^\star}\mathsf{A}(G)\mathsf{R}_{\mathrm{O}}(\mathsf{A}(G))\mathrm{d}\mathsf{A}(G)=\sum_{k\in k^\star}\alpha_k\frac{\nu_{Ik}(G)+\nu_\lambda(G)}{\sum_{G\in G^\star}(\nu_{Ik}(G)+\nu_\lambda(G))},\quad\forall G\in G^\star,\tag{19}$$

for which the equality (18) is reachable iff $(\nu_{Ik}(G))_{k\in k^\star,G\in G^\star}=0$. Thus, only identical improper components $\mathsf{D}(\mathsf{A}|\nu_{Ik}=0)$ in (16) meet (18). For them, the mixture (16) reduces to the single-component improper ideal decision rule $\mathsf{R}_I(\mathsf{A})=\mathsf{D}(\mathsf{A}|\nu_I=0)$. This ideal decision rule always gives the proper optimal rule R_{O} (14) for the considered $\lambda>0$, which makes $\nu_\lambda>0$. Indeed, in the linear combination ν_λ (14) at least one e_p assigns a positive value to any $G\in G^\star$ as $G^\star$ is delimited by this requirement.

The choice of Φ (15): The above considerations uniquely specified the improper ideal decision rule R_I and the form of the optimal rule R_{O}

$$\mathsf{R}_I(\mathsf{A})=\mathsf{D}(\mathsf{A}|\nu_I=0)\propto\prod_{G\in G^\star}\mathsf{A}(G)^{-1}\overset{(17)}{\Rightarrow}\mathsf{R}_{\mathrm{O}}(\mathsf{A})=\mathsf{D}(\mathsf{A}|\nu_\lambda).\tag{20}$$

It also provided $|p^\star|-1$ conditions (15) for $|p^\star|$ Kuhn-Tucker multipliers $\lambda>0$. Thus, it remains to specify the constant Φ in (15).

For a given $\nu_\lambda>0$ (14), and arbitrary $p\in p^\star$, the left-hand side of (15) reads

$$\begin{aligned}\Phi_p &= \int_{\mathsf{A}^\star}\mathsf{R}_{\mathsf{O}}(\mathsf{A})\mathsf{KL}(\mathsf{e}_p||\mathsf{A})\mathrm{d}\mathsf{A} \qquad (21)\\ &= -\sum_{G\in G^\star}\mathsf{e}_p(G)\int_{\mathsf{A}^\star}\mathsf{D}(\mathsf{A}||\nu_\lambda)\ln(\mathsf{A}(G))\mathrm{d}\mathsf{A}-\underbrace{\sum_{G\in G^\star}\mathsf{e}_p(G)\ln(\mathsf{e}_p^{-1}(G))}_{\mathsf{H}(\mathsf{e}_p)}\\ &= -\sum_{G\in G^\star}\mathsf{e}_p(G)\left[\Psi(\nu_\lambda(G))-\Psi\Big(\sum_{\tilde G\in G^\star}\nu_\lambda(\tilde G)\Big)\right]-\mathsf{H}(\mathsf{e}_p),\end{aligned}$$

where Ψ is the digamma function, the derivative of logarithm of the gamma function, Abramowitz and Stegun (1972). The formula for the expectation of $\ln(\mathsf{A}(G))$ with respect to Dirichlet pd is, e.g., in Kárný et al. (2006). The λ-independent summand $\mathsf{H}(\mathsf{e}_p)$ is the entropy of the processed pd e_p. For a simple presentation, let us denote[3]

$$\begin{aligned}\mu &= \sum_{p\in p^\star}\lambda_p,\quad \zeta_p=\frac{\lambda_p}{\sum_{p\in p^\star}\lambda_p},\quad \mathsf{e}_\zeta(G)=\sum_{p\in p^\star}\zeta_p\mathsf{e}_p(G)\ \overbrace{\Rightarrow}^{(14),(21)} \qquad (22)\\ \Phi_p &= -\int_{G^\star}\mathsf{e}_p(G)\Psi(\mu\mathsf{e}_\zeta(G))\mathrm{d}G+\Psi(\mu)-\mathsf{H}(\mathsf{e}_p).\end{aligned}$$

There, the pd e_ζ on $G^\star$ is the mixture of the pds supplied by the respective participants. Given (22), the choice of $\lambda>0$ is equivalent to the choice of the free $\mathsf{p}-1$ positive weighting probabilities $\zeta=(\zeta_p)_{p\in p^\star}$ and the scaling factor $\mu>0$. For a given pd $\zeta>0$, determined so that all $\Phi_p=\Phi$, the choice of Φ is equivalent to the choice of $\mu>0$.

The value Φ_p is a decreasing function of μ and its absolutely smallest value $-\mathsf{H}(\mathsf{e}_p)$ is reached for $\mu=\infty$. Thus, μ is found by solving (15) with the biggest $\mu<\infty$, for which its solution exist. With the introduced notations and performed evaluations, (15) gets the form of $|p^\star|-1$ equations for the probabilistic weights $\zeta>0$

$$\int_{G^\star}(\mathsf{e}_p(G)-\mathsf{e}_1(G))\Psi(\mu\mathsf{e}_\zeta(G))\mathrm{d}G=\mathsf{H}(\mathsf{e}_1)-\mathsf{H}(\mathsf{e}_p),\ \forall p\in p^\star, \qquad (23)$$

and the scalar $\mu>0$ is selected according to the above dictum.

Remark An analysis is needed for whether a solution of (23) exists, i.e. whether $\mathsf{R}^\star\neq\emptyset$. Also uniqueness is to be inspected and an efficient algorithm for its construction designed. All these important tasks are out of the paper scope.

5. Extension of Processed PDs c_r to PDs e Acting on Group Behaviour

The proposed merging works with the collection of pds $\mathsf{e}=(\mathsf{e}_p)_{p\in p^\star}$ defined over the group ignorance $G\in G^\star$, but participants $\mathcal{P}_p$, $p\in p^\star$, provide their closed-loop models $\mathsf{c}_{\mathsf{r}_p}$ assigning

3. The integral notation used hereafter underlines the conjecture that the *results* hold for $|B^\star|=\infty$.

probabilities to $b_p^\star$ only. The needed extension $\mathsf{c_r} = (\mathsf{c_r}_p)_{p\in p^\star} \to \mathsf{e} = (\mathsf{e}_p)_{p\in p^\star}$ is to be done the advisor $\mathcal{A}$. Altogether, $\mathcal{A}$ has to select a decision rule $\mathsf{R}_E : \mathsf{c_r} \to (\mathsf{A}, \mathsf{E})$, generating the action pair consisting of the derived merger A and the pd $\mathsf{E} \in \mathsf{E}^\star$ on the possible extensions $\mathsf{e} = (\mathsf{e}(G)_p)_{p\in p^\star, G\in G^\star}$ of $\mathsf{c_r} = ((\mathsf{c_r}_p(b_p))_{b_p\in b_p^\star})_{p\in p^\star}$. The subscript $_E$ stresses that the discussed decision rule extends that treated in previous sections.

The randomised rule $\mathsf{R}_E(\mathsf{A}, \mathsf{E})$ describes action pairs $(\mathsf{A}(G|\mathsf{E}), \mathsf{E}(\mathsf{e}))$ occurring in the chain-rule factorisation

$$\mathsf{R}_E(\mathsf{A}, \mathsf{E}) = \mathsf{R}_E(\mathsf{A}|\mathsf{E})\mathsf{R}_E(\mathsf{E}). \tag{24}$$

The (extended) group behaviour is $B = (G, (\mathsf{A}, \mathsf{E})) = (\text{group ignorance}, (\text{action pair}))$ parts and the closed-loop model

$$\mathsf{C}_{\mathsf{R}_E}(B) = \mathsf{C}_{\mathsf{R}_E}(G|\mathsf{A}, \mathsf{E})\mathsf{C}_{\mathsf{R}_E}(\mathsf{A}|\mathsf{E})\mathsf{C}_{\mathsf{R}_E}(\mathsf{E}) = \mathsf{A}(G|\mathsf{E})\mathsf{R}_E(\mathsf{A}|\mathsf{E})\mathsf{R}_E(\mathsf{E}), \tag{25}$$

where the plain chain rule gives the first equality and the second one directly follows from the definitions of individual factors. The ideal closed-loop model is

$$\mathsf{C}_{IE}(B) = \mathsf{A}(G|\mathsf{E})\mathsf{R}_{IE}(\mathsf{A}|\mathsf{E})\mathsf{R}_{IE}(\mathsf{E}), \tag{26}$$

where the chain rule is used and the wish that A should describe well group ignorance is again applied. With this, the FPD-optimal extended decision rule $\mathsf{R}_{\mathrm{O}E}(\mathsf{A}, \mathsf{E}) = \mathsf{R}_{\mathrm{O}E}(\mathsf{A}|\mathsf{E})\mathsf{R}_{\mathrm{O}E}(\mathsf{E})$ is the minimising argument in

$$\begin{aligned} &\min_{\mathsf{R}_E\in\mathsf{R}_E{}^\star} \mathsf{KL}(\mathsf{C}_{\mathsf{R}_E}||\mathsf{C}_{IE}) = \min_{\mathsf{R}_E\in\mathsf{R}_E{}^\star} \mathsf{KL}(\mathsf{R}_E||\mathsf{R}_{IE}) \qquad (27)\\ = &\min_{\mathsf{R}_E(\mathsf{E})} \int_{\mathsf{E}^\star} \mathsf{R}_E(\mathsf{E}) \left[\ln\left(\frac{\mathsf{R}_E(\mathsf{E})}{\mathsf{R}_{EI}(\mathsf{E})}\right) + \underbrace{\min_{\mathsf{R}_E(\mathsf{A}|\mathsf{E})} \int_{\mathsf{A}^\star} \mathsf{R}_E(\mathsf{A}|\mathsf{E}) \ln\left(\frac{\mathsf{R}_E(\mathsf{A}|\mathsf{E})}{\mathsf{R}_{IE}(\mathsf{A}|\mathsf{E})}\right) \mathrm{d}\mathsf{A}}_{\Phi(\mathsf{E})} \right] \mathrm{d}\mathsf{E}, \end{aligned}$$

where again $\mathsf{A}(G|\mathsf{E})$ has cancelled and G integrated out.

As demonstrated below, the treatment of the second summand in (27) reduces to that specified and optimised in previous sections. Thus, the ideal pd $\mathsf{R}_{IE}(\mathsf{E})$ determining the first summand only needs to be chosen. Its support has to allow only the extensions, which preserve the closed-loop models $\mathsf{c_r}$ provided by individual participants, i.e.

$$\mathrm{supp}[\mathsf{R}_E(\mathsf{E})] \;=\; \{\mathsf{E} : \mathrm{supp}[\mathsf{E}] = \{\mathsf{e} = (\mathsf{e}_p)_{p\in p^\star} : \mathsf{e}_p(G) = \mathsf{e}_p(G_{-p}|b_p)\mathsf{c_r}(b_p)\}\}, \tag{28}$$

where G_{-p} complements b_p to G, i.e. $G = (G_{-p}, b_p),\ \forall p \in p^\star$.

The following adopted leave-to-the-fate option, Kárný et al. (2006),

$$\mathsf{R}_{IE}(\mathsf{E}) = \mathsf{R}_E(\mathsf{E}) \text{ on } \mathrm{supp}[\mathsf{R}_E(\mathsf{E})] \tag{29}$$

respects that no other requirements exist with respect to $\mathsf{R}_E(\mathsf{E})$. Under (29), the minimised functional (27) is linear in the optimised $\mathsf{R}_E(\mathsf{E})$ and minimum is reached by the deterministic rule $\mathsf{R}_{\mathrm{O}E}(\mathsf{E})$ concentrated on a minimiser E_O of $\Phi(\mathsf{E})$ defined in (27).

For a fixed E concentrated on a point e in $\mathrm{supp}[\mathsf{E}]$ (28), the second minimisation in (27), defining $\Phi(\mathsf{E})$, coincides with the optimisation in Section 3, Proposition 1. Also, the choice of the tuning knobs, Section 4, is the same when taking into account the correspondence

$$\mathsf{R}(\mathsf{A}) \leftrightarrow \mathsf{R}_E(\mathsf{A}|\mathsf{E}),\ \mathsf{R}_I(\mathsf{A}) \leftrightarrow \mathsf{R}_{IE}(\mathsf{A}|\mathsf{E}). \tag{30}$$

This uniquely determines $\mathsf{R}_{IE}(\mathsf{A}|\mathsf{E}) = \mathsf{D}(\mathsf{A}|\nu_I = 0)$ and the optimal $\mathsf{R}_{OE}(\mathsf{A}|\mathsf{E}) = \mathsf{D}(\mathsf{A}|\nu_\lambda)$ and the function $\Phi(\mathsf{E})$ to be minimised with respect to free factors $(\mathsf{e}(G_{-p}|b_p))_{p\in p^\star}$ coincides with the common value Φ (15).

A general pd E in the support of $\mathsf{R}_{IE}(\mathsf{E})$ (28) is a convex combination of pds E concentrated on pds $\mathsf{e} \in \mathsf{e}^\star = \mathrm{supp}[\mathsf{E}]$. The complete analogy of reasoning made in Section 4 recommends the ideal decision rule $\mathsf{R}_{IE}(\mathsf{A}|\mathsf{E}) = \mathsf{D}(\mathsf{A}|\nu_I = 0)$. The set $\mathsf{e}^\star$ defined by constraints on possible e (28) is the convex set. Thus, the optimisation over $\mathsf{E} \in \mathsf{E}^\star$ reduces to optimisation over those concentrated on points $\mathsf{e} = (\mathsf{e}(G_{-p}|b_p)\mathsf{c}_{\mathsf{rp}}(b_p))_{p\in p^\star, G\in G^\star}$. This finalises the solution of the general merging case summarised now for reference purposes.

Proposition 2 (General Merging) *Let the closed-loop models of neighbours*[4] $\mathsf{c}_\mathsf{r}(G) = (\mathsf{c}_{\mathsf{r}_p}(b_p))_{p\in p^\star}$ *be given and the group ignorance* $G \in G^\star$ *be the concatenation of all variables occurring in* b_p, $p \in p^\star$, *while for each* $G \in G^\star$ *at least one* $\mathsf{c}_{\mathsf{r}_p}(b_p)$ *is positive. Then, the FPD-optimal (extended) advising rule of an impartial advisor* $\mathcal{A}$ *respecting (18) is described as follows. Among pds of the form* $\mathsf{e}_p(G) = \mathsf{e}(G_{-p}|b_p)\mathsf{c}_{\mathsf{r}_p}(b_p)$, $p \in p^\star$, $G = (G_{-p}, b_p)$, *find*

$$\min_{(\mathsf{e}(G_{-p}|b_p))_{p\in p^\star}} \Phi(\mathsf{e}) = \tag{31}$$

$$\min_{(\mathsf{e}(G_{-p}|b_p))_{p\in p^\star},\, \mu>0,\, \zeta\in\zeta^\star} \int_{G^\star} \mathsf{e}_1(G)\Psi(\mu\mathsf{e}_\zeta(G))\mathrm{d}G + \Psi(\mu) - \mathsf{H}(\mathsf{e}_1)$$

with the optimised $\mu > 0$ *and the probabilistic weights* $\zeta > 0$ *entering*

$$\mathsf{e}_\zeta(G) = \sum_{p\in p^\star} \zeta_p \mathsf{e}(G_{-p}|b_p)\mathsf{c}_{\mathsf{r}_p}(b_p) \text{ and}$$

$$\int_{G^\star} (\mathsf{e}_p(G_{-p}|b_p)\mathsf{c}_{\mathsf{r}_p}(b_p) - \mathsf{e}_1(G_{-1}|b_1)\mathsf{c}_{\mathsf{r}1}(b_1))\Psi(\mu\mathsf{e}_\zeta(G))\mathrm{d}G = \mathsf{H}(\mathsf{e}_1) - \mathsf{H}(\mathsf{e}_p),\ \forall p \in p^\star,$$

with $\Psi(\mu)$ *being digamma function and the entropy* $\mathsf{H}(\mathsf{e}_p)$ *defined*

$$\mathsf{H}(\mathsf{e}_p) = -\int_{G^\star} \mathsf{e}_p(G_{-p}|b_p)\mathsf{c}_{\mathsf{r}_p}(b_p)\ln(\mathsf{e}_p(G_{-p}|b_p)\mathsf{c}_{\mathsf{r}_p}(b_p))\mathrm{d}G. \tag{32}$$

The minimiser $(\mathsf{e}_O, \mu_O, \zeta_O)$ *determines the optimal advisory rule with* $\mathsf{R}_{OE}(\mathsf{E})$ *concentrated on* e_O *and* $\mathsf{R}_{OE}(\mathsf{A}|\mathsf{E}) = \mathsf{R}_{OE}(\mathsf{A}|\mathsf{e}_O) = \mathsf{D}(\mathsf{A}|\mu_O\mathsf{e}_{O\zeta_O})$, *with* $\mathsf{e}_{O\zeta_O} = \sum_{p\in p^\star} \zeta_{Op}\mathsf{e}_{Op}(G_{-p}|b_b)\mathsf{c}_{\mathsf{r}_p}(b_p)$.

6. Concluding Remarks

The contribution to a prescriptive DM theory, which respects constraints of real DM participants in knowledge sharing, is the main message brought by this paper. The proposed merging of probabilistic knowledge of neighbours is based on a systematic use of FPD and the impartiality requirement suppressing undesirable tuning knobs. Limited space prevents us in describing how to support participants in transforming their non-probabilistic knowledge into probabilities (crisp values as measures concentrated on them, marginal pds and deterministic relations, Sečkárová (2015), extended by minimum cross-entropy principle, Shore and Johnson (1980), re-interpretation of fuzzy rules as conditional pds, etc.)

The paper described how to construct advices but does not guide when accept them. It has, however, a relatively clear conceptual solution. After merging, the advisor $\mathcal{A}$ generates

4. The same procedure applies even when factors of $\mathsf{c}_{\mathsf{r}_p}(b_b)$ are communicated to $\mathcal{A}$ only.

an advice A_O either via sampling from $\mathsf{R}_{\mathsf{O}E}(\mathsf{A}|\mathsf{e}_\mathsf{O})$, Proposition 2, or by taking A_O as the expected value of the optimal advising rule. Then the advice $\mathsf{c}_{\mathcal{A}p}$ is simply the marginal pd of $\mathsf{A}_\mathsf{O}(G),\ G \in G^\star$

$$\mathsf{c}_{\mathcal{A}p}(b_p) = \int_{G^\star_{-p}} \mathsf{A}_\mathsf{O}(G_{-p}, b_p)\mathrm{d}G_{-p}. \tag{33}$$

It operates on variables known to the participant $\mathcal{P}_p$ and as such it is understandable to her. If the participant replaces $\mathsf{c}_{\mathsf{r}_p}$ or its factor by the corresponding factor of the advice A_O she gets the modified closed-loop model $\mathsf{c}_{\mathcal{A}p}$. Naturally, she takes it as a good advice iff

$$\mathsf{D}(\mathsf{c}_{\mathcal{A}p}||\mathsf{c}_{Ip}) \leq \mathsf{D}(\mathsf{c}_{\mathsf{r}_p}||\mathsf{c}_{Ip}). \tag{34}$$

Thus, $\mathcal{P}_p$ evaluates the quality of the advice according to her original selfish aim. The improvement (34) leading to advice acceptance is possible due to the fact that the advisor $\mathcal{A}$ operates on a wider knowledge and may respect supportive as well as competitive tendencies in the group interactions.

Specialisations to subclasses of our general solution, like Quinn et al. (2016), theoretical analysis of the proposed solution and its steps as well as the conversion of the conceptual solution into a practical tool are obvious directions to be addressed.

Good news are that preliminary brute-force numerical experiments (made without extensions only) indicate desirable properties of the merger: i) the average is the optimal merger iff the merged pds have identical entropy but its weight $\mu < |p^\star|$; ii) the weight ζ_p assigned to e_p increases with entropy $\mathsf{H}(\mathsf{e}_p)$, which makes the merger robust; iii) the Bayes' rule is the optimal merger if the merged pds are concentrated on crisp values.

It remains to discuss the concessions made, see Section 2. The obtained results are conjectured to be amendable to unbounded refinements of the discretisation mappings. This makes the assumption about finite cardinality of the behaviours'sets (3) non-restrictive. Also, dynamic DM is expected to be solvable in the exactly the same way as the considered static case. Note that whenever an external decision layer provides relative importance degrees of the respective participants within a group then the advisor can respect them by making the value Φ (15) the participant-specific.

Acknowledgments

The research reflected in this paper has been supported by GAČR GA16-09848S. A thorough feedback from anonymous reviewer help to improve significantly the presentation. Finally, I would like to apologise to all on whose shoulders this paper stands: starting from A. Wald, G.N. Saridis, V. Peterka up to communities dealing with KL control, compromise or goal programming. The available space did not allowed me even to mention them.

References

M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, 1972.

C.E. Antoniak. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics*, 2(6):1152–1174, 1974. ISSN 00905364.

J.O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer, New York, 1985.

J.M. Bernardo. Expected information as expected utility. *The Annals of Statistics*, 7(3): 686–690, 1979.

G. Debreu. Representation of a preference ordering by a numerical function. In R.M. Thrall, C.H. Coombs, and R.L. Davis, editors, *Decision Processes*, New York, 1954. Wiley.

M. Kárný. Towards fully probabilistic control design. *Automatica*, 32(12):1719–1722, 1996.

M. Kárný and T. V. Guy. On dynamic decision-making scenarios with multiple participants. In J. Andrýsek, M. Kárný, and J. Kracík, editors, *Multiple Participant Decision Making*, pages 17–28, Adelaide, 2004. Advanced Knowledge International.

M. Kárný and T. V. Guy. Fully probabilistic control design. *Systems & Control Letters*, 55 (4):259–265, 2006.

M. Kárný and T.V. Guy. On support of imperfect Bayesian participants. In T.V. Guy, M. Kárný, and D.H. Wolpert, editors, *Decision Making with Imperfect Decision Makers*, volume 28. Springer, Berlin, 2012. Intelligent Systems Reference Library.

M. Kárný and T. Kroupa. Axiomatisation of fully probabilistic design. *Information Sciences*, 186(1):105–113, 2012.

M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař. *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer, 2006.

M. Kárný, T. V. Guy, A. Bodini, and F. Ruggeri. Cooperation via sharing of probabilistic information. *Int. J. of Computational Intelligence Studies*, pages 139–162, 2009.

H.W. Kuhn and A.W. Tucker. Nonlinear programming. In *Proc. of 2nd Berkeley Symposium*, pages 481–492. Univ. of California Press, 1951.

S. Kullback and R. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–87, 1951.

A. Quinn, M. Kárný, and T.V. Guy. Fully probabilistic design of hierarchical Bayesian models. *Information Sciences*, 369:532–547, 2016. doi: http://dx.doi.org/10.1016/j.ins.2016.07.035.

L.J. Savage. *Foundations of Statistics*. Wiley, NY, 1954.

V. Sečkárová. *Cross-entropy based combination of discrete probability distributions for distributed decision making*. PhD thesis, Charles University in Prague, Faculty of Mathematics and Physics, Dept. of Probability and Mathematical Statistics., Prague, 2015. Submitted in May 2015. Successfully defended on 14.09.2015.

V. Sečkárová. On supra-Bayesian weighted combination of available data determined by Kerridge inaccuracy and entropy. *Pliska Stud. Math. Bulgar.*, 22:159–168, 2013.

J. Shore and R. Johnson. Axiomatic derivation of the principle of maximum entropy & the principle of minimum cross-entropy. *IEEE Tran. on Inf. Th.*, 26(1):26–37, 1980.

Proceedings of Machine Learning Research 58 (2016) 31-40

Forecasting Spot Oil Price Using Google Probabilities

Krzysztof Drachal KDRACHAL@WNE.UW.EDU.PL
Faculty of Economic Sciences
University of Warsaw
Warsaw, Poland

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

In this paper DMA (Dynamic Averaging Model) is expanded by adding certain probabilities based on Google Trends. Such a method is applied to forecasting spot oil prices (WTI). In particular it is checked whether a dynamic model including stock prices in developed markets, stock prices in China, stock prices volatility, exchange rates, global economic activity, interest rates, production, consumption, import and level of inventories as independent variables might be improved by including a certain measure of Google searches. Monthly data between 2004 and 2015 were analysed. It was found that such a modification leads to slightly better forecast. However, the weight ascribed to Google searches should be quite small. Except that it was found that even unmodified DMA produced better forecast than that based on futures contracts or naive forecast.

Keywords: Dynamic Model Averaging, Dynamic Model Selection, Google Trends, Internet Search Data, Spot Oil Price

1. Introduction

The oil market is quite a complex one. As a result, for practitioners there is no commonly accepted model of forecasting spot oil prices. Usually, forecasts based on futures contracts are used (Yang et al., 2002). Despite that obstacles, well-performing forecasts are crucial for the whole energy market, oil-importing and oil-exporting countries, macroeconomic forecasters, etc. (Alquist et al., 2013).

Generally, various forecasting methods on the oil market can be classified as: time series models, financial models, structural models, artificial neural networks based models, support vector machines models and qualitative method. The review of them can be found, for example, in papers of Drachal (2016a) or Behmiri and Pires Manso (2013).

In this paper quite a novel method, called DMA (Dynamic Model Averaging) is explored (Raftery et al., 2010). However, following Koop and Onorante (2014) the original method is slightly modified to include Internet searches. Recently, Internet search data were extensively applied in economics (Pavlicek and Kristoufek, 2015; Bangwayo-Skeete and Skeete, 2015; Scott and Varian, 2015; Choi and Varian, 2012; Dangl and Halling, 2012; Wu and Brynjolfsson, 2010; Choi and Varian, 2009; Schmidt and Vosen, 2009).

Actually, also for the oil market Google searches were applied (Li et al., 2015; Fantazzini and Fomichev, 2014), but not within the context of DMA. In particular, herein Google variables are not used as independent variables in the regression equation, but they are used to construct certain probabilities, which are used in computation of posteriori probabilities

in DMA recursive estimation. This is motivated by the assumption that a surge in searches about certain determinant might indicate the relevance of this variable in the model.

Indeed, many researches indicated that the impact of various determinants of the oil price might change in time (Aastveit and Bjornland, 2015; Zhang and Wu, 2014; Baumeister and Peersman, 2013). DMA seems to be a very good method in such a case. Actually, the interest in this method grows rapidly in finance (Naser, 2016; Aye et al., 2015; Bork and Moller, 2015; Baur et al., 2014; Koop and Korobilis, 2012; Nicoletti and Passaro, 2012).

Herein, in particular the following problem was addressed: Do Google Trends might somehow improve DMA forecast of the spot oil price? This question is positively answered further in the text.

2. Data

Now, a short review of potential oil price determinants is presented. This serves as an argument for the further data selection to the model.

According to Hotelling (1931) the price of non-renewable commodity should depend on the interest rate. Up to 1980s it was commonly agreed that the most important oil price determinant is OPEC decisions. Later attention was shifted to gross domestic product, stock market activity and exchange rates (Bernabe et al., 2004; Yousefi and Wirjanto, 2004), as well as, emerging markets (Basher et al., 2012).

Indeed, emerging markets were suggested to significantly impact the oil market (Li and Leung, 2011). There is some evidence that oil price changes between 2007 and 2008 might have happened due to the halt in the Chinese demand and supply (Kaufmann, 2011).

Within this context the global economic activity was also analysed. One of the approaches, which allows to use monthly frequency data is by the Kilian index (He et al., 2010; Kilian, 2009).

Moreover, Du and He (2015), as well as many other researchers focused on the impact of stock market volatility. Recently, much attention was put on speculative pressures (Carmona, 2015; Kilian and Murphy, 2014; Fattouh et al., 2013). Usually, they were measured by the level of inventories (Hamilton, 2009).

Much more extensive literature review on the oil price determinants within the context of a time-varying framework can be found, for example, in a paper by Drachal (2016a).

According to the presented literature review 10 potential oil price determinants were identified (Tab. 1). Strategic Petroleum Reserves were excluded from the level of inventories (Bu, 2014). Monthly data beginning on 2004 and ending on 2015 were taken. (Only for IMP the weekly data were aggregated to monthly one by taking the daily mean for the corresponding month.) This results in relatively short (in the context of DMA approach) data set. However, this is an obstacle that cannot be overcome, because Google Trends date back to 2004 only. (This is a common problem for all research with Google Trends statistics.)

If there is much easily available data for U.S., there is a lack of global time series. However, following, for example, Hamilton (2009) and Kilian and Murphy (2014), U.S. data can be taken as proxies.

The estimated forecast was compared to the one obtained by NFP (Alquist et al., 2013), i.e., 1-month NYMEX WTI futures prices (in USD).

Table 1: Data description

name	description
WTI	WTI spot price (in USD)
MSCI	MSCI World Index
TB3MS	U.S. 3-month treasury bill secondary market rate (in percentages)
KEI	Kilian index of global economy activity (Kilian, 2009)
TWEXM	trade weighted U.S. dollar index (Mar, 1973 = 100)
PROD	U.S. crude oil production (in 1'000 barrels)
IMP	daily average of U.S. crude oil import (in 1'000 barrels / day)
INV	U.S. total ending stocks of commercial crude oil (in 1'000 barrels)
VIX	implied volatility of S&P 500
CONS	total consumption of petroleum products in OECD (in quad BTU)
CHI	Shanghai Composite Index

It should be noticed that DMA does not require data to be stationary. On the other hand, it is desirable to normalize data, as this might significantly improve the outcomes (Drachal, 2016b).

In case of search terms the following ones were taken. For MSCI: "stocks", "developed markets", "msci index", "stock prices", "stock market", "stock quotes", "equity performance". For TB3MS: "market rates", "interest rates", "cpi", "inflation", "bond rates", "treasury bill", "fed", "libor". For KEI: "world economic activity", "gdp growth", "economic activity", "economy", "economic growth", "business cycle", "industrial production". For TWEXM: "exchange rates", "USD". For PROD: "oil production", "energy production", "oil supply", "opec". For IMP: "oil import". For INV: "oil inventories", "oil speculation". For VIX: "stock market volatility", "market stress", "market risk", "implied volatility", "vix", "volatility index". For CONS: "oil consumption", "energy consumption", "oil demand", "opec". For CHI: "china", "chinese economy", "china market", "shanghai composite index". For WTI: "wti", "oil price", "crude oil price". Then for each variable the mean of corresponding Google Trends was computed. This choice of search terms is quite general and arbitrary. It might be desirable to include more search terms. However, it was left for the further and much more extensive study, as even the choice of search terms is quite a challenging task (Stephens-Davidowitz and Varian, 2015).

All calculations were done in R (2015) software.

3. Methodology

Herein, just a brief sketch of DMA (Dynamic Model Averaging) and DMS (Dynamic Model Selection) is presented in order to explain the proposed modification with Google probabilities. The detailed explanation can be found in the original paper (Raftery et al., 2010).

Let there be m determinants. Then, $K = 2^m$ different models can be constructed, including the one with constant solely. Let us index time by $_t$, and let the dependent variable (the oil price) be y_t. Let x_t^k denote determinants in the kth model ($k = \{1, \ldots, K\}$). Notice, that for independent variables 1st lags were taken in all estimated models. Then, the state

space model is given by

$$y_t = x_t^k \theta_t^k + \epsilon_t^k \quad , \tag{1}$$

$$\theta_t^k = \theta_{t-1}^k + \delta_t^k \quad . \tag{2}$$

The regression parameters are denoted by θ_t^k and $\epsilon_t^k \sim N(0, V_t^k)$ and $\delta^k \sim N(0, W_t^k)$. Initially, $V_0^k := I$ and W_0^k is set according to the algorithm given by Raftery et al. (2010). I denotes the unit matrix. (Setting $V_0^k := I$ is reasonable, because data were normalized. In particular, let Y_t denote the core data. Then, the normalization is done with the formula $y_t := \frac{Y_t - \min(Y_0,\ldots,Y_t,\ldots)}{\max(Y_0,\ldots,Y_t,\ldots) - \min(Y_0,\ldots,Y_t,\ldots)}$.) Further, V_t^k is estimated by a recursive method of moments estimator, and W_t^k by the Kalman filter updating. This needs a certain forgetting factor $\lambda \in (0, 1]$ to be specified (Raftery et al., 2010; Dedecius et al., 2012). (Notice, that if $\lambda = 1$ there is no forgetting.)

The estimation is done recursively. First, it is set $\pi_{0|0,k} := \frac{1}{K}$. Then, it is proceeded with

$$\pi_{t|t-1,k} = \frac{(\pi_{t-1|t-1,k})^\alpha + c}{\sum_{i=1}^K (\pi_{t-1|t-1,i})^\alpha + c} \quad , \tag{3}$$

$$\pi_{t|t,k} = \frac{\pi_{t|t-1,k} f_k(y_t | y_0, y_1, \ldots, y_{t-1})}{\sum_{i=1}^K \pi_{t|t-1,i} f_i(y_t | y_0, y_1, \ldots, y_{t-1})} \quad . \tag{4}$$

The above equations contain the second forgetting factor $\alpha \in (0, 1]$ and $f_k(y_t | y_0, y_1, \ldots, y_{t-1})$ denotes the predictive density of the kth model at y_t given the data from the previous periods. $\pi_{t|t,k}$ are called posteriori inclusion probabilities. Also, a small constant is specified, for example, $c := K \cdot 10^{-3}$ in order to avoid reducing the probabilities to zero due to numerical approximations during computations.

Then, the DMA forecast is formulated in the following way

$$\widehat{y_t} = \sum_{k=1}^K \pi_{t|t-1,k} \widehat{y_t^k} \tag{5}$$

where $\widehat{y_t^k}$ is the forecast produced by the kth model.

According to Raftery et al. (2010), if $\alpha = 1 = \lambda$, there is no forgetting, and the method is still recursive but not dynamic.

However, let $\widehat{\pi_{t|t-1,k}} := \max_{i=\{1,\ldots,K\}} \{\pi_{t|t-1,i}\}$, where $\pi_{t|t-1,i}$ are computed as in Eq. 3. In DMS the Eq. 5 is modified in the following way

$$\widehat{y_t} = \widehat{\pi_{t|t-1,k}} \widehat{y_t^k} \quad . \tag{6}$$

Now, let g_i be the Google Trends variable corresponding to Internet search of ith term. This variable is between 0 and 100. So, it can be easily rescaled to fit between 0 and 1 and interpreted as Google probability (Koop and Onorante, 2014). (Notice that Google Trends correspond not to the absolute volume of Internet search terms, but to the relative one, i.e., relative to all Internet searches.) Now let

$$p_{t,k} := \prod_{i \in IN} g_i \cdot \prod_{j \in OUT} (1 - g_j) \quad , \tag{7}$$

where IN correspond to variables included in the kth model at time t, and OUT correspond to variables not included in the kth model at time t.

Then, Eq. 3 can be modified in the following way

$$\pi_{t|t-1,k} = \omega \cdot \frac{(\pi_{t-1|t-1,k})^{\alpha} + c}{\sum_{i=1}^{K}(\pi_{t-1|t-1,i})^{\alpha} + c} + (1-\omega) \cdot p_{t,k} \quad , \tag{8}$$

where ω is a parameter from $[0, 1]$. In this way DMA with Google probabilities and DMS with Google probabilities are obtained, by repeating the rest of procedures since Eq. 3 unmodified. If $\omega = 1$ then the standard DMA (DMS) is obtained. If $\omega = 0$ the procedure is highly changed and Google Trends play the whole role.

4. Results

First of all, it was imposed in all tested models that $\alpha = \lambda$, because these parameters correspond to the weight of information from the past that is put in the present. Therefore, this restriction can be nicely interpreted.

First, DMA models were estimated with variables as that in Tab. 1. Initially, models with $\omega = \{1, 0.75, 0.50, 0.25, 0\}$ and $\alpha = \{1, 0.99, \ldots, 0.90\} = \lambda$ were estimated. Unfortunately, in all cases the modification of DMA with Google probabilities lead to worse forecast (with respect to MSE), i.e., for all $\alpha = \lambda$ the smallest MSE was given by the model with $\omega = 1$, and decline in ω resulted in higher MSE for all $\alpha = \lambda$ fixed. The model with the best forecast (with respect to minimise MSE) was that with $\alpha = 0.93 = \lambda$. In such a case MSE was 0.0072153. Unfortunately, it is higher than that of the naive forecast, i.e., 0.0043214.

Therefore, having fixed $\alpha = 0.93 = \lambda$ more detailed examination was performed, i.e., with $\omega = \{0.99, 0.98, 0.97, 0.96, 0.95, 0.90, 0.85, 0.80\}$. It was found that for high ω slight forecast improvements are present. In particular, if $\omega = 0.99$ then the forecast of DMA with Google probabilities is approximately 2% better than the forecast of a standard DMA, i.e., it is 0.0070549. However, it is better than the forecast based on futures contracts (its MSE = 0.0118336).

Notice that, if $\mathcal{E}_t$ are residuals for Y_t, and ϵ_t are residuals for y_t, where $Y_t = a \cdot y_t + b$ (which corresponds to normalization), then $\mathcal{E}_t = a \cdot \epsilon_t$. So, the futures based forecast can be computed for the initial data and then its errors can be rescaled to be comparable with the estimated DMA/DMS models.

Still, this is a very small improvement. Moreover, the weight put to Google probabilities in Eq. 8 is marginal.

As a result, in this case it cannot be said that Google internet searches lead to significant improvement of DMA forecast performance. However, in case of relatively small $\alpha = \lambda$ even other (but still high ones) values of $\omega \neq 1$ lead to outperforming the forecast based on futures contracts. (Due to the limited space the details are not reported.)

Following Koop and Onorante (2014) it was also checked whether switching to DMS with Google probabilities would give better forecast. Unfortunately, for $\alpha = 0.93 = \lambda$ and $\omega = 0.99$ the forecast is slightly worse than in the corresponding case of DMA with Google probabilities (MSE = 0.0070922).

Some attempts were taken to improve the above model, for example, by reducing the set of oil price determinants up to those variables which met the following criterion. Notice

Table 2: MSE for DMA

$\omega \setminus \alpha = \lambda$	1	0.99	0.98	0.97	0.96	0.95
1	0.0043768	0.0042989	0.0042445	**0.0042226**	0.0042234	0.0042408
0.75	0.0059041	0.0055031	0.0051488	0.0048673	0.0046572	0.0045124
0.50	0.0089551	0.0079541	0.0070400	0.0062966	0.0057389	0.0053560
0.25	0.0121312	0.0104178	0.0088705	0.0076298	0.0067211	0.0061154
0	0.0148107	0.0125550	0.0105091	0.0088525	0.0076452	0.0068500

that, posteriori inclusion probabilities for every model which contains a given variable can be summed. Now, let reduce the set of oil price determinants up to the variables which posteriori inclusion probabilities are over 50% for most of the time since 2010. (Before 2010 it could have been observed that there is much variation of these probabilities, because the model "learns".) As a result, the set of determinants was reduced to MSCI, TB3MS, TWEXM and CHI. Unfortunately, it did not lead to any significant improvement of forecast.

Therefore, following Drachal (2016b) the following oil price determinants were considered: 1st and 2nd lags of WTI, MSCI, CHI, and 2nd lag of VIX. It can be seen that in such a case, if $\alpha = \{1, 0.99, \ldots, 0.90\} = \lambda$ and $\omega = \{1, 0.75, 0.50, 0\}$ the naive forecast was outperformed, but only if $\omega = 1$. (See Tab. 2.) The best forecast was obtained for $\alpha = 0.97 = \lambda$ and $\omega = 1$. (Due to the limited space herein only the most important part of results is presented.)

Switching to DMS method did not help, as for the estimations with the same spread of parameters α, λ and ω the minimal MSE was 0.0044905.

Further, $\omega = \{0.99, 0.98, \ldots, 0.90\}$ for fixed $\alpha = 0.97 = \lambda$ was examined (see Tab. 3). It can be seen that the best forecast was given by DMA with $\omega = 0.96$. However, this lead to only 2% improvement with respect to MSE in comparison with DMA without Google probabilities. Moreover, the weight put to Google probabilities is very small.

Finally, it was also checked if DMS model with $\alpha = 0.97 = \lambda$ and $\omega = 0.96$ would give better forecast. It was not so, as its MSE = 0.0046273.

5. Conclusions

It can be concluded that, even if Google probabilities can slightly improve DMA performance (by approx. 2%) the current findings are not much amazing, as this improvement is marginal. However, it can be expected that further investigation will provide better results. Due to the limited space herein, they will be presented elsewhere.

For further research it seems that it should be rather DMA framework explored than DMS. It would be also interesting to consider more indices from financial markets as oil price determinants, and omitting macroeconomics factors and supply-demand forces indicators. Such data are obtainable in higher frequencies, which allows to prepare a better data sample. Secondly, Google Trends statistics should be examined more carefully, and some other variations of Eq. 7 could be proposed.

Table 3: MSE for DMA with $\alpha = 0.97 = \lambda$

ω	
0.99	0.0041645
0.98	0.0041381
0.97	0.0041266
0.96	**0.0041249**
0.95	0.0041303
0.94	0.0041414
0.93	0.0041572
0.92	0.0041771
0.91	0.0042005
0.90	0.0042271
naive	0.0043214
futures	0.0118336

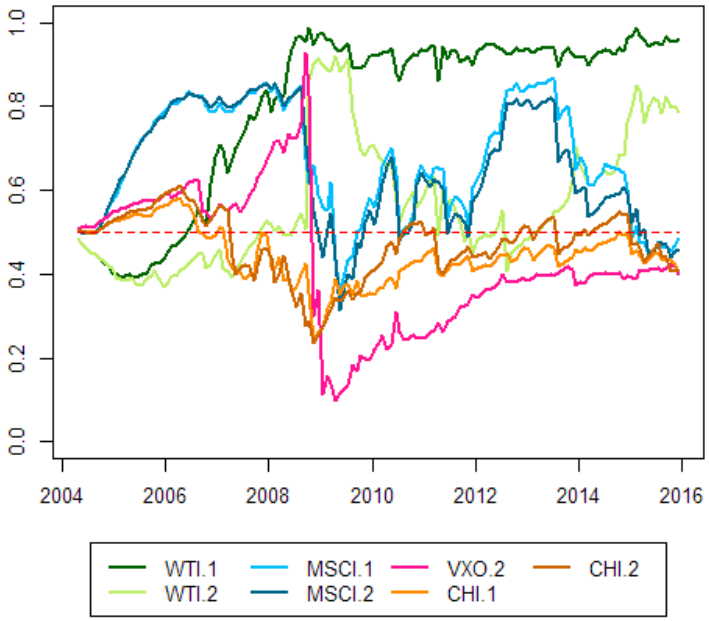

Figure 1: Posteriori probabilities for DMA with $\alpha = 0.97 = \lambda$ and $\omega = 0.96$

Acknowledgments

Research funded by the Polish National Science Centre grant under the contract number DEC-2015/19/N/HS4/00205.

Data sources

WTI, PROD, IMP, INV, CONS and NFP (U.S. Energy Information Administration)
http://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbbl_m.htm
http://www.eia.gov/dnav/pet/pet_move_wkly_dc_NUS-Z00_mbblpd_4.htm
http://www.eia.gov/dnav/pet/pet_stoc_wstk_dcu_nus_m.htm
http://www.eia.gov/dnav/pet/pet_pri_spt_s1_m.htm
http://www.eia.gov/countries
http://www.eia.gov/dnav/pet/pet_pri_fut_s1_m.htm
MSCI (MSCI World)
http://www.msci.com/end-of-day-data-search
TB3MS and TWEXM (Federal Reserve Bank of St. Louis)

http://research.stlouisfed.org/fred2/series/TB3MS
http://research.stlouisfed.org/fred2/series/TWEXBMTH
KEI (Kilian, 2009)
http://www-personal.umich.edu/∼lkilian/paperlinks.html
VIX (Chicago Board Options Exchange)
http://www.cboe.com/micro/buywrite/monthendpricehistory.xls
CHI (Stooq)
http://stooq.com
GOOGLE SERACH
http://www.google.com/trends

References

K.A. Aastveit and H.C. Bjornland. What drives oil prices? Emerging versus developed economies. *Journal of Applied Econometrics*, 30:1013–1028, 2015.

R. Alquist, L. Kilian, and R.J. Vigfusson. *Handbook of Economic Forecasting*, chapter Forecasting the price of oil. Elsevier, 2013.

G. Aye, R. Gupta, S. Hammoudeh, and W.J. Kim. Forecasting the price of gold using dynamic model averaging. *International Review of Financial Analysis*, 41:257–266, 2015.

P.F. Bangwayo-Skeete and R.W. Skeete. Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46: 454–464, 2015.

S.A. Basher, A.A. Haug, and P. Sadorsky. Oil prices, exchange rates and emerging stock markets. *Energy Economics*, 34:227–240, 2012.

C. Baumeister and G. Peersman. Time-varying effects of oil supply shocks on the US economy. *American Economic Journal: Macroeconomics*, 5:1–28, 2013.

D.G. Baur, J. Beckmann, and Czudaj R. Gold price forecasts in a dynamic model averaging framework - Have the determinants changed over time? *Ruhr Economic Papers*, 506, 2014.

N.B. Behmiri and J.R. Pires Manso. Crude oil price forecasting techniques: a comprehensive review of literature. *CAIA Alternative Investment Analyst Review*, 2:30–48, 2013.

A. Bernabe, E. Martina, J. Alvarez-Ramirez, and C. Ibarra-Valdez. A multi-model approach for describing crude oil price dynamics. *Physica A*, 338:567–584, 2004.

L. Bork and S.V. Moller. Forecasting house prices in the 50 states using dynamic model averaging and dynamic model selection. *International Journal of Forecasting*, 31:63–78, 2015.

H. Bu. Effect of inventory announcements on crude oil price volatility. *Energy Economics*, 46:485–494, 2014.

R. Carmona. *Commodities, Energy and Environmental Finance*, chapter Financialization of the commodities markets: A non-technical introduction. Springer, 2015.

H. Choi and H. Varian. Predicting initial claims for unemployment insurance using Google Trends. http://ssrn.com/abstract=1659307, 2009.

H. Choi and H. Varian. Predicting the present with Google Trends. *Economic Record*, 88: 2–9, 2012.

T. Dangl and M. Halling. Predictive regressions with time varying coefficients. *Journal of Financial Economics*, 106:157–181, 2012.

K. Dedecius, I. Nagy, and M. Karny. Parameter tracking with partial forgetting method. *International Journal of Adaptive Control and Signal Processing*, 26:1–12, 2012.

K. Drachal. Forecasting spot oil price in a dynamic model averaging framework - Have the determinants changed over time? *Energy Economics*, 60:35–46, 2016a.

K. Drachal. Determining time-varying drivers of spot oil price in a dynamic model averaging framework. Working Paper, 2016b.

L. Du and Y. He. Extreme risk spillovers between crude oil and stock markets. *Energy Economics*, 51:455–465, 2015.

D. Fantazzini and N. Fomichev. Forecasting the real price of oil using online search data. *International Journal of Computational Economics and Econometrics*, 4:4–31, 2014.

B. Fattouh, L. Kilian, and L. Mahadeva. The role of speculation in oil markets: What have we learned so far? *Energy Journal*, 34:20–30, 2013.

J.D. Hamilton. Causes and consequences of the oil shock of 2007-08. *Brookings Papers on Economic Activity*, 40:215–259, 2009.

Y. He, S. Wang, and K.K. Lai. Global economic activity and crude oil prices: A cointegration analysis. *Energy Economics*, 32:868–876, 2010.

H. Hotelling. The economics of exhaustible resources. *Journal of Political Economy*, 39: 137–175, 1931.

R.K. Kaufmann. The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy*, 39:105–115, 2011.

L. Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99:1053–1069, 2009.

L. Kilian and D.P. Murphy. The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, 29:454–478, 2014.

G. Koop and D. Korobilis. Forecasting inflation using dynamic model averaging. *International Economic Review*, 53:867–886, 2012.

G. Koop and L. Onorante. Macroeconomic nowcasting using Google probabilities. http://www.ecb.europa.eu/events/pdf/conferences/140407/OnoranteKoop_MacroeconomicNowcastingUsingGoogleProbabilities.pdf, 2014.

R. Li and G.C.K. Leung. The integration of china into the world crude oil market since 1998. *Energy Policy*, 39:5159–5166, 2011.

X. Li, J. Ma, S. Wang, and X. Zhang. How does Google search affect trader positions and crude oil prices? *Economic Modelling*, 49:162–171, 2015.

H. Naser. Estimating and forecasting the real prices of crude oil: A data rich model using a dynamic model averaging (DMA) approach. *Energy Economics*, 56:75–87, 2016.

G. Nicoletti and R. Passaro. Sometimes it helps: Evolving predictive power of spreads on GDP. *European Central Bank*, 1447, 2012.

J. Pavlicek and L. Kristoufek. Nowcasting unemployment rates with Google searches: Evidence from the Visegrad Group countries. *PLoS ONE*, 10, 2015. http://dx.doi.org/10.1371/journal.pone.0127084.

A.E. Raftery, M. Karny, and P. Ettler. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics*, 52:52–66, 2010.

T. Schmidt and S. Vosen. Nowcasting private consumption: Survey-based indicators vs. Google trends. *Ruhr Economic Papers*, 155, 2009.

S. Scott and H. Varian. *Economic Analysis of the Digital Economy*, chapter Bayesian variable selection for nowcasting economic time series. National Bureau of Economic Research, 2015.

S. Stephens-Davidowitz and H. Varian. A hands-on guide to Google data. http://people.ischool.berkeley.edu/∼hal/Papers/2015/primer.pdf, 2015.

L. Wu and E. Brynjolfsson. The future of prediction: How Google searches foreshadow housing prices and sales. Technical report, MIT, 2010.

C.W. Yang, M.J. Hwang, and B.N. Huang. An analysis of factors affecting price volatility of the us oil market. *Energy Economics*, 24:107–119, 2002.

A. Yousefi and T.S. Wirjanto. The empirical role of the exchange rate on the crude-oil price information. *Energy Economics*, 26:783–799, 2004.

Y.-J. Zhang and G. Wu. Does China factor matter? An econometric analysis of international crude oil prices. *Energy Policy*, 72:78–86, 2014.

Proceedings of Machine Learning Research 58 (2016) 41-50

Performance of Kullback-Leibler Based Expert Opinion Pooling for Unlikely Events

Vladimíra Sečkárová SECKAROV@UTIA.CAS.CZ
Department of Adaptive Systems
Institute of Information Theory and Automation
Pod Vodárenskou věží 4, 182 00 Prague 8, CZ

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

The aggregation of available information is of great importance in many branches of economics, social sciences. Often, we can only rely on experts' opinions, i.e. probabilities assigned to possible events. To deal with opinions in probabilistic form, we focus on the Kullback-Leibler (KL) divergence based pools: linear, logarithmic and KL-pool (Sečkárová, 2015). Since occurrence of events is subject to random influences of the real world, it is important to address events assigned lower probabilities (unlikely events). This is done by choosing pooling with a higher entropy than standard linear or logarithmic options, i.e. the KL-pool. We show how well the mentioned pools perform on real data using absolute error, KL-divergence and quadratic reward. In cases favoring events assigned higher probabilities, the KL-pool performs similarly to the linear pool and outperforms the logarithmic pool. When unlikely events occur, the KL-pool outperforms both pools, which makes it a reasonable way of pooling.

Keywords: Opinion Pooling, Combining Probability Distributions, Minimum Kullback-Leibler Divergence

1. Introduction

Problem of information aggregation from multiple sources is of interest in decision making and its applications in areas such as social sciences, economics and business. The choice of the final decision is a delicate process and if an unexpected situation occurs, it can have big psychological or financial impact. In this contribution, we address this problem by considering statistical and information-theoretic techniques and show that, if necessary, impact of the final decision in unexpected situations can be softened.

In applications of decision making (elections, information markets) we often rely on information formulated as an opinion. The process of aggregating this type of information is then referred to as (expert) opinion pooling. By the term expert we not only mean a person, who has knowledge about the variable of interest based on previous experience. We also include sources which, by investing reasonable time and energy, exploit the today's amount of information distributed by media into their advantage. Some of the experts may be certain about the decision (event), some prefer to express the uncertainty among possible events. To include the uncertainty in predicting possible events, we consider the expert opinion in the form of probability distribution. We assume, that experts are able to

form probability distribution as a probability vector and we will not include the details of its elicitation; an extensive discussion about elicitation of probability distributions can be found in (Garthwaite et al., 2005).

Although sources form their opinion to their best knowledge and abilities and assign high probability to the most probably event, the occurrence of the described event is subject to the random influences of the real world such as nature, political situation, ... Then, an event assigned lower probability (an unlikely event) causing non-negligible loss may occur. Thus, the pool admitting unlikely events should assign more uncertainty in events; but still give a reasonable result in the regular case, when event assigned high probability from experts occurred. To measure the amount of uncertainty for probability distribution resulting from combining experts' opinions we exploit information theory, i.e. the Shannon entropy (Shannon and Weaver, 1963). To measure the utility of resulting probability distribution we consider the Kullback-Leibler (KL) divergence as recommended in (Bernardo, 1979).

There are many algorithms for combining probability distributions available and being developed in areas such as risk analysis (Clemen and Winkler, 1999), weather forecasting (Ranjan and Gneiting, 2010), economics (Chen et al., 2005), parameter estimation with knowledge elicitation (Kárný et al., 2014), knowledge sharing with deliberator (Azizi and Quinn, 2016). We focus on the basic combining algorithms used for opinion pooling: linear and logarithmic pool. Both can be obtained via unconstrained optimisation, i.e., minimisation of the KL-divergence (Abbas, 2009). We also consider more sophisticated version of linear pool (KLp) inspired by (Kárný et al., 2009) and introduced in (Sečkárová, 2015), arising from the constrained minimisation of the KL-divergence. The result of any pooling can be viewed as a compromise among considered sources; the compromise is usually derived to satisfy group aims and every included individual has to sacrifice its own aims. Consequently, when majority of experts' opinions are similar, the influence of an expert with different (opposite) opinion is suppressed.

We address these shortcomings by focusing on a combining approach KLp that yields combination with higher entropy than, e.g., linear or logarithmic pool. In particular, we consider constraints on acceptance of the resulting combination of opinions by individual sources, again in the sense of the KL-divergence. KLp, being a conservative compromise, is thus an appropriate pool when unlikely event occurs, especially if reward for formulated opinion is included.

The aim of this contribution is to verify our hypothesis, that the KLp behaves similarly to standard pooling options (linear, logarithmic) in regular case, but outperforms these in case of unlikely events. For comparison purposes, we use the absolute error and the KL-divergence with respect to the perfect prediction (probability 0 for the losing team) together with the quadratic reward.

Next section contains the overview of the construction of the KL-divergence pool KLp and shows behavior of the linear, logarithmic and KLp pools in terms of the entropy of pooled probabilities. In the third section we apply considered opinion pools on the real data obtained from contest for National Football League in USA (NFL games). The fourth section concludes the work.

2. Combining Experts' Discrete Distributions

Consider a finite number of experts labeled by $j = 1, \ldots, s$ providing discrete probability distributions represented by n-dimensional probability vectors:

$$\mathbf{p}_j = (p_{j1}, \ldots, p_{jn}) : \quad p_{ji} \geq 0,\ \sum_{i=1}^{n} p_{ji} = 1,\ n < \infty,\ j = 1, \ldots, s.$$

where n denotes the number of possible events, i.e., outcomes of an underlying random variable.

Let $\mathbf{q}$ denote an unknown probability vector representing the combination of $\mathbf{p}_1, \ldots, \mathbf{p}_s$. Its estimator $\hat{\mathbf{q}}$ is chosen as the minimiser the expected Kullback-Leibler divergence (Bernardo, 1979):

$$\hat{\mathbf{q}} \in \underset{\tilde{\mathbf{q}}(\mathbf{p}_1,\ldots,\mathbf{p}_s)}{\arg\min}\ \mathrm{E}_{\pi(\mathbf{q}|\mathbf{p}_1,\ldots,\mathbf{p}_s)}\mathrm{KLD}(\mathbf{q}||\tilde{\mathbf{q}}). \tag{1}$$

The minimisation task (1) follows the theory of Bayesian decision making (Savage, 1972) and yields:

$$\hat{\mathbf{q}} = (\hat{q}_1, \ldots, \hat{q}_n) = \mathrm{E}_{\pi(\mathbf{q}|\mathbf{p}_1,\ldots,\mathbf{p}_s)}[\mathbf{q}|\mathbf{p}_1, \ldots, \mathbf{p}_s]. \tag{2}$$

The expectation in (2) depends on the conditional probability density function (pdf) $\pi(\mathbf{q}|\mathbf{p}_1, \ldots, \mathbf{p}_s)$, which we specify in two consequent steps:

1. We assume that each expert, if considered as the 'best' representation of value of unknown $\mathbf{q}$, has a finite expected divergence from $\mathbf{q}$

   $$\mathrm{E}_{\pi(\mathbf{q}|\mathbf{p}_1,\ldots,\mathbf{p}_s)}[\mathrm{KLD}(\mathbf{p}_j||\mathbf{q})|\mathbf{p}_1, \ldots, \mathbf{p}_s] < \infty, \quad j = 1, \ldots, s, \tag{3}$$

   and he is willing to include other experts' opinions in the final combination if they represent $\mathbf{q}$ equally or worse in terms of (3). Applying this condition to every expert in the group we obtain that conditional pdfs have to satisfy:

   $$\mathrm{E}_{\pi(\mathbf{q}|\mathbf{p}_1,\ldots,\mathbf{p}_s)}[\mathrm{KLD}(\mathbf{p}_j||\mathbf{q})|\mathbf{p}_1, \ldots, \mathbf{p}_s] = \mathrm{E}_{\pi(\mathbf{q}|\mathbf{p}_1,\ldots,\mathbf{p}_s)}[\mathrm{KLD}(\mathbf{p}_s||\mathbf{q})|\mathbf{p}_1, \ldots, \mathbf{p}_s], \tag{4}$$

   $j = 1, \ldots, s-1$. Previous attempts in terms of the bounded Kerridge inaccuracy (Kerridge, 1961) can be found in (Sečkárová, 2013).

2. We assume that the set of all pdfs satisfying (4) is non-empty. We exploit the minimum cross-entropy principle (Shore and Johnson, 1980) and choose the pdf solving the following optimisation task

   $$\min_{\pi(\mathbf{q}|\mathbf{p}_1,\ldots,\mathbf{p}_s)\ \text{satisfying (4)}} \mathrm{KLD}(\pi(\mathbf{q}|\mathbf{p}_1, \ldots, \mathbf{p}_s)||\pi_0(\mathbf{q})), \tag{5}$$

   with

   $$\pi_0(\mathbf{q}) = \pi_0(\mathbf{q}|\mathbf{p}_1, \ldots, \mathbf{p}_s)$$

   being the prior guess on the conditional pdf $\pi(\mathbf{q}|\mathbf{p}_1, \ldots, \mathbf{p}_s)$.

The constrained optimisation task (5) yields

$$\pi(\mathbf{q}|\mathbf{p}_1,\dots,\mathbf{p}_s) \propto \pi_0(\mathbf{q}) \prod_{i=1}^{n} q_i^{\sum_{j=1}^{s-1} \lambda_j (p_{ji}-p_{si})}, \tag{6}$$

where λ_j are the Lagrange multipliers resulting from the minimisation of (5) with respect to $(s-1)$ equations in (4).

To obtain more specific form of the combination (2) we next specify prior pdf in (5). This pdf is defined over $(n-1)$-dimensional probability simplex - a set of all $\mathbf{p}_j$. Based also on the relation given in (6), we exploit numerically appealing Dirichlet distribution. We then obtain that pdf $\pi(\mathbf{q}|\mathbf{p}_1,\dots,\mathbf{p}_s)$ is also the pdf of the Dirichlet distribution. Its parameters have the following form:

$$\hat{\nu}_i = \nu_{0i} + \sum_{j=1}^{s-1} \lambda_j (p_{ji} - p_{si}), \quad i = 1,\dots,n,$$

where $\nu_{01},\dots,\nu_{0n}$ are parameters of the prior Dirichlet distribution, $\lambda_1,\dots,\lambda_{s-1}$ are the Lagrange multipliers from the constrained optimisation task.

Based on the properties of the Dirichlet distribution

$$\mathrm{E}_{\pi(\mathbf{q}|\mathbf{p}_1,\dots,\mathbf{p}_s)}[q_i|\mathbf{p}_1,\dots,\mathbf{p}_s] = \frac{\hat{\nu}_i}{\sum_{i=1}^{n} \hat{\nu}_i},$$

the estimator (2) is

$$\hat{q}_i = \frac{\nu_{0i}}{\sum_{k=1}^{n} \nu_{0k}} + \sum_{j=1}^{s-1} \frac{\lambda_j}{\sum_{k=1}^{n} \nu_{0k}} (p_{ji} - p_{si}), \quad i = 1,\dots,n. \tag{7}$$

Equation (7) represents the KL-pool (KLp) of expert opinions $\mathbf{p}_1,\dots,\mathbf{p}_s$.

In many real-life inspired cases (elections, betting predictions), there is very little or no prior information available before processing expert opinions. We are thus forced to exploit given opinions in prior guess for parameters of the Dirichlet distribution. Since no new information is included in combining, it is natural that the sum of prior parameters ν_{0i} and parameters after combining $\hat{\nu}_i$ to be equal (see (2)). In particular, we set this sum of parameters to be equal to the number of opinions (or generally observations). Each parameter ν_{0i} is then assigned value relative to the arithmetic mean of the given opinions

$$\nu_{0i} = s \sum_{j=1}^{s} \frac{p_{ji}}{s} = \sum_{j=1}^{s} p_{ji}, \quad i = 1,\dots,n. \tag{8}$$

For comparison purposes, we use

- linear pool (linp):

$$\hat{\mathbf{q}} \propto \sum_{j=1}^{s} w_j p_j, \tag{9}$$

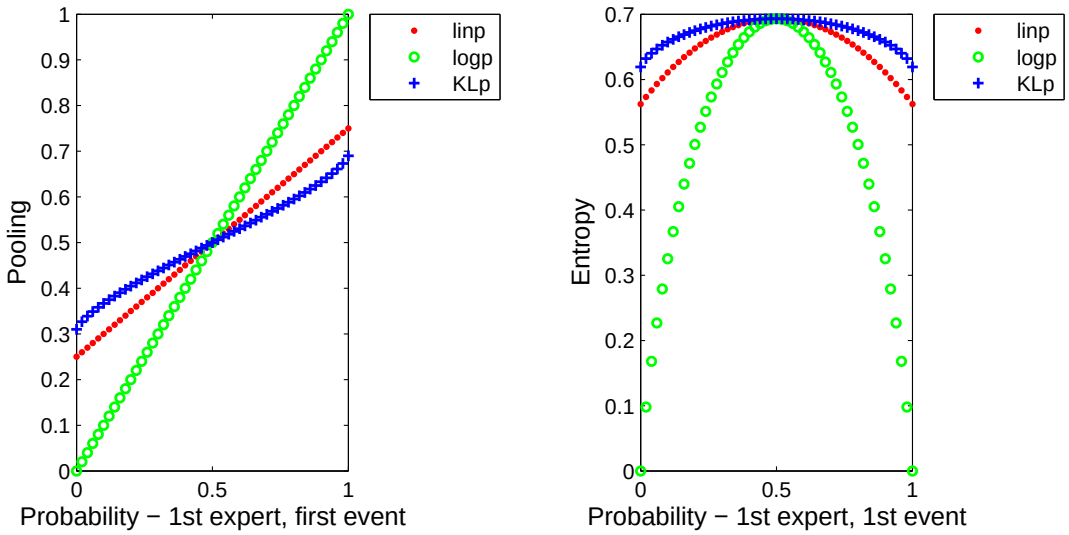

Figure 1: Values of pooling for $\hat{q}_1$ and amount of uncertainty (entropy) for the linear (9), logarithmic (10) and KL-pool (7).

- logarithmic pool (logp):

$$\hat{\mathbf{q}} \propto \prod_{j=1}^{s} p_j^{w_j}, \tag{10}$$

in this contribution, with w_j denoting weights, i.e., subjective preferences among experts $1, \ldots, s$. These two combining approaches arise from the minimisation of both versions of asymmetric KL-divergence without constraints (Abbas, 2009).

2.1 KL-pool as Treatment for Unlikely Events

We expect the experts posses certain level of expertise when assigning the probability to possible events. However, they often have incomplete information about random influences in real world such as nature, political situation, physical and mental abilities of others. Even minor changes in these factors can have huge impact on the occurrence of improbable events. Thus, it is important to reflect also events with lower probabilities assigned by experts – by allowing more uncertainty in the set of probable events. To measure the amount of uncertainty allowed by the combination in (7) we focus on the entropy (Rényi, 1961), exploited in environmental scenarios, e.g., methane emissions from wetlands (Sabolová et al., 2015). Higher values of entropy indicate that more uncertainty is present and that unlikely events should be assigned higher probability.

The formula (7) is not closed; the values of $\hat{\mathbf{q}} = (\hat{q}_1, \ldots, \hat{q}_n)$ are obtained numerically via optimisation with respect to either $\lambda_1, \ldots, \lambda_{s-1}$ or $\nu_1, ..., \nu_n$. The direct theoretical comparison of KLp with linp and logp, by exploiting constraints (4) rewritten for the Dirichlet distribution or by using an approximation of $\hat{\mathbf{q}}$, is a part of the future research.

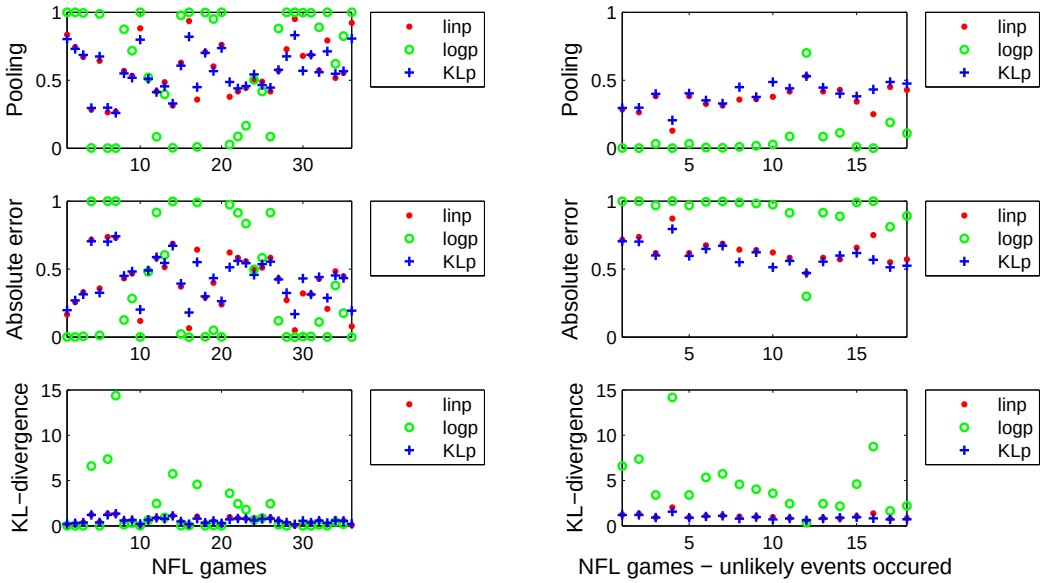

Figure 2: **Left column** Results for pooling based on the linear (linp), logarithmic (logp) and KL-pool (KLp), the absolute error (11) of the pools and the KL-divergence (13) for randomly picked games. **Right column** Results for pooling based on the linp, logp and KLp, the absolute error of the pools and the KL-divergence for the subset of games to which the players assigned low probabilities to the winning team.

We now consider a numerical example to see that the proposed KL-pool has this desirable property: let us thus combine two opinions $\mathbf{p}_j = (p_{j1}, p_{j2})$, $j = 1, 2$, where

$$\mathbf{p}_1 = (0, 1), \ldots, (1, 0) \quad \text{and} \quad \mathbf{p}_2 = (0.5, 0.5),$$

i.e., $\mathbf{p}_1$ varies in the above range with increments in p_{11} equal to 0.02. Figure 1 shows how KLp, with $\nu_{01}, \ldots, \nu_{0n}$ given by (8), behaves in comparison with linear (9) and logarithmic (10) pools arising for unit weights w_j, $j = 1, 2$. In Figure 1 on the left we see the outcomes of considered pools for the first event ($i = 1$). On the right, we see that the KL-pool has equal or higher entropy than linear and logarithmic pool for low-probability events ($p_{11} \leq 0.5$) and thus fits better for the treatment of unlikely events.

3. Real Data Application

In this section we demonstrate that the behaviour of the KL-pool, illustrated on the toy example in the previous section, is favourable when processing real data. We obtained data from `http://probabilityfootball.com/2005/`, an online football contest also referred to as "a game of skill", where the players "estimate, for each football game, the probability that each NFL team will win based on how strong they believe each team to be". Players pick a probability of winning for each team, ranging from 0 to 1 (according to webpage in percentile ranging from 0% to 100%), which they think accurately describes the strength of teams.

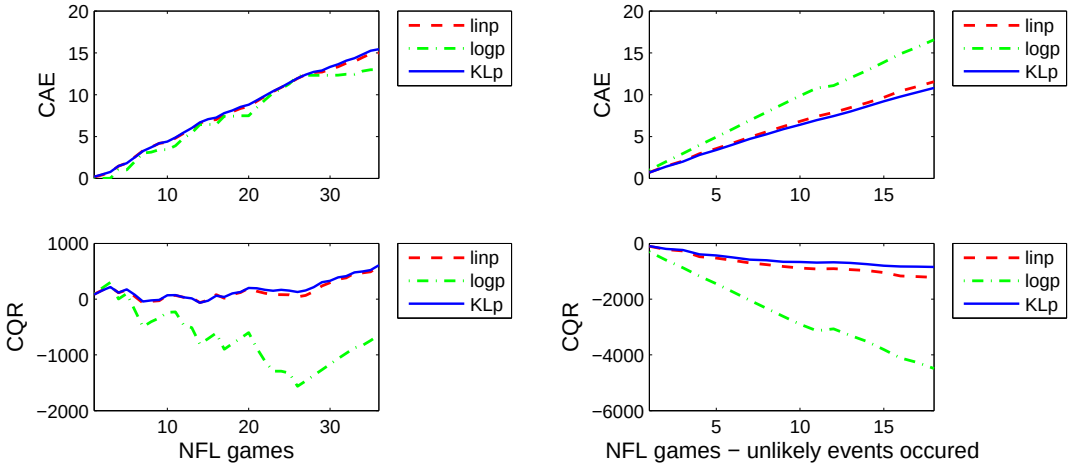

Figure 3: **Left column** Results for the linear (linp), logarithmic (logp) and KL-pool (KLp) in terms of cumulative absolute error (12) and cumulative quadratic reward (14) for randomly chosen games. **Right column** Results for the linp, logp and KLp in terms of CAE and CQR for games, when players assigned low probabilities to the winning team.

We apply previously mentioned linp (9), logp (10) and KLp (7) to demonstrate their behaviour and compare:

- values of pooling for the first event (the team that eventually won),
- absolute error measuring the difference between perfect prediction (probability 0 for losing team) and the actual prediction (Chen et al., 2005)

$$\text{AE} = |0 - \hat{q}_{\text{losing team}}|, \quad \hat{q}_{\text{losing team}} \in \{\text{linp, logp, KLp}\} \tag{11}$$

  and its cumulative version

$$\text{CAE}_t = \text{CAE}_{t-1} + |0 - \hat{q}_{\text{losing team},t}|, \quad \hat{q}_{\text{losing team},t} \in \{\text{linp, logp, KLp}\} \tag{12}$$

  where CAE_{t-1} is the sum of absolute errors from $(t-1)$ previous games and $\hat{q}_{\text{losing team},t}$ is the pooling of opinions for current game,

- and KL-divergence

$$\text{KLDiv} = \text{KLD}(\mathbf{q}_{\text{perfect}}||\hat{\mathbf{q}}), \quad \hat{\mathbf{q}} \in \{\text{linp, logp, KLp}\} \tag{13}$$

  where $\mathbf{q}_{\text{perfect}}$ is a perfect prediction (probability 0 for losing team),

- cumulative quadratic reward

$$\text{CQR} = \text{CQR}_{t-1} + 100 - 400(\hat{q}_{\text{losing team},t})^2, \quad \hat{q}_{\text{losing team},t} \in \{\text{linp, logp, KLp}\} \tag{14}$$

  with quadratic reward used in the contest (Chen et al., 2005).

Firstly, focus on the first 6 weeks of NFL in 2005 for 7 successful players, each week 6 games were chosen. Results in the left column of Figure 2 show, that KL-pool performs similarly to the linear pool in terms of absolute error (11) and KL-divergence (13). This result is expected because of the additive nature of the KL-pool. The performance of the logarithmic pool is due to its multiplicative form oscillatory - it performs either very good or very poor.

Secondly, we focus on 18 games, when the majority of picked players put more believe into a team, which eventually lost the game. Figure 2 (in the right column) shows that in such case our method performs better than the linear pool, as intuitively expected.

Because of the design of this example, i.e., the variability in playing football teams, it is difficult to study the properties of considered pools sequentially (based on estimates from the previous game). Thus, we next exploit performance of these pools with respect to the cumulative versions of the absolute error (12) and the quadratic reward (14). The results depicted in Figure 3 are similar to the non-cumulative case: for randomly chosen games, the KL-pool performs similarly to the linear pool and both pools outperform the logarithmic pool (their CAE is lower and CQR is higher). In case of games, where the losing team was assigned lower probability by the players, the KL-pool outperforms both, the linear and the logarithmic pool.

4. Conclusion and Future Work

In this contribution we focused on problem of pooling expert opinions when events assigned lower probabilities occur (unlikely events). This is especially important in cases when also a reward for formulated opinion is also included: the higher probability assigned to the event that did not occur can yield huge loss. To treat this we need an opinion pool resulting in a combination with a higher entropy than standard pooling ways, so that unlikely events obtain reasonable probability. We considered the Kullback-Leibler based opinion pool (Sečkárová, 2015), values of which were obtained via constrained non-linear optimisation. This pool was constructed as the compromise for group of experts a) without sacrificing their own aims and b) without suppressing opinion which significantly differs from other opinions.

We showed on numerical example that KL-pool reaches higher values of entropy than generally known pools: (equally weighted) linear pool and logarithmic pool. We then applied these pools on real data and compared them using following performance measures: the absolute error, the KL-divergence and the quadratic reward. Because of the additive form of the KL-pool, its performance was similar to the performance of the linear pool for regular data (not many unlikely events occurred). In case of unlikely events, the KL-pool outperformed the linear and the logarithmic pool in terms of cumulative version of the absolute error and the quadratic reward and thus is a reasonable tool for pooling expert opinions. The future work includes the theoretical comparison of KL-pool with other opinion pools by using its approximation. Also, the application of the KL-pool on other sets of real data including betting with knowledge of fixed-odds, handling financial contracts, weather forecasts, is of interest.

Acknowledgements

This work has been supported by the GAČR grant GA16-09848S.

References

A. E. Abbas. A Kullback-Leibler View of Linear and Log-Linear Pools. *Decision Analysis*, 6(1):25–37, 2009.

S. Azizi and A. Quinn. Hierarchical fully probabilistic design for deliberator-based merging in multiple participant systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, PP(99):1–9, 2016. ISSN 2168-2216. doi: 10.1109/TSMC.2016.2608662.

J. M. Bernardo. Expected information as expected utility. *Ann. Stat.*, 7:686–690, 1979.

Y. Chen, C. Chu, T. Mullen, and D. M. Pennock. Information Markets vs. Opinion Pools: An Empirical Comparison. In *Proceedings of the 6th ACM Conference on Electronic Commerce*, pages 58–67, 2005.

R. T. Clemen and R. L. Winkler. Combining probability distributions from experts in risk analysis. *Risk Analysis*, 19(2):1539–6924, 1999.

P. H. Garthwaite, J. B. Kadane, and A. O'Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–701, 2005.

M. Kárný, T. V. Guy, A. Bodini, and F. Ruggeri. Cooperation via sharing of probabilistic information. *Int. J. of Computational Intelligence Studies*, 1(5):139–162, 2009.

M. Kárný, A. Bodini, T.V. Guy, J. Kracík, P. Nedoma, and F. Ruggeri. Fully probabilistic knowledge expression and incorporation. *Statistics and Its Interface*, 7(4):503–515, 2014.

D. F. Kerridge. Inaccuracy and inference. *J. R. Stat. Soc., Ser. B*, 23:184–194, 1961.

R. Ranjan and T. Gneiting. Combining probability forecasts. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(1):71–91, 2010.

A. Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 547–561, Berkeley, Calif., 1961. University of California Press.

R. Sabolová, V. Sečkárová, J. Dušek, and M. Stehlík. Entropy based statistical inference for methane emissions released from wetland. *Chemometrics and Intelligent Laboratory Systems*, 141(0):125 – 133, 2015. ISSN 0169-7439.

L. J. Savage. *The Foundations of Statistics*. Dover Books on Mathematics Series. Dover Publications, 1972.

V. Sečkárová. A Note on Weighted Combination Methods for Probability Estimation. In *Proccedings of the 3rd International Workshop on Scalable Decision Making held in conjunction with ECML/PKDD 2013 , Eds: Guy Tatiana V., Kárný Miroslav*. UTIA AV CR, Prague, 2013. ISBN 978-80-903834-8-7.

V. Sečkárová. Weighted probabilistic opinion pooling based on cross-entropy. In *Neural Information Processing, 22nd International Conference, ICONIP 2015, Istanbul, Turkey, November 9-12, 2015, Proceedings, Part II*, pages 623–629, Istanbul, Turkey, 2015.

C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. Number pt. 11 in Illini books. University of Illinois Press, 1963.

J. E. Shore and R. W. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory*, 26:26–37, 1980.

Proceedings of Machine Learning Research 58 (2016) 51-60

Experimental Performance of Deliberation-Aware Responder in Multi-Proposer Ultimatum Game

Tatiana V. Guy GUY@UTIA.CAS.CZ
Marko Ruman MARKO.RUMAN@GMAIL.COM
František Hůla HULA.FRANTISEK@GMAIL.COM
Miroslav Kárný SCHOOL@UTIA.CAS.CZ
Department of Adaptive Systems
Institute of Information Theory and Automation
Czech Academy of Sciences
Prague, P.O.Box 18, 182 08, CZ

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

The ultimatum game serves for studying various aspects of decision making (DM). Recently, its multi-proposer version has been modified to study the influence of deliberation costs. An optimising policy of the responder, switching between several proposers at non-negligible deliberation costs, was designed and successfully tested in a simulated environment. The policy design was done within the framework of Markov Decision Processes with rewards also allowing to model the responder's feeling for fairness. It relies on simple Markov models of proposers, which are recursively learnt in a Bayesian way during the game course. This paper verifies, whether the gained theoretically plausible policy, suits to real-life DM. It describes experiments in which this policy was applied against human proposers. The results – with eleven groups of three independently acting proposers – confirm the soundness of this policy. It increases the responder's economic profit due to switching between proposers, in spite of the deliberation costs and the used approximate modelling of proposers. Methodologically, it opens the possibility to learn systematically willingness of humans to spent their deliberation resources on specific DM tasks.

Keywords: decision making; deliberation effort; Markov decision process; ultimatum game

1. Introduction

Maximizing of expected utility is perceived as "rational" within traditional economic models (von Neumann and Morgenstern (1944); Thaler (2000)). The observed discrepancies between theoretically optimal DM and real DM, e.g. (Gong et al. (2013), Jones (1999), Regenwetter et al. (2011)), can be diminished by changing the behaviour of DM subjects (by educating them) and by modifying decision rewards and models used in prescriptive theories. Our paper deals with the latter case and focuses on the influence of *deliberation effort* in DM.

A proper theory respecting deliberation effort should take into account that any decision made, either by humans or machines, costs time (Ortega and Stocker (2016)) energy and possibly other, limited, resources (Ortega et al. (2016)); a sample of different application

domains and related references are in Ruman et al. (2016). In this paper, we rely on, the design of DM policies respecting deliberation effort treated as an application of standard Markov Decision Process (MDP, Puterman (1994)) with the reward explicitly influenced by deliberation costs and with the environment model learnt in a Bayesian way (Peterka (1981)). The solution was developed for designing policy of the responder in multi-proposer ultimatum games (UG, Rubinstein (1994)). The simplicity of the UG enables extensive tests confronting prescriptive and human DM. In the UG, the proposer offers the split of a given amount of money and the responder either accepts this split, and the money are split, or refuses it, and none of players gets anything. In multi-proposer versions, the responder has the right to select the proposer among several of them. Any change of the proposer between rounds is penalized. In this way, the influence of deliberation effort is respected. Simulations documented in Ruman et al. (2016) confirmed the expected behaviour of the proposed responder's policy. However, a key question remained open: Will this policy be successful in real-life? It is a specific case of the generally inspected question: Does a prescriptive, theoretically justified, solution suit real life DM? Our experiments with human proposers, which form the core of this paper, provide answers to the posed questions.

The paper layout is as follows. Section 2 recalls formalization and the optimal design of the responder's policy in the multi-proposer UG. Section 3 describes the performed experiments and their results. Section 4 contains discussion. Section 5 provides concluding remarks.

2. Tested Decision-Making Policy

The section formalizes the multi-proposer UG and recalls the essence of the tested policy proposed in Ruman et al. (2016) as an application of MDP (Puterman (1994)).

2.1 Preliminaries

Throughout, bold capital $\mathbf{X}$ is a set of x-values; x_t is the value of x at the *decision epoch* $t \in \mathbf{T} = \{1, \ldots, N\}$ bounded by a horizon $N \in \mathbb{N}$ (here, the number of game rounds); $p(x|y)$ is a conditional probability. MDP provides a general framework for describing an *agent* (here, responder), which interacts with an *environment* (here, available proposers) by taking appropriate actions to achieve her goal. The decisions about *actions* $a_t \in \mathbf{A}$ (here, select the proposer to play with and accept or reject her offer) made are only influenced by the observed environment *state* $s_{t-1} \in \mathbf{S} \subset \mathbb{N}$, not by the whole environment history. The state s_{t-1} evolves to s_t according to the *transition probabilities*, $p = (p(s_t|a_t, s_{t-1}))_{t\in\mathbf{T}}$, influenced by actions $(a_t)_{t\in\mathbf{T}}$, and the agent receives rewards $r = (r(s_t, a_t, s_{t-1}))_{t\in\mathbf{T}}$. Given the initial state $s_0 = 0$, the tuple $(\mathbf{T}, \mathbf{S}, \mathbf{A}, r, p)$ describes MDP. The agent evaluates *randomized DM policies* $\pi = \left((p(a_t|s_{t-1}))_{a_t\in\mathbf{A}, s_{t-1}\in\mathbf{S}} \right)_{t\in\mathbf{T}}$ — formed by randomized *decision rules* $p(a_t|s_{t-1})$, $a_t \in \mathbf{A}$, $s_{t-1} \in \mathbf{S}$, $t \in \mathbf{T}$ — based on the *expected reward*

$$E_\pi[r(s_t, a_t, s_{t-1})] = \sum_{a_t\in\mathbf{A}} \sum_{s_t\in\mathbf{S}} \sum_{s_{t-1}\in\mathbf{S}} r(s_t, a_t, s_{t-1}) p(s_t|a_t, s_{t-1}) p(a_t|s_{t-1}) p(s_{t-1}), \quad (1)$$

where state probabilities $p(s_t)$ evolve according to

$$p(s_t) = \sum_{a_t\in\mathbf{A}} \sum_{s_{t-1}\in\mathbf{S}} p(s_t|a_t, s_{t-1}) p(a_t|s_{t-1}) p(s_{t-1})$$

with $p(s_0) = \delta(s_0, 0)$. The used Kronecker symbol $\delta(x, y) = 1$ if $x = y$, $\delta(x, y) = 0$ if $x \neq y$.

The agent seeks for the *optimal policy* π^{opt} maximizing the sum of expected rewards (1) up to the horizon N

$$\pi^{opt} \in \operatorname*{argmax}_{\pi} \sum_{t \in \mathbf{T}} E_{\pi}[r(s_t, a_t, s_{t-1})]. \tag{2}$$

2.2 Deliberation-Aware Multi-Proposer Ultimatum Game

The considered *multi-proposer N-round UG* assumes a *group* of $K \in \mathbb{N}$ proposers $\mathcal{P} \in \boldsymbol{\mathcal{P}} = \{\mathcal{P}^1, \ldots, \mathcal{P}^K\}$ and one responder $\mathcal{R}$. The responder's goal is the same as in the traditional UG, i.e. to influence her accumulated profit R_t, see (4), while accepting or rejecting the offers. The main difference is that at the beginning of each round $t \in \mathbf{T}$ the responder chooses a proposer $\mathcal{P}_t \in \boldsymbol{\mathcal{P}} = \{\mathcal{P}^1, \ldots, \mathcal{P}^K\}$ to play with. The choice of a proposer $\mathcal{P}_t$ at round $t \in \mathbf{T}$ is the first responder's action $a_{1t} = \mathcal{P}_t \in \boldsymbol{\mathcal{P}} = \mathbf{A}_1$. To model deliberation costs, the choice of a proposer different from that in the previous round is penalized by a *deliberation penalty* $d \in \mathbb{N}$. This leads to the *accumulated deliberation cost* of the responder

$$\mathsf{D}_t = d \sum_{\tau=1}^{t} (1 - \delta(\mathcal{P}_\tau, \mathcal{P}_{\tau-1})). \tag{3}$$

Then, as in the original UG (Rubinstein (1994)) the chosen proposer $\mathcal{P}_t$ offers a split $o_t \in \mathbf{O} = \{1, 2, \ldots, q-1\}$, $q \in \mathbb{N}$, for the responder and $(q - o_t)$ for herself. Money split according to the proposal if the responder *accepts the offer*, if she chooses the action $a_{2t} = 2$. None of the players get anything if the responder *rejects the offer*, if she chooses the action $a_{2t} = 1$. The accumulated responder's (economic) *profit*, R_t, at round $t \in \mathbf{T}$, is

$$\mathsf{R}_t = \sum_{\mathcal{P} \in \boldsymbol{\mathcal{P}}} \mathsf{P}_{\mathcal{P}t}, \qquad \mathsf{P}_{\mathcal{P}t} = \sum_{\tau=1}^{t} o_\tau (a_{2\tau} - 1) \delta(a_{1\tau}, \mathcal{P}), \qquad \mathcal{P} \in \boldsymbol{\mathcal{P}}. \tag{4}$$

Proposers play a passive role whenever they are not selected in the round. The accumulated proposers' (economic) profits are

$$Z_{\mathcal{P}t} = \sum_{\tau=1}^{t} (q - o_\tau)(a_{2\tau} - 1) \delta(a_{1\tau}, \mathcal{P}), \qquad \forall \mathcal{P} \in \boldsymbol{\mathcal{P}}. \tag{5}$$

In the multi-proposer UG, the responder R can be modelled as an agent in MDP, which tries to maximize her accumulated profit while minimising her accumulated deliberation cost.

Definition 1 *The multi-proposer UG in the MDP framework, with epochs $t \in \mathbf{T}$ identified with* game rounds*, is described through*

- *the environment state at $t \in \mathbf{T}$*

$$s_t = (o_t, \sigma_t) \;\; \textit{with } \sigma_t = (\mathcal{P}_t, \mathsf{D}_t, \mathsf{R}_t, Z_{\mathcal{P}^1 t}, Z_{\mathcal{P}^2 t}, ..., , Z_{\mathcal{P}^K t}), \quad \textit{where} \tag{6}$$

  $o_t \in \mathbf{O}$ is an offer made by the proposer $\mathcal{P}_t$. The accumulated deliberation cost D_t, the accumulated responder's (economic) profit R_t and the accumulated (economic) profits of proposers $Z_{\mathcal{P}t}$, $\mathcal{P} \in \boldsymbol{\mathcal{P}}$, are defined by (3), (4) and (5), respectively.

- *the two-dimensional action* $a = (a_1, a_2) \in \mathbf{A}_1 \times \mathbf{A}_2$ *consists of the selection* $a_1 \in \mathbf{A}_1 = \boldsymbol{\mathcal{P}}$ *of the proposer to play with and of* $a_2 \in \mathbf{A}_2 = \{1, 2\} = \{reject, accept\}$ *the offered split made by the selected proposer.*

  The selection of the proposer $a_{1t} = \mathcal{P}_t \in \boldsymbol{\mathcal{P}}$ *is based on the state* s_{t-1} *(6), while the action* $a_{2t} \in \mathbf{A}_2$ *also depends on the offer* $o_t \in \mathbf{O}$ *of the selected proposer* $\mathcal{P}_t \in \boldsymbol{\mathcal{P}}$*. Thus,*

$$p(a_t|o_t, s_{t-1}) = p(a_{1t}, a_{2t}|o_t, s_{t-1}) = p(a_{1t}|s_{t-1})p(a_{2t}|o_t, a_{1t}, s_{t-1}). \tag{7}$$

  Consequently, the optimal policy π^{opt} *is searched among sequences of functions*

$$\pi = \big(p(a_{1t}|s_{t-1}), p(a_{2t}|o_t, a_{1t}, s_{t-1})\big)_{t\in\mathbf{T}}. \tag{8}$$

- *The reward function with the penalty for the deliberation costs and respecting also self-fairness (Guy et al. (2015)) is considered*

$$r(s_t, a_t, s_{t-1}) = w(\mathsf{R}_t - \mathsf{R}_{t-1}) - (1-w)(Z_{\mathcal{P}_t t} - Z_{\mathcal{P}_t(t-1)}) - (\mathsf{D}_t - \mathsf{D}_{t-1}),\ w \in [0, 1]. \tag{9}$$

- *The transition probabilities* $p = p(s_t|a_{1t}, s_{t-1})$ *are assumed to be known, possibly as point estimates resulting from recursive estimation (Hůla et al. (2016)).*

For the inspection of the influence of deliberation costs, the risk neutral *economic responder*, caring about pure economic profit balanced with deliberation costs, is of interest. It is a special case of (9) with the weight $w = 1$. The results in Hůla et al. (2016), were adaptive proposer was studied, indicate that the economic player generates insufficiently exciting actions causing non-convergence of parameter estimates. In the cited case, the economic proposer generated very narrow range of offers and could not learn reactions of the responder to values out of this range. To avoid it, the self-fair modification of the responder's policy with weight $w \neq 1$ in (9), was used, but the results were judged according to the responder's profit.

2.3 Optimal Deliberation-Aware Responder Policy

Dynamic programming (Bellman (1957); Bertsekas (2001)) is used to solve (2). The special structure of policies (8) calls for a specific construction of the optimal policy. The following theorem — a tailored dynamic programming presented in Ruman et al. (2016) — provides the optimal strategy of the responder.

Theorem 2 (Optimal policy of the deliberation-aware responder) *The optimal policy* π^{opt} *constrained by (7) is formed by the sequence of decision rules*

$$\left\{(p^{opt}(a_{1t}|s_{t-1}), p^{opt}(a_{2t}|o_t, a_{1t}, s_{t-1}))\right\}_{t=1}^{N},$$

with $s_t = (o_t, \sigma_t)$ (6), which are evaluated against game course, starting with the value function $\varphi_N(s_N) = 0$, $\forall s_N \in \mathbf{S}$,

$$\begin{aligned}
\varphi_{t-1}(s_{t-1}) &= E[r(o_t, \sigma_t, a^\star_{1t}, a^\star_{2t}, s_{t-1}) + \varphi_t(s_t) | a^\star_{1t}, a^\star_{2t}, s_{t-1}] \\
a^\star_{1t}(s_{t-1}) &\in \operatorname*{argmax}_{a_{1t} \in \mathbf{A}_1} E\left[r(o_t, \sigma_t, a_{1t}, a_{2t}, s_{t-1}) + \varphi_t(s_t) \mid a_{1t}, s_{t-1}\right] \\
p^{opt}(a_{1t}|s_{t-1}) &= \delta(a_{1t}, a^\star_{1t}(s_{t-1})) \\
a^\star_{2t}(o_t, a^\star_{1t}, s_{t-1}) &\in \operatorname*{argmax}_{a_{2t} \in \mathbf{A}_2} E\left[r(o_t, \sigma_t, a^\star_{1t}, a_{2t}, s_{t-1}) + \varphi_t(s_t) \mid a^\star_{1t}, a_{2t}, o_t, s_{t-1}\right] \\
p^{opt}(a_{2t}|o_t, s_{t-1}) &= \delta(a_{2t}, a^\star_{2t}(o_t, a^*_{1t}, s_{t-1})). \qquad (10)
\end{aligned}$$

For the reward (9), the action $a^\star_{1t}$ in (10) describes the optimal, deliberation-aware, choice of the proposer and $a^\star_{2t}$ the optimal response to her offer.

3. Experiments

Successful simulation experiments of the optimal responder's policy π^{opt} (described by Theorem 2) were presented in Ruman et al. (2016). Its usefulness in real life has not been tested. This section presents results of experiments performed to fill this gap. The assumption that real proposers use time-invariant proposal policies describable by known or well-estimated probabilities $p = p(s_t|a_{1t}, s_{t-1})$ is the key potential weakness of the theoretically optimal responder's policy. Thus, the experiments mainly checked this assumption. In this experiment human-participants played roles of proposers against a virtual responder that was a computer programme implementing the deliberation-aware policy as described in Section 2.3.

3.1 Experimental Setup

Thirty three university students (mostly males, age range 19-25 years) participated in the study. The participants had no or minimal knowledge of the UG. The human-proposers could not interact/communicate during the game as the game ran through a web interface on personal computers. The keyboard was used as a response device. The participants were instructed about the UG rules and their role in the experiment. At each round a participants action was to propose an integer split of $q = 10\,CZK$. Participants were told to play trying to maximize their profit. The real money was at stake. The participants played with virtual money but they were paid their profits won in the game block (see below) at the end of experiment. The virtual responder always played with a group of three human-proposers $\mathcal{P} \in \boldsymbol{\mathcal{P}} = \{\mathcal{P}^1, \mathcal{P}^2, \mathcal{P}^3\}$. While playing, the virtual responder recursively learned the parameters $\alpha_\mathcal{P}, \beta_\mathcal{P} > 0,\ \alpha_\mathcal{P} + \beta_\mathcal{P} < 1$ of the simplified proposer's model

$$p(s_t|a_t, s_{t-1}, \alpha_\mathcal{P}, \beta_\mathcal{P}) = p(o_t|o_{t-1}, \alpha_\mathcal{P}, \beta_\mathcal{P}) = \begin{cases} \alpha_\mathcal{P} & \text{for offers } o_t > o_{t-1} \\ \beta_\mathcal{P} & \text{for offers } o_t < o_{t-1} \\ 1 - \alpha_\mathcal{P} - \beta_\mathcal{P} & \text{for offers } o_t = o_{t-1} \end{cases}. \qquad (11)$$

Learning essentially consists of evaluating the relative frequencies $\hat{\alpha}_\mathcal{P}, \hat{\beta}_\mathcal{P}$ corresponding to (11).

Interaction with a human group was split into *learning block* and *game block*. In the learning block, the virtual responder played $N = 20$ game rounds with each proposer $\mathcal{P} \in \boldsymbol{\mathcal{P}}$ (fixed $a_{1t} = \mathcal{P}$) while maximizing the expected accumulated reward with the reward function (9) and with the transition probabilities $p(s_t|a_t, s_{t-1}) = p(o_t|o_{t-1}, \hat{\alpha}_{\mathcal{P}_t}, \hat{\beta}_{\mathcal{P}_t})$. This made the virtual player adaptive. The learning block also helped the participants familiarize with their task.

In the game block, the virtual responder played $N = 20$ game rounds with all three human-proposers at once $\mathcal{P} \in \boldsymbol{\mathcal{P}}$ ($a_{1t} = \mathcal{P}_t$ was also optimized), while maximizing the expected accumulated reward with the reward function (9) and with the transition probabilities $p(s_t|a_t, s_{t-1}) = p(o_t|o_{t-1}, \hat{\alpha}_{\mathcal{P}_t}, \hat{\beta}_{\mathcal{P}_t})$, where the point estimates $(\hat{\alpha}_P, \hat{\beta}_P)$ of (α, β) were permanently updated. Thus, during each round the virtual responder had selected one of the human-proposers, who offered a split. Then the virtual responder decided on the acceptance/rejection of the split. The deliberation cost was set to $d = 1\,CZK$.

The results presented in Hůla et al. (2016) indicated that the economic player mostly generates insufficiently exciting actions, which result into divergence of parameter estimates. In the case of economic proposer, studied in Hůla et al. (2016), the player generated very narrow range of offers and could not learn reactions of the responder to values out of this range. To suppress this effect in our case, the self-fair modification of the virtual responder's policy with the weight $w = 0.6$ in (9) (Guy et al. (2015)) was used. The results were, however, judged according to the responder's profit.

3.2 Results

The achieved results are summarized in Table 1. It contains the *profits* $\mathsf{Pk} = \mathsf{P}_{\mathcal{P}^k N}$, (4), when playing with the proposer $\mathcal{P}^k$, $k = 1, 2, 3$, in the learning block. The average

$$\mathsf{RM} = \frac{1}{3}\sum_{k=1}^{3} \mathsf{Pk} \tag{12}$$

of these profits is comparable with the virtual *responder's profit* $\mathsf{R} = \mathsf{R}_N$ (4) in the game block. Switching between proposers is profitable if and only if the difference $\mathsf{R} - \mathsf{RM} > 0$. The results in the table are ordered according to this difference. Sample descriptive statistics are provided.

The ordered differences are also presented in Figure 1. Figure 2 presents two more detailed samples of the game results related to games with the worst and the best responder's results (1 and 11, Table 1).

4. Discussion

The experimental results confirm the potential applicability of the proposed responder's policy based on MDP in real life. Table 1 shows that the average and median of responder's profits increased by switching in spite of the relatively high switching cost, see boldface numbers. The increase realized in experiments with seven human groups from the performed eleven experiments, see Figure 1. Three experiments led to quite bad results and the difference of the average and median indicates that the distribution of differences $\mathsf{R} - \mathsf{RM}$ has heavy tail at negative profit differences. Thus, there is a space for improvement of the

Game	P1	P2	P3	RM	R	R − RM
1	70.00	25.00	117.00	70.67	51.00	-19.67
2	95.00	119.00	137.00	117.00	102.00	-15.00
3	163.00	130.00	121.00	138.00	131.00	-7.00
4	81.00	87.00	75.00	81.00	80.00	-1.00
5	96.00	80.00	106.00	94.00	96.00	2.00
6	140.00	94.00	60.00	98.00	101.00	3.00
7	79.00	69.00	58.00	68.67	77.00	8.33
8	76.00	88.00	91.00	85.00	94.00	9.00
9	63.00	76.00	83.00	74.00	86.00	12.00
10	75.00	93.00	76.00	81.33	95.00	13.67
11	64.00	103.00	120.00	95.67	110.00	14.33
mean	91.09	87.64	94.91	91.21	93.00	**1.79**
median	79.00	88.00	91.00	85.00	95.00	**3.00**
minimum	63.00	25.00	58.00	68.67	51.00	-19.67
maximum	163.00	130.00	137.00	138.00	131.00	14.33
standard deviation	32.10	27.47	26.80	20.96	20.32	11.52

Table 1: Pk is the virtual-responder's profit when playing with proposer $\mathcal{P}^k$ with no switching, RM is the average (12), $\mathsf{R} = \mathsf{R}_N$ (4) is the virtual responder's profit in the game block (with penalized switching).

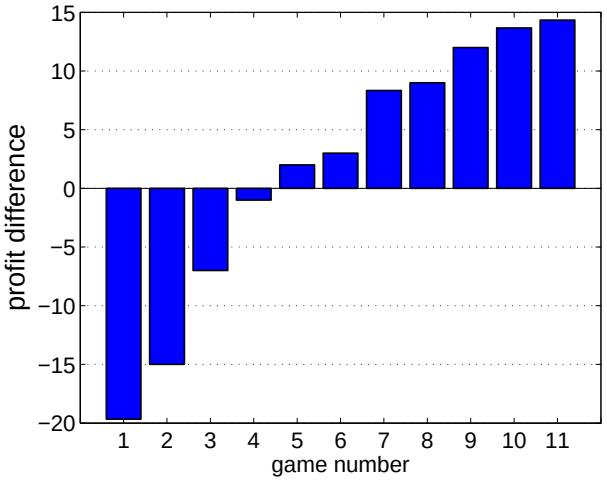

Figure 1: Profit differences in the game block and the average profit (12).

applied policy. Primarily, it has to improve the environment modelling, i.e. modelling of the proposers. Specifically:

- The assumed structure (11) should be refined by making probabilities of offers dependent on the values of differences $o_t - o_{t-1}$. A geometric change can still provide

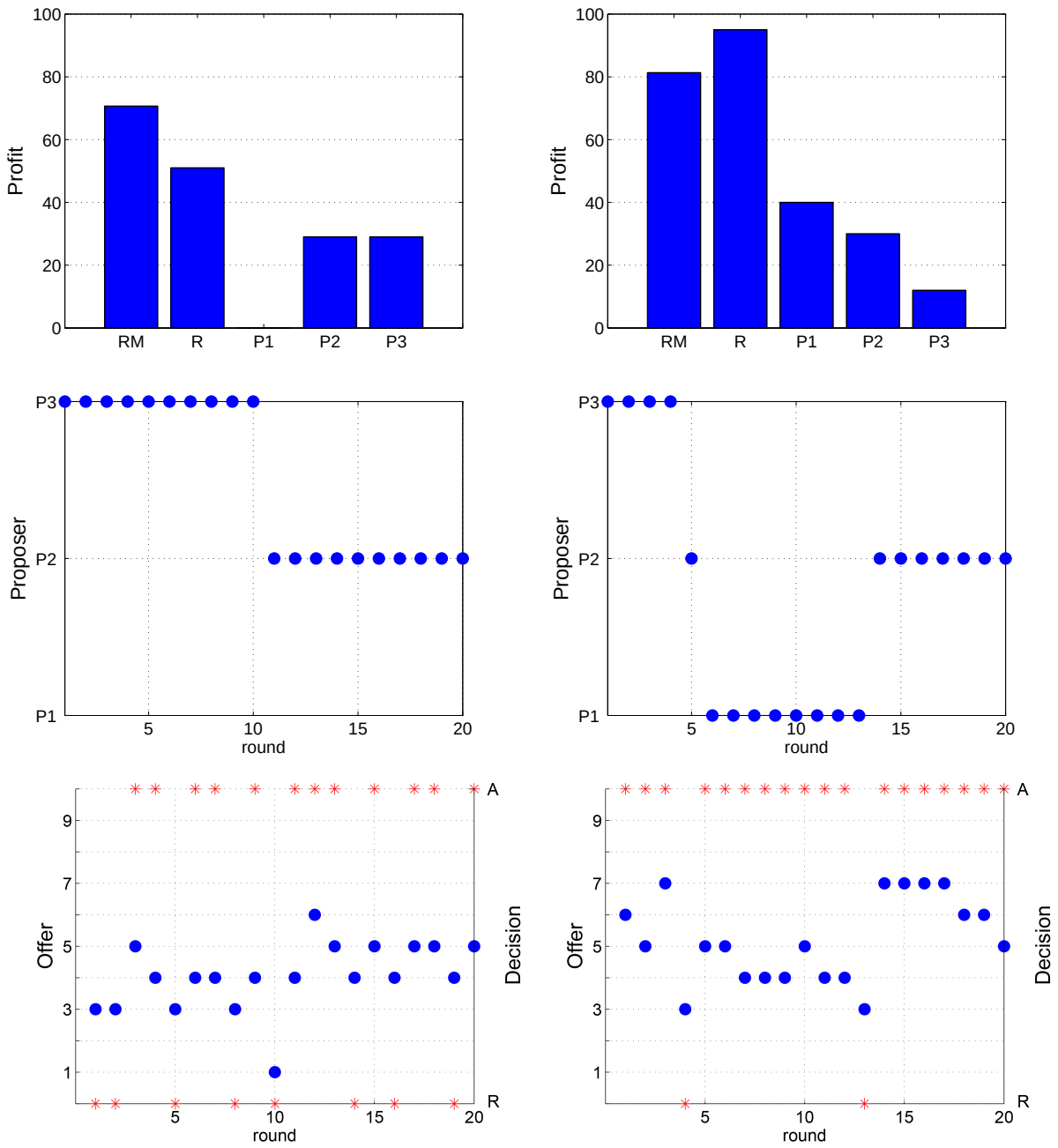

Figure 2: Left column concerns game 1, right column game 11 . First row shows the average profit RM (12), the virtual responder's profit $\mathsf{R} = \mathsf{R}_N$ (4) in the game block, and profits $\mathsf{Pk} = \mathsf{P}_{\mathcal{P}^k N}$ while playing with proposers $\mathcal{P}^k$ in the learning block. Second row shows switching between proposers $\mathcal{P}^k$ in the game block. Third row presents offers (•) and decisions Accept/Reject (∗) within the game course played in the game block.

parsimonious parametrization when considering constant (possibly asymmetric) decrease rates of probabilities $p(o_t|o_{t-1})$.

- Carefully selected prior probabilities of estimated parameters, based, for instance, on the already run games, may speed up estimation and increase the responder's profit. This is the main advantage of the adopted Bayesian estimation (Berger (1985); Garthwaite et al. (2005)), which is confirmed in the UG context in Hůla et al. (2016).
- The adopted exploitive strategy, definitely influenced the results as seen in Figure 2. The exploration-supporting "trick" based on optimization of non-economic profit reward (assuming *self-fair* virtual responder (Guy et al. (2015)) should be replaced by a more systematic treatment based on approximate dynamic programming (Si et al. (2004)) needed when dealing with learnt environment models (Feldbaum (1960)).

5. Concluding Remarks

This paper experimentally examined the influence of deliberation costs of the virtual responder in a multi-proposer Ultimatum Game with human proposers. It confirmed that the use of the MDP machinery while taking proposers as a part of environment is an efficient tool for solving game-like situations. This solution is close to the Bayesian games (Harsanyi (2004)). Our paper supports the claim that the approach caring about dynamics and incomplete knowledge of players makes the adopted theory applicable in real life. It offers extreme flexibility in modelling player's aims. In our case, it respects the influence of deliberation costs on decision making. This makes it not only a useful design tool but also an analytical tool. The analysis concerns a real decision maker, whose acting differs from "rational", purely economic, behaviour. We can analyse it as an inversion problem: assuming that she *is* rational but uses a different than economic reward, we can learn it from her actions. This idea was successfully used in learning of self-fairness in Guy et al. (2015). This paper opens the way of learning *laziness*, the *personal penalty of deliberation effort*. This is one direction of future work, which has to deal with others aspects like: i) extension of the presented experimental study to more groups of proposers with improved learning; ii) fighting with the curse of dimensionality inherent to the Bayesian games; iii) joint modelling of non-profit influences (deliberation costs, fairness, emotions, etc.).

Acknowledgments

The research reflected in this paper has been supported by GAČR GA16-09848S. Thorough feedbacks from anonymous reviewer help to improve the paper significantly.

References

R. Bellman. *Dynamic Programming.* Princeton University Press, New York, 1957.

J.O. Berger. *Statistical Decision Theory and Bayesian Analysis.* Springer, New York, 1985.

D.P. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, US, 2001.

A.A. Feldbaum. Theory of dual control. *Autom. Remote Control*, 21(9), 1960.

A.A. Feldbaum. Theory of dual control. *Autom. Remote Control*, 22(2), 1961.

P.H. Garthwaite, J.B. Kadane, and A. O´Hagan. Statistical methods for eliciting probability distributions. *J. of the American Statistical Association*, 100(470):680–700, 2005.

J. Gong, Y. Zhang, Z. Yang, Y. Huang, J. Feng, and W. Zhang. The framing effect in medical decision-making: a review of the literature. *Psychology, Health and Medicine*, 18 (6):645–653, 2013.

T.V. Guy, M. Kárný, A. Lintas, and A.E.P. Villa. Theoretical models of decision-making in the ultimatum game: Fairness vs. reason. In R. Wang and Xiaochuan X. Pan, editors, *Advances in Cognitive Neurodynamics (V), Proceedings of the Fifth International Conference on Cognitive Neurodynamics*. Springer, 2015.

F. Hůla, M. Ruman, and M. Kárný. Adaptive proposer for ultimatum game. In *Artificial Neural Networks and Machine Learning – Proceedings ICANN 2016*, pages 330–338. Barcelona, 2016.

J.C. Harsanyi. Games with incomplete information played by Bayesian players, I–III. *Management Science*, 50(12), 2004. ISSN 0025-1909. Supplement.

B.D. Jones. Bounded rationality. *Annual Rev. Polit. Sci.*, 2:297–321, 1999.

P.A. Ortega and A.A. Stocker. Human decision-making under limited time. *arXiv preprint arXiv:1610.01698*, 2016.

P.A. Ortega, D.A. Braun, J. Dyer, K.E. Kim, and N. Tishby. Information-theoretic bounded rationality. *arXiv preprint arXiv:1512.06789v1*, 2016.

V. Peterka. Bayesian system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.

M.L. Puterman. *Markov Decision Processes*. Wiley, 1994.

M. Regenwetter, J. Dana, and C.P. Davis-Stober. Transitivity of preferences. *Psychological Review*, 118(1):42–56, 2011.

A. Rubinstein. Perfect equilibrum in a bargaining model. *Econometrica*, 50:97–109, 1994. 1982.

M. Ruman, F. Hůla, M. Kárný, and T.V. Guy. Deliberation-aware responder in multi-proposer ultimatum game. In *Artificial Neural Networks and Machine Learning – Proceedings ICANN 2016*, pages 230–237. Barcelona, 2016.

J. Si, A.G. Barto, W.B. Powell, and D. Wunsch, editors. *Handbook of Learning and Approximate Dynamic Programming*, Danvers, May 2004. Wiley-IEEE Press. ISBN 0-471-66054-X.

R.H. Thaler. From homo economicus to homo sapiens. *Journal of Economic Perspectives*, 14:133–141, 2000.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, New York, London, 1944.

Proceedings of Machine Learning Research 58 (2016) 61-74

Optimal Number of Choices in Rating Contexts

Sam Ganzfried SGANZFRI@CIS.FIU.EDU
Florida International University
School of Computing and Information Sciences

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

In many settings people must give numerical scores to entities from a small discrete set. For instance, rating physical attractiveness from 1–5 on dating sites, or papers from 1–10 for conference reviewing. We study the problem of understanding when using a different number of options is optimal. For concreteness we assume the true underlying scores are integers from 1–100. We consider the case when scores are uniform random and Gaussian. While in theory for this setting it would be optimal to use all 100 options, in practice this is prohibitive, and it is preferable to utilize a smaller number of options due to humans' cognitive limitations. Our results suggest that using a smaller number of options than is typical could be optimal in certain situations. This would have many potential applications, as settings requiring entities to be ranked by humans are ubiquitous.

1. Introduction

Humans rate items or entities in many important settings. For example, users of dating websites and mobile applications rate other users' physical attractiveness, teachers rate scholarly work of students, and reviewers rate the quality of academic conference submissions. In these settings, the users assign a numerical (integral) score to each item from a small discrete set. However, the number of options in this set can vary significantly between applications, and even within different instantiations of the same application. For instance, for rating attractiveness, three popular sites all use a different number of options. On "Hot or Not," users rate the attractiveness of photographs submitted voluntarily by other users on a scale of 1–10 (Figure 1[1]). These scores are aggregated and the average is assigned as the overall "score" for a photograph. On the dating website OkCupid, users rate other users on a scale of 1–5 (if a user rates another user 4 or 5 then the rated user receives a notification)[2] (Figure 2[3]). And on the mobile application Tinder users "swipe right" (green heart) or "swipe left" (red X) to express interest in other users (two users are allowed to message each other if they mutually swipe right), which is essentially equivalent to using a binary $\{1, 2\}$ scale (Figure 3[4]). Education is another important application area requiring human ratings. For the 2016 International Joint Conference on Artificial Intelligence, reviewers assigned a "Summary Rating" score from -5–5 (equivalent to 1–10) for each submitted paper (Figure 4).[5]). The papers are then

1. http://blog.mrmeyer.com/2007/are-you-hot-or-not/
2. The likelihood of receiving an initial message is actually much more highly correlated with the variance—and particularly the number of "5" ratings—than with the average rating (Fry, 2015).
3. http://blog.okcupid.com/index.php/the-mathematics-of-beauty/
4. https://tctechcrunch2011.files.wordpress.com/2015/11/tinder-two.jpg
5. https://easychair.org/conferences/?conf=ijcai16

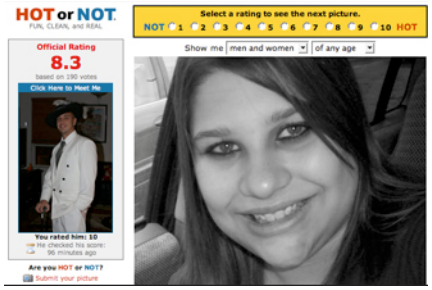

Figure 1: Hot or Not users rate attractiveness 1–10.

Figure 2: OkCupid users rate attractiveness 1–5.

Figure 3: Tinder users rate attractiveness 1–2.

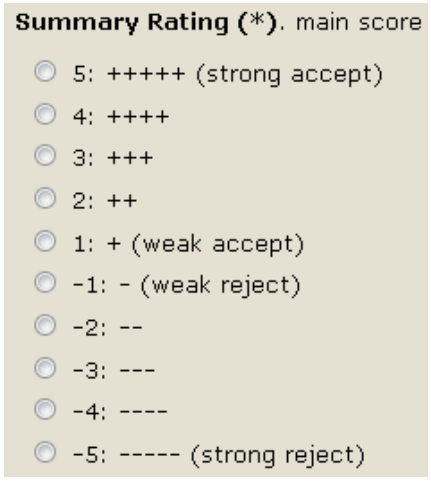

Figure 4: IJCAI reviewers rate papers -5–5.

discussed and scores aggregated to produce an acceptance or rejection decision based largely on the average of the scores.

Despite the importance and ubiquity of the problem, there has been little fundamental research done on the problem of determining the optimal number of options to allow in such settings. We study a model in which users have a underlying integral ground truth score for each item in $\{1,\ldots,n\}$ and are required to submit an integral rating in $\{1,\ldots,k\}$, for $k \ll n$. (For ease of presentation we use the equivalent formulation $\{0,\ldots,n-1\}$, $\{0,\ldots,k-1\}$.) We use two generative models for the ground truth scores: a uniform random model in which the fraction of scores for each value from 0 to $n-1$ is chosen uniformly at random (by choosing a random value for each and then normalizing), and a model where scores are chosen according to a Gaussian distribution with a given mean and variance. We then compute a "compressed" score distribution by mapping each full score s from $\{0,\ldots,n-1\}$ to $\{0,\ldots,k-1\}$ by applying

$$s \leftarrow \left\lfloor \frac{s}{\binom{n}{k}} \right\rfloor . \tag{1}$$

We then compute the average "compressed" score a_k, and compute its error e_k according to

$$e_k = \left| a_f - \frac{n-1}{k-1} \cdot a_k \right|, \tag{2}$$

where a_f is the ground truth average score. The goal is to pick $\text{argmin}_k e_k$. While there are many possible generative models and cost functions to use, these seemed like the most natural ones to start with. We leave study of alternative choices for future work.

We derive a closed-form expression for e_k that depends on only a small number (k) of parameters of the underlying distribution for an arbitrary distribution.[6] This allows us to exactly characterize the performance of using each number of choices. In computational simulations we repeatedly compute e_k and compare the average values. We focus on $n = 100$ and $k = 2, 3, 4, 5, 10$, which we believe are the most natural and interesting choices for initial study.

One could argue that this model is somewhat "trivial" in the sense that it would be optimal to set $k = n$ to permit all the possible scores, as this would result in the "compressed" scores agreeing exactly with the full scores. However, there are several reasons that would lead us to prefer to select $k \ll n$ in practice (as all of the examples previously described have done), thus making this "thought experiment" worthwhile. It is much easier for a human to assign a score from a small set than from a large set, particularly when rating many items under time constraints. We could have included an additional term into the cost function e_k that explicitly penalizes larger values of k, which would have a significant effect on the optimal value of k (providing a favoritism for smaller values). However the selection of this function would be somewhat arbitrary and would make the model more complex, and we leave this for future study. Given that we do not include such a penalty term, one may expect that increasing k will always decrease e_k in our setting. While the simulations show a clear negative relationship between k and e_k, we show that smaller values of k can actually lead to smaller e_k surprisingly often. These smaller values would receive further preference with a penalty term.

The most closely related theoretical work studies the impact of using finely grained numerical grades (e.g., 100, 99, 98) vs. coarse letter grades (e.g., A, B, C) (Dubey and Geanakoplos, 2010). They conclude that if students care primarily about their rank relative to the other students, they are often best motivated to work by assigning them to coarse categories (letter grades) than by the exact numerical exam scores. In a specific setting of "disparate" student abilities they show that the optimal absolute grading scheme is always coarse. Their model is game-theoretic; each player (student) selects an effort level, seeking to optimize a utility function that depends on both the relative score and effort level. Their setting is quite different from ours in many ways. For one, they study a setting where it is assumed that the underlying "ground truth" score is known, yet may be disguised for strategic reasons. In our setting the goal is to approximate the ground truth score as closely as possible.

While we are not aware of prior theoretical study of our problem, there have been experimental studies on the optimal number of options on a "Likert scale" (Matell and Jacoby, 1971; Wildt and Mazis, 1978; Cox III, 1980; Friedman et al., 1981). The general conclusion is that "the optimal number of scale categories is content specific and a function of the conditions of measurement" (Garland,

6. For theoretical simplicity we theoretically study a continuous version where scores are chosen according to a distribution over $(0, n)$ (though the simulations are for the discrete version) and the compressed scores are over $\{0, \ldots, k-1\}$. Continuous approximations for large discrete spaces have been studied in other settings; for instance, they have led to simplified analysis and insight in poker games with continuous distributions of private information (Ankenman and Chen, 2006).

1991). There has been study of whether including a "mid-point" option (i.e., the middle choice from an odd number) is beneficial. One experiment demonstrated that the use of the mid-point category decreases as the number of choices increases: 20% of respondents choose the mid-point for 3 and 5 options while only 7% did for $7, 9, \ldots, 19$ options (Matell and Jacoby, 1972). They conclude that it is preferable to either not include a mid-point at all or use a large number of options. Subsequent experiments demonstrated that eliminating a mid-point option can reduce social desirability bias which results from respondents' desires to please the interviewer or not give what they perceive to be a socially unacceptable answer (Garland, 1991).

2. Theoretical characterization

Suppose scores are given by continuous pdf f (with cdf F) on $(0, 100)$, and we wish to compress them to two options, $\{0, 1\}$. Scores below 50 are mapped to 0 and scores above 50 to 1.

The average of the full distribution is

$$a_f = E[X] = \int_{x=0}^{100} xf(x)dx.$$

The average of the compressed version is

$$\begin{aligned} a_2 &= \int_{x=0}^{50} 0f(x)dx + \int_{x=50}^{100} 1f(x)dx = \int_{x=50}^{100} f(x)dx \\ &= F(100) - F(50) = 1 - F(50). \end{aligned}$$

So $e_2 = |a_f - 100(1 - F(50))| = |E[X] - 100 + 100F(50)|$.

For three options,

$$\begin{aligned} a_3 &= \int_{x=0}^{100/3} 0f(x)dx + \int_{x=100/3}^{200/3} 1f(x)dx + \int_{x=200/3}^{100} 2f(x)dx \\ &= F(200/3) - F(100/3) + 2(1 - F(200/3)) \\ &= 2 - F(100/3) - F(200/3) \\ e_3 &= |a_f - 50(2 - F(100/3) - F(200/3))| \\ &= |E[X] - 100 + 50F(100/3) + 50F(200/3)| \end{aligned}$$

In general for n total and k compressed options,

$$\begin{aligned} a_k &= \sum_{i=0}^{k-1} \int_{x=\frac{ni}{k}}^{\frac{n(i+1)}{k}} if(x)dx \\ &= \sum_{i=0}^{k-1} \left[i \left(F\left(\frac{n(i+1)}{k}\right) - F\left(\frac{ni}{k}\right) \right) \right] \\ &= (k-1)F(n) - \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right) \\ &= (k-1) - \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right) \end{aligned}$$

$$\begin{aligned} e_k &= \left| a_f - \frac{n}{k-1}\left((k-1) - \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right)\right)\right| \\ &= \left| E[X] - n + \frac{n}{k-1}\sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right)\right| \end{aligned} \tag{3}$$

Equation 3 allows us to characterize the relative performance of choices of k for a given distribution f. For each k the characterization requires only knowing k statistics of f (the $k-1$ values of $F\left(\frac{ni}{k}\right)$ plus $E[X]$). In practice these could likely be closely approximated from historical data for small values of k.

As an example we see that $e_2 < e_3$ iff

$$|E[X] - 100 + 100F(50)| < |E[X] - 100 + 50F(100/3) + 50F(200/3)|$$

Consider a full distribution that has half its mass right around 30 and half its mass right around 60 (Figure 5). Then $a_f = E[X] = 0.5 \cdot 30 + 0.5 \cdot 60 = 45$. If we use $k = 2$, then the mass at 30 will be mapped down to 0 (since $30 < 50$) and the mass at 60 will be mapped up to 1 (since $60 > 50$) (Figure 6). So $a_2 = 0.5 \cdot 0 + 0.5 \cdot 1 = 0.5$. Using normalization of $\frac{n}{k} = 100$, $e_2 = |45 - 100(0.5)| = |45 - 50| = 5$. If we use $k = 3$, then the mass at 30 will also be mapped down to 0 (since $0 < \frac{100}{3}$); but the mass at 60 will be mapped to 1 (not the maximum possible value of 2 in this case), since $\frac{100}{3} < 60 < \frac{200}{3}$ (Figure 7). So again $a_3 = 0.5 \cdot 0 + 0.5 \cdot 1 = 0.5$, but now using normalization of $\frac{n}{k} = 50$ we have $e_3 = |45 - 50(0.5)| = |45 - 25| = 20$. So, surprisingly, in this example allowing more ranking choices actually significantly increases error.

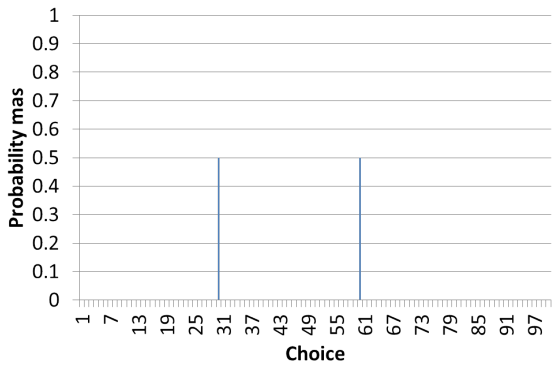

Figure 5: Example distribution for which compressing with $k = 2$ produces lower error than $k = 3$.

3. Rounding compression

An alternative model we could have considered is to use rounding to produce the compressed scores as opposed to using the floor function from Equation 1. For instance, for the case $n = 100, k = 2$, instead of dividing s by 50 and taking the floor, we could partition the points according to whether

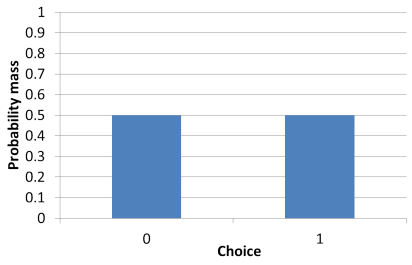

Figure 6: Compressed distribution using $k = 2$.

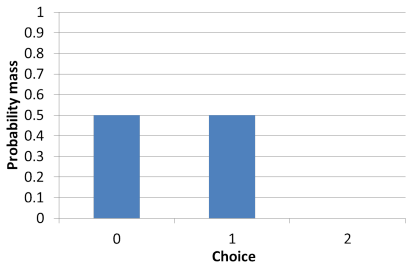

Figure 7: Compressed distribution using $k = 3$.

they are closer to $t_1 = 25$ or $t_2 = 75$. In the example above, the mass at 30 would be mapped to t_1 and the mass at 60 would be mapped to t_2. This would produce a compressed average score of

$$a_2 = \frac{1}{2} \cdot 25 + \frac{1}{2} \cdot 75 = 50.$$

No normalization would be necessary, and this would produce error of

$$e_2 = |a_f - a_2| = |45 - 50| = 5,$$

as the floor approach did as well. Similarly, for $k = 3$ the region midpoints will be $q_1 = \frac{100}{6}$, $q_2 = 50$, $q_3 = \frac{500}{6}$. The mass at 30 will be mapped to $q_1 = \frac{100}{6}$ and the mass at 60 will be mapped to $q_2 = 50$. This produces a compressed average score of

$$a_3 = \frac{1}{2} \cdot \frac{100}{6} + \frac{1}{2} \cdot 50 = \frac{100}{3}.$$

This produces an error of

$$e_3 = |a_f - a_3| = \left|45 - \frac{100}{3}\right| = \frac{35}{3} = 11.67$$

Although the error for $k = 3$ is smaller than for the floor, it is still significantly larger than the error for $k = 2$, and using two options still outperforms using three for the example in this new model.

In general, this approach would create k "midpoints" $\{m_i^k\}$:

$$m_i^k = \frac{n(2i-1)}{2k}$$

For $k = 2$ we have

$$\begin{aligned} a_2 &= \int_{x=0}^{50} 25 + \int_{x=50}^{100} 75 = 25F(50) + 75(1 - F(50)) = 75 - 50F(50) \\ e_2 &= |a_f - (75 - 50F(50))| = |E[X] - 75 + 50F(50)| \end{aligned}$$

One might wonder whether the floor approach would ever outperform the rounding approach (in the example above the rounding approach produced lower error $k = 3$ and the same error for $k = 2$). As a simple example to see that it can, consider the distribution with all mass on 0. The

floor approach would produce $a_2 = 0$ giving an error of 0, while the rounding approach would produce $a_2 = 25$ giving an error of 25. Thus, the superiority of the approach is dependent on the distribution. We explore this further in the experiments.

For three options,

$$\begin{aligned}
a_3 &= \int_{x=0}^{100/3} 100/6 f(x)dx + \int_{x=100/3}^{200/3} 50 f(x)dx + \int_{x=200/3}^{100} 500/6 f(x)dx \\
&= 100/6F(100/3) + 50(F(200/3) - F(100/3)) + 500/6(1 - F(200/3) \\
&= 100/6F(100/3) + 50F(200/3) - 50F(100/3) + 500/6 - 500/6F(200/3) \\
&= 500/6 - 100/3F(100/3) - 100/3F(200/3) \\
e_3 &= |a_f - (500/6 - 100/3F(100/3) - 100/3F(200/3))| \\
&= |E[X] - 500/6 + 100/3F(100/3) + 100/3F(200/3)|
\end{aligned}$$

For general n and k,

$$\begin{aligned}
a_k &= \sum_{i=0}^{k-1} \int_{x=\frac{ni}{k}}^{\frac{n(i+1)}{k}} m_{i+1}^k f(x)dx \\
&= \sum_{i=0}^{k-1} \int_{x=\frac{ni}{k}}^{\frac{n(i+1)}{k}} \frac{n(2i+1)}{2k} f(x)dx \\
&= \sum_{i=0}^{k-1} \left[\frac{n(2i+1)}{2k} \left(F\left(\frac{n(i+1)}{k}\right) - F\left(\frac{ni}{k}\right) \right) \right] \\
&= \frac{n}{k} \sum_{i=0}^{k-1} \left[i \left(F\left(\frac{n(i+1)}{k}\right) - F\left(\frac{ni}{k}\right) \right) \right] + \frac{n}{2k} \sum_{i=0}^{k-1} \left[\left(F\left(\frac{n(i+1)}{k}\right) - F\left(\frac{ni}{k}\right) \right) \right] \\
&= \frac{n}{k} \left[(k-1) - \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right) \right] + \frac{n}{2k} \\
&= \frac{n(2k-1)}{2k} - \frac{n}{k} \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right) \\
e_k &= \left| a_f - \left[\frac{n(2k-1)}{2k} - \frac{n}{k} \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right) \right] \right| = \left| E[X] - \frac{n(2k-1)}{2k} + \frac{n}{k} \sum_{i=1}^{k-1} F\left(\frac{ni}{k}\right) \right| \quad (4)
\end{aligned}$$

Like for the floor model e_k requires only knowing k statistics of f. The rounding model has an advantage over the floor model that there is no need to convert scores between different scales and perform normalization. One drawback is that it requires knowing n (the expression for m_i^k is dependent on n), while the floor model does not. In our experiments we assume $n = 100$, but in practice it may not be clear what the agents' ground truth granularity is and may be easier to just deal with scores from 1 to k. Furthermore, it may seem unnatural to essentially ask people to rate items as "$\frac{100}{6}, 50, \frac{200}{6}$" rather than "$1, 2, 3$" (though the conversion between the score and m_i^k could be done behind the scenes essentially circumventing the potential practical complication). We note that one could generalize both the floor and rounding model by using a score of $s(n,k)_i$

for the i'th region. For the floor setting we set $s(n,k)_i = i$, and for the rounding setting we set $s(n,k)_i = m_i^k = \frac{n(2i+1)}{2k}$.

4. Computational simulations

The above analysis leads to the immediate question of whether the example for which $e_2 < e_3$ was just a fluke or whether using fewer choices can actually reduce error under reasonable assumptions on the generative model. We study this question using simulations which we believe are the two most natural models. While we have studied the continuous setting where the full set of options is continuous over $(0, n)$ and the compressed set of options is the discrete space $\{0, \ldots, k-1\}$, we will now consider the perhaps more realistic setting where the full set is the discrete set $\{0, \ldots, n-1\}$, and the compressed set is $\{0, \ldots, k-1\}$ (though it should be noted that the two settings are likely extremely similar qualitatively).

The first generative model we consider is a uniform model in which the values of the pmf p_f for each of the n possible values are chosen independently and uniformly at random. Formally, we construct a histogram of n scores for p_f according to Algorithm 1. We then compress the full scores to a compressed distribution p_k by applying Algorithm 2.

The second generative model is a Gaussian model in which the values are generated according to a normal distribution with specified mean μ and standard deviation σ. This model also takes as a parameter a number of samples s to use for generating the scores. The procedure is given by Algorithm 3. As for the uniform setting, Algorithm 2 is then used to compress the scores.

Algorithm 1 Procedure for generating full scores in uniform model

Inputs: Number of scores n

```
scoreSum ← 0
for i = 0 : n do
    r ← random(0,1)
    scores[i] ← r
    scoreSum = scoreSum +r
for i = 0 : n do
    scores[i] = scores[i] / scoreSum
```

Algorithm 2 Procedure for compressing scores

Inputs: scores[], number of total scores n, desired number of compressed scores k

$Z(n,k) \leftarrow \frac{n}{k}$ ▷ Normalization

for $i = 0 : n$ **do**

    scoresCompressed $\left[\left\lfloor \frac{i}{Z(n,k)} \right\rfloor\right]$ += scores[i]

For our simulations we used $n = 100$, and considered $k = 2, 3, 4, 5, 10$, which are popular and natural values. For the Gaussian model we used $s = 1000$, $\mu = 50$, $\sigma = \frac{50}{3}$. For each set of simulations we computed the errors for all considered values of k for $m = 100,000$ "items" (each corresponding to a different distribution generated according to the specified model). The main quantities we are interested in computing are the number of times that each value of k produces the lowest error over the m items, and the average value of the errors over all items for each k value.

Algorithm 3 Procedure for generating full scores in Gaussian model

Inputs: Number of scores n, number of samples s, mean μ, standard deviation σ

```
for i = 0 : s do
    r ← randomGaussian(μ, σ)
    if r < 0 then
        r = 0
    else if r > n − 1 then
        r ← n − 1
    ++scores[round(r)]
for i = 0 : n do
    scores[i] = scores[i] / s
```

In the first set of experiments, we compared performance between using $k = 2, 3, 4, 5, 10$ to see for how many of the trials each value of k produced the minimal error. The results are in Table 1. Not surprisingly, we see that the number of victories increases monotonically with the value of k, while the average error decreased monotonically (recall that we would have zero error if we set $k = 100$). However, what is perhaps surprising is that using a smaller number of compressed scores produced the optimal error in a far from negligible number of the trials. For the uniform model, using 10 scores minimized error only around 53% of the time, while using 5 scores minimized error 17% of the time, and even using 2 scores minimized it 5.6% of the time. The results were similar for the Gaussian model, though a bit more in favor of larger values of k, which is what we would expect because the Gaussian model is less likely to generate "fluke" distributions that could favor the smaller values.

	2	3	4	5	10
Uniform # victories	5564	9265	14870	16974	53327
Uniform average error	1.32	0.86	0.53	0.41	0.19
Gaussian # victories	3025	7336	14435	17800	57404
Gaussian average error	1.14	0.59	0.30	0.22	0.10

Table 1: Number of times each value of k in $\{2,3,4,5,10\}$ produces minimal error and average error values, over 100,000 items generated according to both models.

We next explored the number of victories between just $k = 2$ and $k = 3$, with results in Table 2. Again we observed that using a larger value of k generally reduces error, as expected. However, we find it extremely surprising that using $k = 2$ produces a lower error 37% of the time. As before, the larger k value performs relatively better in the Gaussian model. We also looked at results for the most extreme comparison, $k = 2$ vs. $k = 10$. These results are provided in Table 3. Using 2 scores outperformed 10 8.3% of the time in the uniform setting, which was larger than we expected. In Figures 8–10, we present a distribution for which $k = 2$ particularly outperformed $k = 10$.

We next repeated the extreme $k = 2$ vs. 10 comparison, but we imposed a restriction that the $k = 10$ option could not give a score below 3 or above 6. (If it selected a score below 3 then we set it to 3, and if above 6 we set it to 6). These result are given in Table 4. For some settings, for instance the paper reviewing setting, extreme scores are very uncommon, and we strongly suspect that the

	2	3
Uniform number of victories	36805	63195
Uniform average error	1.31	0.86
Gaussian number of victories	30454	69546
Gaussian average error	1.13	0.58

Table 2: Number of times each value of k in $\{2,3\}$ produces minimal error and average error values, over 100,000 items generated according to both generative models.

	2	10
Uniform number of victories	8253	91747
Uniform average error	1.32	0.19
Gaussian number of victories	4369	95631
Gaussian average error	1.13	0.10

Table 3: Number of times each value of k in $\{2,10\}$ produces minimal error and average error values, over 100,000 items generated according to both generative models.

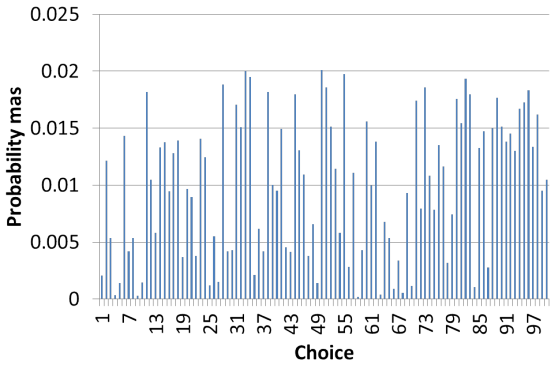

Figure 8: Example distribution for which compressing with $k = 2$ produces significantly lower error than $k = 10$. The full distribution has mean 54.188, while the $k = 2$ compression has mean 0.548 (54.253 after normalization) and the $k = 10$ compression has mean 5.009 (55.009 after normalization). The normalized errors between the means were 0.906 for $k = 10$ and 0.048 for $k = 2$, yielding a difference of 0.859 in favor of $k = 2$.

vast majority of scores are in this middle range. Some possible explanations are that reviewers who give extreme scores may be required to put in additional work to justify their scores, and are more likely to be involved in arguments with the other reviewers (or with the authors in the rebuttal). Reviewers could also experience higher regret or embarrassment for being "wrong" and possibly off-base in the review by missing an important nuance. In this setting using $k = 2$ outperforms $k = 10$ nearly $\frac{1}{3}$ of the time in the uniform model.

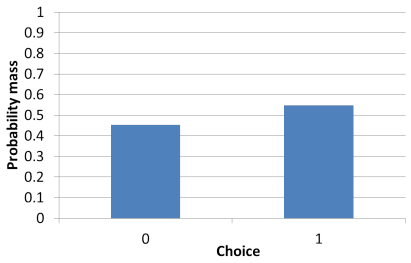

Figure 9: Compressed distribution for $k = 2$.

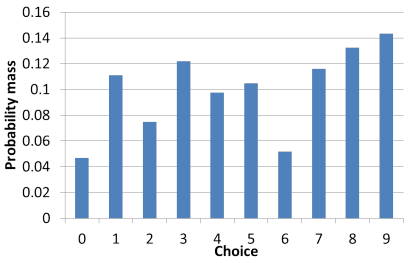

Figure 10: Compressed distribution for $k = 10$.

	2	10
Uniform number of victories	32250	67750
Uniform average error	1.31	0.74
Gaussian number of victories	10859	89141
Gaussian average error	1.13	0.20

Table 4: Number of times each k in $\{2,10\}$ produces minimal error and average error values, over 100,000 items generated according to both models. For $k = 10$, we only permitted scores between 3 and 6 (inclusive). If a score was below 3 we set it to be 3, and above 6 to 6.

We also considered the situation where we restricted the $k = 10$ scores to fall between 3 and 7 (as opposed to 3 and 6). Note that the possible scores range from 0–9, so this restriction is asymmetric in that the lowest three possible scores are eliminated while only the highest two are. This is motivated by the intuition that raters may be less inclined to give extremely low scores which may hurt the feelings of an author (for the case of paper reviewing). In this setting, which is seemingly quite similar to the 3–6 setting, $k = 2$ produced lower error 93% of the time in the uniform model!

	2	10
Uniform number of victories	93226	6774
Uniform average error	1.31	0.74
Gaussian number of victories	54459	45541
Gaussian average error	1.13	1.09

Table 5: Number of times each k in $\{2,10\}$ produces minimal error and average error values, over 100,000 items generated according to both models. For $k = 10$, we only permitted scores between 3 and 7 (inclusive). If a score was below 3 we set it to be 3, and above 7 to 7.

We next repeated these experiments for the rounding compression function. There are several interesting observations from Table 6. In this setting, $k = 3$ is the clear choice, performing best in both generative models (by a very significant margin for the Gaussian model). The smaller values of k perform significantly better in the rounding model than in the floor model (as indicated by lower average errors) while the larger values perform significantly worse, and their errors seem to

approach 0.5 for both models. Taking both compression functions into account, the optimal overall approach would still be to use the floor approach with $k = 10$, which produced the smallest average errors of 0.19 and 0.1 in the two models, while using $k = 3$ in the rounding setting produced errors of 0.47 and 0.24. The 2 vs. 3 experiments produced very similar results for the two compressions (Table 7). The 2 vs. 10 results were quite different, with 2 performing better almost 40% of the time with rounding, vs. less than 10% with the floor function (Table 8). In the 2 vs. 10 truncated 3–6 experiments 2 performed relatively better in the rounding setting for both models (Table 9), and for the 2 vs. 10 truncated 3–7 experiments $k = 2$ performed better nearly all the time (Table 10).

	2	3	4	5	10
Uniform # victories	15766	33175	21386	19995	9678
Uniform average error	0.78	0.47	0.55	0.52	0.50
Gaussian # victories	13262	64870	10331	9689	1848
Gaussian average error	0.67	0.24	0.50	0.50	0.50

Table 6: Number of times each value of k in $\{2,3,4,5,10\}$ produces minimal error and average error values, over 100,000 items generated from both models with rounding compression.

	2	3
Uniform number of victories	33585	66415
Uniform average error	0.78	0.47
Gaussian number of victories	18307	81693
Gaussian average error	0.67	0.24

Table 7: Number of times each value of k in $\{2,3\}$ produces minimal error and average error values, over 100,000 items generated according to both models with rounding compression.

	2	10
Uniform number of victories	37225	62775
Uniform average error	0.78	0.50
Gaussian number of victories	37897	62103
Gaussian average error	0.67	0.50

Table 8: Number of times each value of k in $\{2,10\}$ produces minimal error and average error values, over 100,000 items generated according to both models with rounding compression.

5. Conclusion

Settings in which humans must rate items or entities from a small discrete set of options are ubiquitous. We have singled out several important applications—rating attractiveness for dating websites and mobile applications, assigning grades to students, and reviewing academic papers for confer-

	2	10
Uniform number of victories	55676	44324
Uniform average error	0.79	0.89
Gaussian number of victories	24128	75872
Gaussian average error	0.67	0.34

Table 9: Number of times each value of k in $\{2,10\}$ produces minimal error and average error values, over 100,000 items generated according to both models with rounding compression. For $k = 10$, we only permitted scores between 3 and 6 (inclusive). If a score was below 3 we set it to be 3, and above 6 to 6.

	2	10
Uniform number of victories	99586	414
Uniform average error	0.78	3.50
Gaussian number of victories	95692	4308
Gaussian average error	0.67	1.45

Table 10: Number of times each value of k in $\{2,10\}$ produces minimal error and average error values, over 100,000 items generated according to both generative models with rounding compression. For $k = 10$, we only permitted scores between 3 and 7 (inclusive). If a score was below 3 we set it to be 3, and above 7 to 7.

ences. The number of available options can vary considerably, even within different instantiations of the same application. For instance, we saw that three popular sites for the attractiveness rating problem use completely different systems: Hot or Not uses a 1–10 system, OkCupid uses 1–5 "star" system, and Tinder uses a binary 1–2 "swipe" system.

Despite the ubiquity and importance of the problem of selecting the optimal number of rating choices, we have not seen it studied theoretically previously. Our goal is to select k to minimize the average (normalized) error between the compressed average score and the ground truth average. We studied two natural models for generating the scores. The first is a uniform model where the scores are selected independently and uniformly at random, and the second is a Gaussian model where they are selected according to a more structured procedure that gives more preference for the options near the center.

We provided a closed-form solution for continuous distributions with arbitrary cdf. This allows us to characterize the relative performance of choices of k for a given distribution. We saw that, counterintuitively, using a smaller value of k can actually produce a smaller error for some distributions (even though we know that as k approaches n the error approaches 0). We presented a specific example distribution f for which using $k = 2$ outperforms $k = 3$.

We performed numerous computational simulations, comparing the performance between different values of k for different generative models and metrics. The main metric we used was the absolute number of times for which values of k produced the minimal error. We also considered the average error over all simulated items. Not surprisingly, we observe that performance generally improves monotonically with increased k, and more so for the Gaussian model than the uniform.

However, we observe that small k values can be optimal a non-negligible amount of the time, which is perhaps counterintuitive. In fact, using $k = 2$ outperformed $k = 3, 4, 5$, and 10 on 5.6% of the trials in the uniform setting. Just comparing 2 vs. 3, $k = 2$ performed better around 37% of the time. Using $k = 2$ outperformed $k = 10$ 8.3% of the time, and significantly more as we imposed some very natural restrictions on the $k = 10$ setting that are motivated by intuitive phenomena. When we restricted the $k = 10$ to only assign values between 3 and 7 (inclusive), using $k = 2$ actually produced lower error 93% of the time! This could correspond to a setting where raters are ashamed to assign extreme scores (particularly extreme low scores).

We compared two different natural compression rules—one based on the floor function and one based on rounding—and weighed the pros and cons of each. For smaller values of k the rounding approach leads to significantly lower error than the floor approach, with $k = 3$ being the clear optimal choice, while for larger values of k rounding leads to significantly higher error.

One avenue for future study would be to extend our theoretical characterization analysis in order to get a better understanding of the specific distributions for which different values of k are optimal, as opposed to our experimental results which are in aggregate over many distributions. Specific application domains will have distributions with different properties, and improved understanding will allow us to determine which k value is optimal for the types of distributions we expect to encounter for a given domain. This improved theoretical understanding can be coupled with exploring data on specific applications of interest.

References

Jerrod Ankenman and Bill Chen. *The Mathematics of Poker*. ConJelCo LLC, 2006.

Eli P. Cox III. The optimal number of response alternatives for a scale: A review. *Journal of Marketing Research*, 17:407–442, 1980.

Pradeep Dubey and John Geanakoplos. Grading exams: 100, 99, 98, ... or A, B, C? *Games and Economic Behavior*, 69:72–94, May 2010.

Hershey H. Friedman, Yonah Wilamowsky, and Linda W. Friedman. A comparison of balanced and unbalanced rating scales. *The Mid-Atlantic Journal of Business*, 19(2):1–7, 1981.

Hannah Fry. *The Mathematics of Love*. TED Books, 2015.

Ron Garland. The mid-point on a rating scale: is it desirable? *Marketing Bulletin*, 2:66–70, 1991. Research Note 3.

Michael S. Matell and Jacob Jacoby. Is there an optimal number of alternatives for likert scale items? Study 1: reliability and validity. *Educational and psychological measurements*, 31:657–674, 1971.

Michael S. Matell and Jacob Jacoby. Is there an optimal number of alternatives for likert scale items? Effects of testing time and scale properties. *Journal of Applied Psychology*, 56(6):506–509, 1972.

Albert R. Wildt and Michael B. Mazis. Determinants of scale response: label versus position. *Journal of Marketing Research*, 15:261–267, 1978.

Proceedings of Machine Learning Research 58 (2016) 75-85

On Learning Decision Heuristics

Özgür Şimşek OZGUR@MPIB-BERLIN.MPG.DE
Marcus Buckmann BUCKMANN@MPIB-BERLIN.MPG.DE
Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, Germany

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

Decision heuristics are simple models of human and animal decision making that use few pieces of information and combine the pieces in simple ways, for example, by giving them equal weight or by considering them sequentially. We examine how decision heuristics can be learned—and modified—as additional training examples become available. In particular, we examine how additional training examples change the variance in parameter estimates of the heuristic. Our analysis suggests new decision heuristics, including a family of heuristics that generalizes two well-known families: lexicographic heuristics and tallying. We evaluate the empirical performance of these heuristics in a large, diverse collection of data sets.

1. Introduction

Decision heuristics (Gigerenzer et al., 1999, 2011) are models of human and animal decision making. They use few pieces of information and combine the pieces in simple ways, for example, by giving them equal weight or by considering them sequentially. We examine how such heuristics can be learned from training examples. Our motivation is both computational and cognitive. From a computational viewpoint, we want to develop heuristics that are fast in computation, frugal in information use, and effective in making good decisions. From a cognitive viewpoint, we want to understand how people and animals create and modify decision heuristics over time, as they accumulate experiences. While there is a large literature on decision heuristics, few studies have examined the learning process.

Simple decision heuristics are most widely studied within the context of comparison problems, where the objective is to identify which of a number of alternatives has the highest value in a specified (unobserved) criterion. We study the learning process in this context, examining two well-known families of heuristics: lexicographic heuristics and tallying. In particular, we examine the variance resulting from small training-set sizes and its effect on the building blocks of these heuristics.

Our theoretical analysis suggests new heuristics, including a generalization of lexicographic heuristics and tallying. We examine the performance of these heuristics in a large, diverse collection of data sets, comparing their performance to random forests. We find that sampling variance can have a large impact on the learning rate of heuristics and that very simple methods for accounting for sampling variance can substantially improve predictive accuracy.

2. Background

The comparison problem asks which of a given set of objects has the highest value on an unobserved criterion, given a number of attributes of the objects. We focus on pairwise comparisons, where exactly two objects are being compared. In the heuristics literature, *attributes* are called *cues*; we will follow this custom. We use x_A and y_A to denote the cue and criterion value, respectively, of object A.

For the comparison problem, two well-known families of heuristics are lexicographic heuristics and tallying. Both families decide by comparing the objects on one or more cues, asking which object has the more favorable cue value. Each cue is associated with a direction of inference, known as the *cue direction*, which can be positive or negative, favoring the object with the higher or lower cue value, respectively. Neither family requires the difference in cue values to be quantified. For example, if *height of a person* is a cue, one needs to be able to determine which of two people is taller but it is not necessary to know the height of either person or the magnitude of the difference.

Lexicographic heuristics (Fishburn, 1974) consider the cues one at a time, in a specified order, until they find a cue that *discriminates* between the objects, that is, one whose value differs on the two objects. The heuristic then decides based on that cue alone. An example is take-the-best (TTB; Gigerenzer and Goldsein, 1996), which orders cues with respect to decreasing validity on the training sample, where *validity* is the accuracy of the cue among pairwise comparisons on which the cue discriminates between the objects.

Tallying (Czerlinski et al., 1999) is a voting model. It determines how each cue votes on its own (selecting one or the other object or abstaining from voting) and selects the object with the highest number of votes, breaking ties randomly. Cue directions are set to the direction with the highest validity in the training set.

Note that the comparison problem has a symmetry: a comparison of A to B and B to A should agree on which object has the higher criterion value.

3. Distribution of sample statistics

A primary building block of decision heuristics is how a cue decides on its own, independently of the other cues. This is determined by the cue direction. If the direction is positive, the cue favors the alternative with the higher cue value; if it is negative, the cue favors the alternative with the lower cue value.

A second building block is how well a cue decides on its own, in other words, how accurate it is when it discriminates among the alternatives. This building block informs how the various cues should be integrated within the heuristic, for example, how they should be ordered in a lexicographic decision rule. For this building block, two quantities are relevant: the positive and negative validity, which are the probability that the cue makes the correct decision given that the cue discriminates between the alternatives if the cue is used in the positive or negative direction, respectively. *Cue validity* is the larger of positive and negative validity—it is the accuracy of the cue when it discriminates between the alternatives if the cue is used in the correct direction.

In earlier work (Şimşek and Buckmann, 2015), these building blocks were examined with a focus on expected rate of learning. Here, our main focus is on sampling variance in cue parameters.

When learning decision heuristics, cue directions and cue validities are estimated from a training sample, where each training instance corresponds to a single pairwise comparison between two objects. We assume that the instances in the training sample are independent.

From a comparison of object A to object B, the information we need is a single variable with three possible values: *positive* if the cue and the criterion move in the same direction, that is, if $(x_A - x_B) \times (y_A - y_B) > 0$; *negative* if the cue and the criterion move in opposite directions, that is, if $(x_A - x_B) \times (y_A - y_B) < 0$; and *neutral* otherwise. We can therefore denote a training sample with three numbers, $\{a, b, c\}$, where a is the number of positive instances, b the negative instances, and c the neutral instances.

Given a training sample, the estimate of cue direction, $\hat{d}$, is positive if $a > b$, negative if $a < b$, and positive or negative with equal probability if $a = b$. The estimate of cue validity, $\hat{v}$, is $\max\{a, b\}/(a + b)$. Notice that the value of c does not play a role in these estimates. Our analysis therefore focuses on samples with no neutral instances, where the sample size is $n = a + b$. We call n the number of *informative* instances. We denote the true validity and direction of the cue with v and d, respectively. In the analysis that follows, we assume, without loss of generality, that d is positive.

We examine the sample distributions of three variables. The first is $\hat{d}$, which is 1 if the cue-direction estimate from the sample is identical to the cue direction in the population, and 0 otherwise. The second variable is $\hat{v}$, and the third is o, which is the expected accuracy of the cue on an unseen test instance where the cue discriminates between the alternatives if the cue is used in the direction inferred from the sample. Our main objective in this section is to examine the variance in $\hat{d}$, $\hat{v}$, and o.

Lemma 1 Random variable $\hat{d}$ follows a Bernoulli distribution with probability of success $p_1 = \sum_{k=\lfloor n/2 \rfloor + 1}^{n} B(k, n, v) + 0.5 \times B(n/2, n, v)$, expected value p_1, and variance $p_1(1 - p_1)$, where $B(x, n, p) = \binom{n}{x} p^x (1 - p)^{n-x}$ denotes the binomial function.

Lemma 2 Random variable $\hat{v}$ has expected value v and variance $v(1 - v)/n$. This follows from $\hat{v} = a/n$ and the fact that a follows the binomial distribution with parameters n and v, with expected value nv and variance $nv(1 - v)$.

Lemma 3 Random variable o has expected value $p_1(2v - 1) + 1 - v$ and variance $p_1(1 - p_1)(2v - 1)^2$. Proof: $o = \hat{d}v + (1 - \hat{d})(1 - v) = \hat{d}(2v - 1) + 1 - v$. It follows that $E(o) = E(\hat{d})(2v - 1) + 1 - v$ and $Var(o) = Var(\hat{d})(2v - 1)^2$.

First, we briefly examine the expected prediction error (E) of a single cue:

$$\begin{aligned} E &= (1 - v) \times P(\hat{d} = 1) + v \times (1 - P(\hat{d} = 1)) \\ &= (1 - v)p_1 + v(1 - p_1) \\ &= \underbrace{(1 - v)}_{irreducible} + \underbrace{(2v - 1)(1 - p_1)}_{reducible} \end{aligned} \tag{1}$$

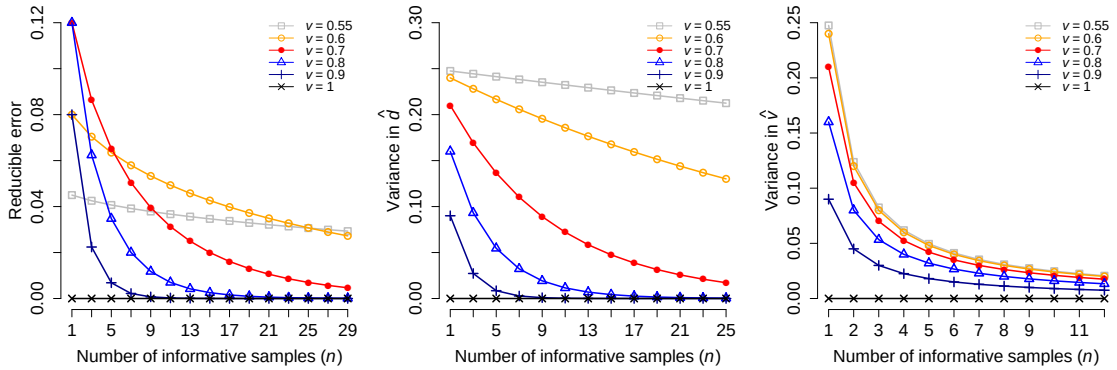

Figure 1: Reducible error, variance in $\hat{d}$, and variance in $\hat{v}$, as the number of informative samples increases.

The first term in Equation 1 is the irreducible error of the cue: This is the error that would be incurred even if the direction was known. The second term results from misestimation of the cue direction and becomes zero as sample size goes to infinity. Figure 1 (leftmost panel) shows the reducible error as a function of n for various v values. The plot shows an interesting pattern: When reducible error is high, it reduces rapidly; when it is low, it reduces very slowly. This conclusion was noted earlier (Şimşek and Buckmann, 2015); the analysis here is an alternative that shows it more clearly.

Figure 1 (middle and rightmost panel) shows how the variance in $\hat{d}$ and $\hat{v}$ decreases as sample size increases. Variance in sample validity reduces fairly rapidly within the first few samples. On the other hand, for cue direction, reduction in variance can be rapid or slow, depending on population validity. For high population validity, variance reduces rapidly. The closer the population validity is to 0.5, the slower the reduction in variance.

These results show that the sample variance of the cue-direction estimate varies substantially with population validity and with the number of informative samples. This has important consequences for learning heuristics. Even when the data set is complete, with no missing cue or criterion values, typically cues will vary (sometimes substantially) in the number of informative samples they have. For example, cues with lower discrimination rates will typically have smaller n.

Existing decision heuristics do not take this information into account. For example, tallying collects votes from all available cues, with no regard for how much uncertainty there is around the cue direction estimates. Similarly, TTB orders cues with respect to their sample validity, with no regard for the uncertainty around cue direction and validity estimates.

When the training set is large enough, the uncertainty in parameter estimates will diminish and not play a role, but with small training sets, there will often be substantial differences in how certain one is about the true direction and validity of the various available cues.

What can be done? In the next section, we turn our attention to the reverse inference problem: Given sample statistics, what is the true cue direction and validity in the population?

4. Inference on population statistics from a sample

We now examine how to make inferences on cue direction, cue validity, and o, given a training set $\{a, b, c\}$. First, we derive the posterior probability that population direction is positive given the training sample, $P(d_+|a, b, c)$. Note that the quantity c is irrelevant and $P(d_+|a, b, c) = P(d_+|a, b)$.

$$\begin{aligned} P(d_+|a,b) &= \frac{P(a,b|d_+)}{P(a,b)}\,P(d_+) = \frac{\int_{v=0}^{1} P(a,b|v,d_+)P(v|d_+)dv}{\int_{v=0}^{1} P(a,b|v)P(v)dv}\,P(d_+) \\ &= \frac{\int_{v=0.5}^{1} \binom{a+b}{a} v^a(1-v)^b\, 2\, P(v)\, dv}{\int_{v=0}^{1} \binom{a+b}{a} v^a(1-v)^b P(v)\, dv} 0.5 \\ &= \frac{\int_{v=0.5}^{1} v^a(1-v)^b P(v)\, dv}{\int_{v=0}^{1} v^a(1-v)^b P(v)\, dv} \end{aligned} \tag{2}$$

To arrive at Equation 2, we first used Bayes's rule, then conditioned on the population validity v, both in the numerator and in the denominator. Due to the symmetry of the comparison problem, $P(d_+)$ (prior probability of positive cue direction) is 0.5, and $P(v|d_+)$ is $2 \times P(v)$ if $v \geq 0.5$ and 0 otherwise.

Equation 2 uses $P(v)$, the prior on v. Figure 2 shows the distribution of cue validities in a large, diverse collection of natural data sets (described in section 6). The triangular distribution matches the validity distribution well.

Next, we derive the posterior distribution of population validity given sample statistics:

$$\begin{aligned} P(v|a,b) &= \frac{P(a,b|v)}{P(a,b)}\,P(v) = \frac{P(a,b|v)}{\int_{v=0}^{1} P(a,b|v)P(v)dv}\,P(v) \\ &= \frac{\binom{a+b}{a} v^a(1-v)^b}{\int_{v=0}^{1} \binom{a+b}{a} v^a(1-v)^b P(v)\, dv} P(v) \\ &= \frac{v^a(1-v)^b}{\int_{v=0}^{1} v^a(1-v)^b P(v)\, dv} P(v) \end{aligned} \tag{3}$$

To arrive at Equation 3, we first used Bayes's rule, then conditioned on the population validity v in the denominator.

And finally, we derive the posterior distribution of o, the probability that the cue will be accurate on a test instance, given that it discriminates between the alternatives, if the cue is used in the direction inferred from the training set.

$$\begin{aligned} P(o|a,b) &= \int_{v=0}^{1} P(o|v,a,b)P(v|a,b)\, dv \\ &= \int_{v=0}^{1} (v\hat{d} + (1-v)(1-\hat{d}))P(v|a,b)\, dv \end{aligned} \tag{4}$$

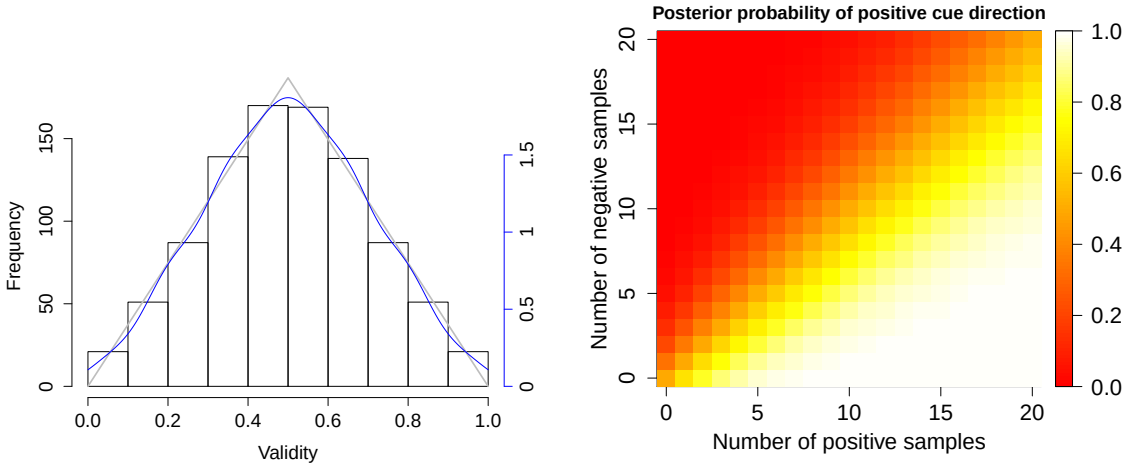

Figure 2: Left: A histogram showing the distribution of cue validity in 56 natural data sets. The blue line is a density estimate with a corresponding vertical axis to the right of the curve. The gray line is a triangular distribution that is a very close fit to the density. Right: The posterior probability that the population direction is positive given the number of positive and negative instances in the training set, computed by setting the prior on population validity to the triangular distribution shown on the left.

We now examine some of these quantities using the triangular distribution in Figure 2 as a prior on cue validity. Figure 2 (right panel) shows the posterior probability of positive cue direction as a function of a and b. The plot has a simple structure. There is a narrow band around the line $a = b$ where each additional informative sample changes the posterior substantially. For example, at (a = 3, b = 1), the posterior probability is 0.62, and after sampling one more positive instance, the posterior becomes 0.73. This band slowly expands in width as a and b increase.

The figure also shows that values of a and b that yield identical sample validities can have vastly different posterior probabilities of positive cue direction. Consider having four cues with training sets $\{2, 1\}$, $\{4, 2\}$, $\{6, 3\}$, and $\{8, 4\}$. All cues have sample validity 2/3 but their posterior probabilities of positive cue direction are 0.69, 0.78, 0.83, and 0.87, respectively. The general pattern is similar for the posterior probability that $o = 1$ (plot not shown).

It is important to note that while the computation of the posteriors is complex, the resulting posterior distributions exhibit simple patterns that simple rules for handling uncertainty would be able to express. This analysis shows that in early stages of learning, there are strong reasons to pay attention to the differences among cues in the level of uncertainty about their key parameters. In the next section, we consider three decision models that are sensitive to sample variance.

5. Decision models to consider

Motivated by the analysis in the previous sections, we consider three decision models. Two of the models are two extremes of the lexicographic decision rule where the training method ignores the dependencies among the cues. The first model answers, in a principled way, how to best take into account the differences in sample variances of the different cues. The second model asks to what extent this problem can be addressed using minimum effort and computation. As training-set size grows, both models converge to TTB. The third model takes a different perspective, and asks, if there is uncertainty about the order of the cues, why should one order the cues at all?

Lexicographic-by-posterior is a lexicographic decision rule that orders cues with respect to the posterior probability of a correct decision ($o = 1$), as computed by Equation 4, breaking ties randomly. This method is not computationally simple but represents an ideal for lexicographic decision rules that ignore the dependencies between the cues—it is the best that can be done. We call this method *lexipost* for short.

TTBS is a variation on TTB. It orders cues in decreasing order of sample validity (as TTB does) but breaks ties in favor of the cue with the higher number of informative samples. Variance in sample estimates of cue direction and cue validity reduces with increasing number of informative samples, and TTBS is one of the simplest, most straightforward methods of being sensitive to sample variance.

Lexicographic-tallying is a family of heuristics that generalizes lexicographic models and tallying. It is characterized by cue directions and cue levels. At decision time, at first, only the cues at level 1 are examined. These cues vote independently, and their independent decisions are tallied. If the result favors one or the other alternative, a decision is reached. Otherwise (if the tally at level 1 is neutral between the two alternatives), the cues at level 2 are tallied, and so on, until a decision is reached. If all cues are at level 1, the method reduces to tallying. If no cues share the same level, it reduces to a lexicographic decision rule. We refer to this model as *lexital* for short. We are not aware of earlier uses of this model even though it is a natural generalization of existing heuristics.

One motivation for using lexital is uncertainty in cue parameters. When there is not enough certainty about how a subset of the cues should be ordered (e.g., if they have equal sample validity and an equal number of informative samples), tallying these cues is a more reasonable approach than using them sequentially. There are other reasons for employing the hierarchical structure of lexital but uncertainty in cue parameters is a natural reason for doing so.

How should the parameters of a lexital model be determined from training samples? This is an open question, with many possible approaches. Here we consider perhaps the simplest approach: Order cues according to decreasing sample validity; when multiple cues tie on their sample validity, assign them to the same level in lexital. We call this model *tally-the-best*.

6. Empirical performance

We examined the performance of various heuristics in a large, diverse collection of natural data sets. The collection included 56 data sets gathered from a wide variety of sources,

including online data repositories, textbooks, packages for R statistical software, statistics and data-mining competitions, research publications, and individual scientists collecting field data. The subjects were diverse, including biology, business, computer science, ecology, economics, education, engineering, environmental science, medicine, political science, psychology, sociology, sports, and transportation. The data sets varied in size, ranging from 13 to 601 objects. Many of the smaller data sets contained the entirety of the population of objects, for example, all 29 islands in the Galápagos archipelago. Most of the data sets were used in earlier studies (Czerlinski et al., 1999; Şimşek, 2013; and Şimşek and Buckmann, 2015). All are publicly available. The data sets are described in the supplementary material.

We tested the following models: TTB, TTBS, lexipost, and tally-the-best. In addition, we tested random forests (Breiman, 2001), one of the very best statistical learning algorithms, to provide a strong benchmark from machine learning. We trained random forests using their implementation in R package *randomForest* (Liaw and Wiener, 2002). Typically, the only parameter tuned when using random forests is *mtry*, which specifies how many cues should be randomly selected for consideration when splitting a branch (Hastie et al., 2009). We tuned *mtry* using 10-fold cross-validation in the training set. A description of our random-forest implementation is provided in the supplementary material.

We focused on cases where sample variance plays a role in the learning process, for example, due to differences in discrimination rates, resulting in differences in the number of informative samples available for different cues (even though all cues were trained on the same set of paired comparisons). We observed (not so small) differences in the learning curves of TTB and lexipost in 25 of the 56 data sets.

Figure 3 (top left) shows the mean accuracy in these 25 data sets as the training-set size grows, starting with one instance. We focus here on differences in early stages of learning. Because some of the data sets are smaller than others, the number of data sets included in the figure decreases as training-set sizes increases (therefore the tail end of the learning curves are not smooth). On the 25 data sets, there is a substantial gap between TTB and random forest. TTBS, lexipost, and tally-the-best close this gap to some extent.

In the heuristics literature, it is common to dichotomize the cues around the median (Czerlinski et al., 1999; Brighton, 2006; Martignon et al., 2008), for which one reason is "to mimic the limited knowledge about cue values that people typically have, and the potential unreliability of precise values" (Gigerenzer et al., 1999). With dichotomized cues, sampling variance almost always plays an important role in the learning process. In Figure 3 (top right), we show mean accuracy in all 56 data sets when the cues were dichotomized around the median. All three models performed better than or as well as random forests. TTB lagged behind in some regions of the learning curve.

The figure shows, in addition, individual learning curves in 9 of the 25 data sets. These plots show two standard errors around each learning curve as a shaded region surrounding the curve above and below. On individual data sets, lexipost frequently made large improvements in performance compared to TTB. Surprisingly, this was also true for the other two methods, tally-the-best and TTBS, despite their very simple handling of sampling uncertainty.

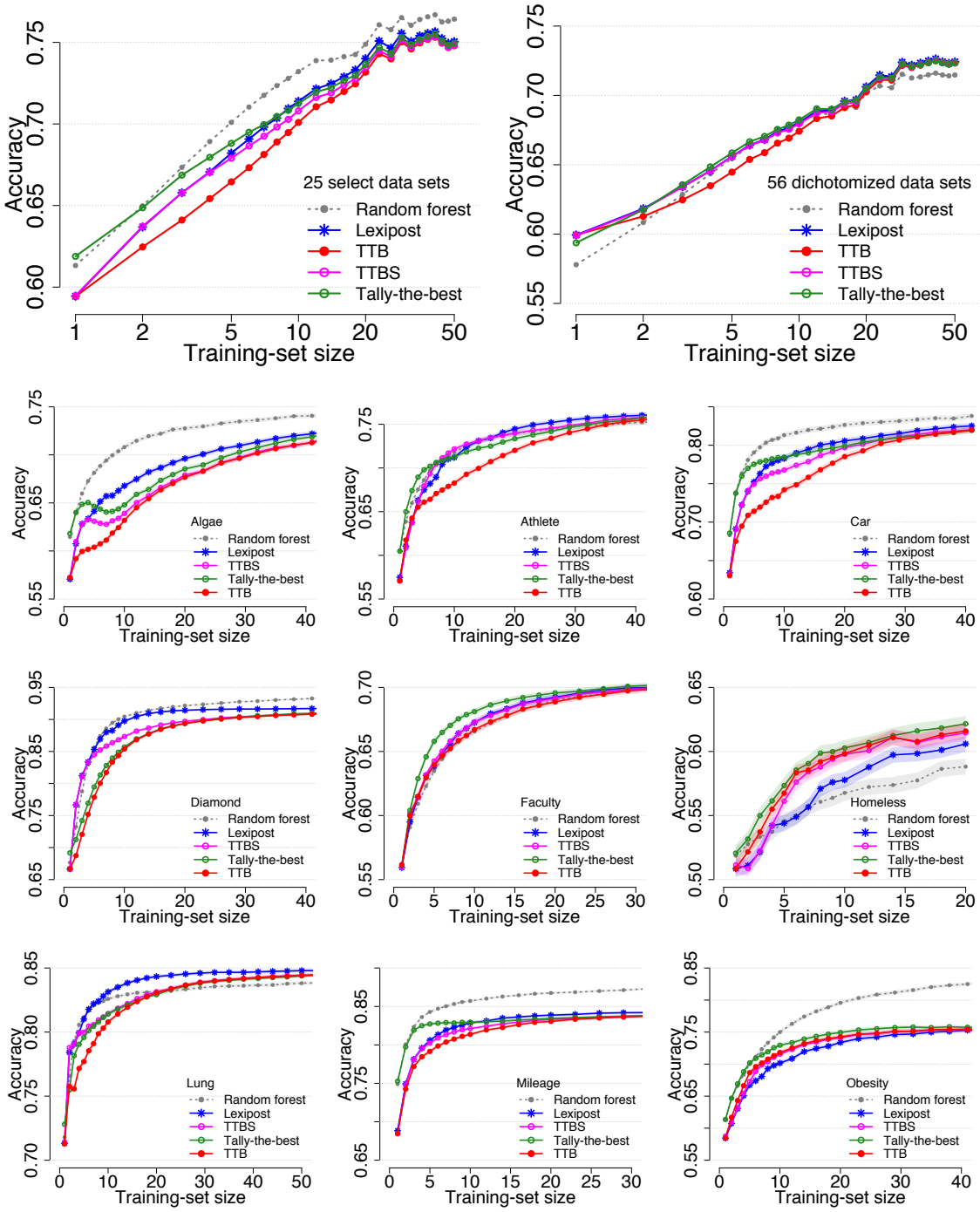

Figure 3: Top left: Mean accuracy in 25 data sets where there was a visible difference between the learning curves of take-the-best (TTB) and lexipost. Top right: Mean accuracy in all 56 data sets, when the cues are dichotomized around the median. In both plots, the horizontal axis is drawn on a log scale. Other plots: Learning curves in 9 of the 25 data sets (no dichotomization of the cues).

7. Discussion

Our results provide a foundation for taking into account sample variance in learning decision heuristics. The far superior performance of lexipost compared to TTB suggests that lexicographic heuristics have large untapped potential. In our simulations, even very simple ways of accounting for sampling uncertainty resulted in large performance improvements in many data sets. Principled methods of handling uncertainty have the potential to further improve performance. Our analysis may be useful in understanding how people take sampling uncertainty into account in learning simple decision rules.

References

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

Henry Brighton. Robust Inference with Simple Cognitive Models. In C. Lebiere and R. Wray, editors, *AAAI Spring Symposium: Cognitive Science Principles Meet AI Hard Problems*, pages 17–22, Menlo Park, CA, 2006. American Association for Artificial Intelligence.

Jean Czerlinski, Gerd Gigerenzer, and Daniel G. Goldstein. How good are simple heuristics? In Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, editors, *Simple heuristics that make us smart*, pages 97–118. Oxford University Press, New York, 1999.

Peter C. Fishburn. Lexicographic orders, utilities and decision rules: A survey. *Management Science*, 20(11):1442–1471, 1974.

Gerd Gigerenzer and Daniel G. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669, 1996.

Gerd Gigerenzer, Peter M. Todd, ABC Research Group, et al. *Simple heuristics that make us smart*. Oxford University Press, New York, 1999.

Gerd Gigerenzer, Ralph Hertwig, and Thorsten Pachur, editors. *Heuristics: The Foundations of Adaptive Behavior*. Oxford University Press, New York, 2011.

Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer New York Inc., New York, NY, USA, 2. edition, 2009.

Andy Liaw and Matthew Wiener. Classification and regression by randomforest. *R News*, 2 (3):18–22, 2002.

Laura Martignon, Konstantinos V. Katsikopoulos, and Jan K. Woike. Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology*, 52 (6):352–361, 2008.

Özgür Şimşek. Linear decision rule as aspiration for simple decision heuristics. In Chris J. C. Burges, Leon Bottou, Max Welling, Zoubin Ghahramani, and Kilian Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 2904–2912. Curran Associates, Inc., Red Hook, New York, 2013.

Özgür Şimşek and Marcus Buckmann. Learning from small samples: An analysis of simple decision heuristics. In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 3159–3167. Curran Associates, Inc., Red Hook, New York, 2015.

Proceedings of Machine Learning Research 58 (2016) 87-96

Quantum Rational Preferences and Desirability

Alessio Benavoli[*] ALESSIO@IDSIA.CH
Alessandro Facchini ALESSANDRO.FACCHINI@IDSIA.CH
Marco Zaffalon ZAFFALON@IDSIA.CH
Dalle Molle Institute for Artificial Intelligence (IDSIA), Switzerland

Editor: T.V. Guy, M. Kárný, D. Rios-Insua, D.H. Wolpert

Abstract

We develop a theory of quantum rational decision making in the tradition of Anscombe and Aumann's axiomatisation of preferences on horse lotteries. It is essentially the Bayesian decision theory generalised to the space of Hermitian matrices. Among other things, this leads us to give a representation theorem showing that quantum complete rational preferences are obtained by means of expected utility considerations.

Keywords: Quantum mechanics, Bayesian decision theory, Horse lotteries, Imprecise probability.

1. Introduction

The aim of this paper is simple. We have recently shown in Benavoli et al. (2016a) that Quantum Mechanics (QM) coincides with the Bayesian theory once this is formulated in the space of Hermitian matrices (so as to make possible to gamble on quantum experiments). Such an identification makes a number of things, at least in principle, straightforward: one of these is the extension of QM to make it handle non-linear utility. We do so by adapting the traditional axiomatisation of rational preferences by Anscombe and Aumann (1963) to the quantum setting.

After axiomatising quantum rational preferences in this way, we move on to give a representation theorem that shows that quantum rational preferences can be obtained by means of expected utility considerations. Our route to this results is based on the approach devised by Zaffalon and Miranda (2015) in the classical case: we show that the axiomatisation of quantum rational preferences is equivalent to the quantum theory of coherent desirable gambles—the very same theory at the heart of Benavoli et al.'s (2016a) formulation of QM—yet formulated so as to consider prizes other then events. Intuitively, this allows us to formally bring quantum rational preferences back to plain QM through an enlargement of the space of possibilities. We eventually show how this leads to quantum probabilities and utilities after enforcing axioms for state independence and completeness. All proofs can be found in the extended version (Benavoli et al., 2016c).

Before briefly reviewing related work and making some final comments, we illustrate one elegant consequence of the correspondence between preferences and desirability: namely, how to derive a coherent rule for updating preferences determining how should a subject rationally change her preferences in the prospect of obtaining new information in the form of an event.

Since the wording "quantum" is used nowadays with a number of acceptations in the literature, we would like to make precise what our framework actually addresses: that is, nothing else but

∗. The authors are listed in alphabetic order.

gambling on quantum mechanics experiments; we are not, at this stage, endorsing any other interpretation of the quantum decision theory developed here. Moreover, we would like to remark that our framework is an actual generalisation of classical decision-theoretic approaches in the tradition of Anscombe and Aumann (1963): in fact, we can recover them by simply focusing on the subset of Hermitian matrices made by diagonal real-valued matrices; that is, by focusing on classical experiments.

2. Rational gambling on quantum experiments

We start by defining a gambling system about the results of a quantum experiment. To this end, we consider two subjects: the bookmaker and the gambler (Alice). The bookmaker prepares the quantum system in some quantum state. Alice has her personal knowledge (beliefs) about the experiment—possibly no knowledge at all.

1. The bookmaker announces that he will measure the quantum system along its n orthogonal directions and so the outcome of the measurement is an element of $\Omega = \{\omega_1, \ldots, \omega_n\}$, with ω_i denoting the elementary event "detection along i". Mathematically, it means that the quantum system is measured along its eigenvectors,[1] i.e., the projectors[2] $\Pi^* = \{\Pi_1^*, \ldots, \Pi_n^*\}$ and ω_i is the event "indicated" by the i-th projector. The bookmaker is fair, meaning that he will correctly perform the experiment and report the actual results to Alice.
2. Before the experiment, Alice declares the set of gambles she is willing to accept. Mathematically, a gamble G on this experiment is a Hermitian matrix in $\mathbb{C}_h^{n\times n}$, the space of all Hermitian $n \times n$ matrices being denoted by $\mathbb{C}_h^{n\times n}$. We will denote the set of gambles Alice is willing to accept by $\mathscr{K} \subseteq \mathbb{C}_h^{n\times n}$.
3. By accepting a gamble G, Alice commits herself to receive $\gamma_i \in \mathbb{R}$ utiles if the outcome of the experiment eventually happens to be ω_i. The value γ_i is defined from G and Π^* through $\Pi_i^* G \Pi_i^* = \gamma_i \Pi_i^*$ for $i = 1, \ldots, n$. It is a real number since G is Hermitian.

Denote by $\mathscr{H}^+ = \{G \in \mathbb{C}_h^{n\times n} : G \gtrsim 0\}$ the subset of all positive semi-definite and non-zero (PSDNZ) matrices in $\mathbb{C}_h^{n\times n}$: we call them the set of *positive gambles*. The set of negative gambles is similarly given by $\mathscr{H}^- = \{G \in \mathbb{C}_h^{n\times n} : G \lesssim 0\}$. Alice examines the gambles in $\mathbb{C}_h^{n\times n}$ and comes up with the subset $\mathscr{K}$ of the gambles that she finds desirable. Alice's rationality is characterised as follows:

1. Any gamble $G \in \mathbb{C}_h^{n\times n}$ such that $G \gtrsim 0$ must be desirable for Alice, given that it may increase Alice's utiles without ever decreasing them (**accepting partial gain**): $\mathscr{H}^+ \subseteq \mathscr{K}$.
2. Any gamble $G \in \mathbb{C}_h^{n\times n}$ such that $G \lesssim 0$ is not desirable for Alice, given that it may only decrease Alice's utiles without ever increasing them (**avoiding partial loss**): $\mathscr{H}^- \cap \mathscr{K} = \emptyset$.
3. If Alice finds G desirable, that is $G \in \mathscr{K}$, then also νG must be desirable for her for any $0 < \nu \in \mathbb{R}$ (**positive homogeneity**).
4. If Alice finds G_1 and G_2 desirable, that is $G_1, G_2 \in \mathscr{K}$, then she must also accept $G_1 + G_2$, i.e., $G_1 + G_2 \in \mathscr{K}$ (**additivity**).

To understand these rationality criteria, originally presented in Benavoli et al. (2016a, Sec. III), we must remember that mathematically the payoff for any gamble G is computed as $\Pi_i^* G \Pi_i^*$ if the outcome of the experiment is the event indicated by Π_i^*. Then the first two rationality criteria above hold no matter the experiment Π^* that is eventually performed. In fact, from the properties of PSDNZ matrices, if $G \gtrsim 0$ then $\Pi_i^* G \Pi_i^* = \gamma_i \Pi_i^*$ with $\gamma_i^* \geq 0$ for any i and $\gamma_j > 0$ for some j.

1. We mean the eigenvectors of the density matrix of the quantum system.
2. A projector Π is a set of n positive semi-definite matrices in $\mathbb{C}_h^{n\times n}$ s.t. $\Pi_i \Pi_k = 0$, $(\Pi_i)^2 = \Pi_i = (\Pi_i)^\dagger$, $\sum_{i=1}^n \Pi_i = I$.

Therefore, by accepting $G \gtrsim 0$, Alice can only increase her utiles. Symmetrically, if $G \lesssim 0$ then $\Pi_i^* G \Pi_i^* = \gamma_i \Pi_i^*$ with $\gamma_i \leq 0$ for any i. Alice must then avoid the gambles $G \lesssim 0$ because they can only decrease her utiles. This justifies the first two rationality criteria. For the last two, observe that

$$\Pi_i^*(G_1+G_2)\Pi_i^* = \Pi_i^* G_1 \Pi_i^* + \Pi_i^* G_2 \Pi_i^* = (\gamma_i + \vartheta_i)\Pi_i^*,$$

where we have exploited the fact that $\Pi_i^* G_1 \Pi_i^* = \gamma_i \Pi_i^*$ and $\Pi_i^* G_2 \Pi_i^* = \vartheta_i \Pi_i^*$. Hence, if Alice accepts G_1, G_2, she must also accept $G_1 + G_2$ because this will lead to a reward of $\gamma_i + \vartheta_i$. Similarly, if G is desirable for Alice, then also $\Pi_i^*(\nu G)\Pi_i^* = \nu \Pi_i^* G \Pi_i^*$ should be desirable for any $\nu > 0$.

In other words, as in the case of classical desirability (Williams, 1975; Walley, 1991), the four conditions above state only minimal requirements: that Alice would like to increase her wealth and not decrease it (conditions 1 and 2); and that Alice's utility scale is linear (conditions 3 and 4). The first two conditions should be plainly uncontroversial. The linearity of the utility scale is routinely assumed in the theories of personal, and in particular Bayesian, probability as a way to isolate considerations of uncertainty from those of value (wealth).

We can characterise $\mathscr{K}$ also from a geometric point of view. In fact, from the above properties, it follows that a coherent set of desirable gambles $\mathscr{K}$ is a convex cone in $\mathbb{C}_h^{n\times n}$ that includes all positive gambles (accepting partial gains) and excludes all negative gambles (avoiding partial losses). Without loss of generality we can also assume that $\mathscr{K}$ is not pointed, i.e., $0 \notin \mathscr{K}$: Alice does not accept the null gamble. A coherent set of desirable gambles is therefore a non-pointed convex cone.

Definition 1 *We say that $\mathscr{K} \in \mathbb{C}_h^{n\times n}$ is a* **coherent quantum set of desirable gambles** *(DG) if*
(S1) *$\mathscr{K}$ is a non-pointed convex-cone (positive homogeneity and additivity);*
(S2) *if $G \gtrsim 0$ then $G \in \mathscr{K}$ (accepting partial gain).*

If in addition a coherent set of desirable gambles satisfies the following property:
(S3) *if $G \in \mathscr{K}$ then either $G \gtrsim 0$ or $G - \varepsilon I \in \mathscr{K}$ for some strictly positive real number ε (openness),*[3]

then it is said to be a coherent quantum set of **strictly** *desirable gambles (SDGs).*

Note that the although the additional openness property (S3) is not necessary for rationality, it is technically convenient as it precisely isolates the kind of models we use in QM (as well as in classical probability), see Benavoli et al. (2016a). Property (S3) has a gambling interpretation too: it means that we will only consider gambles that are *strictly* desirable for Alice; these are the positive ones or those for which she is willing to pay a positive amount to have them. Note that assumptions (S1) and (S2) imply that SDGs also avoids partial loss: if $G \lesssim 0$, then $G \notin \mathscr{K}$ (Benavoli et al., 2016a, Remark III.2).

In Benavoli et al. (2016a), we have shown that maximal set of SDGs, that is SDG sets that are not included in any larger set of SDG, and density matrices are one-to-one. The mapping between them is obtained through the standard inner product in $\mathbb{C}_h^{n\times n}$, i.e., $G \cdot R = Tr(G^\dagger R)$ with $G, R \in \mathbb{C}_h^{n\times n}$. This follows by a representation result whose proof is a direct application of Hahn-Banach theorem:

Theorem 2 (Representation theorem from Benavoli et al. (2016a)) *The map that associates a maximal SDG the unique density matrix ρ such that $Tr(G^\dagger \rho) \geq 0 \ \forall G \in \mathscr{K}$ defines a bijective corre-*

3. In Benavoli et al. (2016a) we used another formulation of openness, namely (S3'): if $G \in \mathscr{K}$ then either $G \gtrsim 0$ or $G - \Delta \in \mathscr{K}$ for some $0 < \Delta \in \mathbb{C}_h^{n\times n}$. (S3) and (S3') are provably equivalent given (S1) and (S2).

spondence between maximal SDGs and density matrices. Its inverse is the map $(\cdot)^\circ$ *that associates each density matrix* ρ *the maximal SDG*[4] $(\rho)^\circ = \{G \in \mathbb{C}_h^{n\times n} \mid G \gtrsim 0\} \cup \{G \in \mathbb{C}_h^{n\times n} \mid Tr(G^\dagger \rho) > 0\}$.

This representation theorem has several consequences. First, it provides a gambling interpretation of the first axiom of QM on density operators. Second, it shows that density operators are coherent, since the set $(\rho)^\circ$ that they induce is a valid SDG. This also implies that QM is self-consistent—a gambler that uses QM to place bets on a quantum experiment cannot be made a partial (and, thus, sure) loser. Third, the first axiom of QM on $\mathbb{C}_h^{n\times n}$ is structurally and formally equivalent to Kolmogorov's first and second axioms about probabilities on $\mathbb{R}^n$ (Benavoli et al., 2016a, Sec. 2). In fact, they can be both derived via duality from a coherent set of desirable gambles on $\mathbb{C}_h^{n\times n}$ and, respectively, $\mathbb{R}^n$. In Benavoli et al. (2016a) we have also derived Born's rule and the other three axioms of QM from the illustrated setting and shown that measurement, partial tracing and tensor product are just generalised probabilistic notions of Bayes rule, marginalisation and independence. Finally, as shown in (Benavoli et al., 2016b), the representation theorem enables us to derive a Gleason-type theorem that holds for any dimension n of a quantum system, hence even for $n = 2$.

3. Quantum horse lotteries and preference relations

As seen in Sec. 2, QM is just the Bayesian theory of probability formulated over Hermitian matrices. Now we proceed to extend such a theory of probability to make it handle non-linear utility. To this end, we work in the tradition of Anscombe and Aumann (1963). Central to this tradition is the notion of a horse lottery.

Consider a set of prizes $X = \{x_1, \ldots, x_m\}$ with $m \geq 2$ (this last constraint will be clarified later). A *horse lottery* is a compound lottery such that if $\omega \in \Omega$ occurs, it returns a simple lottery, which can depend on ω, over the prizes in X. The idea is that at some later point the subject (Alice) will play the simple lottery thus earning one of the prizes. Anscombe and Aumann's setting axiomatises rational preferences over horse lotteries; from this, it follows that there are probabilities and utilities that represent those preferences via maximum expected utility.

3.1 Horse lotteries over complex numbers

As in the classical case, now we consider that when Π_i^* is observed, Alice receives a probability mass function p_i (pmf) over the prices X rather than the value γ_i as in Sec. 2. This framework is a composite system made of a quantum experiment and a classical experiment (on X). To describe it in a mathematically convenient way, we need to define gambles on this composite system. First, we define the form of the gambles. Since the experiment on X is classical, it can be described by the subspace of $\mathbb{C}_h^{m\times m}$ of diagonal matrices; we denote it as $\mathbb{D}^m$. Gambles on this composite system are therefore elements of $\mathbb{D}^m \otimes \mathbb{C}_h^{n\times n} \subset \mathbb{C}_h^{mn\times mn}$, where $\otimes$ denotes the tensor product. It can be observed that G is a block diagonal matrix with elements in $\mathbb{C}_h^{n\times n}$, i.e. $G = \mathrm{Diag}(G_1, \ldots, G_m)$ with $G_k \in \mathbb{C}_h^{n\times n}$.

We are interested in the special case of gambles on $\mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$ that return a pmf p_i when the quantum system is measured along some projector.

Definition 3 *Let* $Q \in \mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$. *$Q$ is said to be a* quantum horse lottery *(QH-lottery) if*

$$\forall \Pi, \forall \Pi_i \in \Pi, \exists p_i \in \mathbb{P}^m : (I_m \otimes \Pi_i) Q (I_m \otimes \Pi_i) = p_i \otimes \Pi_i, \tag{1}$$

4. Here the gambles $G \gtrsim 0$ are treated separately because they are always desirable and, thus, they are not informative on Alice's beliefs about the quantum system. Alice's knowledge is determined by the gambles that are not $G \gtrsim 0$.

where $\mathbb{P}^m \subset \mathbb{D}^m$ denotes the subset of trace one diagonal-matrices whose elements are non-negative.

The set $\mathbb{P}^m$ is isomorphic to the set of all probability mass functions (pmf) on $\mathbb{R}^m$. Therefore, with an abuse of terminology we improperly refer to the diagonal matrix p_i as a pmf. We denote the subspace of $\mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$ of QH-lotteries as QL. By Definition 3, a QH-lottery is therefore a gamble that returns to Alice the pmf p_i on X whenever a measurement Π is performed on the quantum system and the projector $\Pi_i \in \Pi$ is observed. In what follows, we determine some properties of QL.

Consider the matrix Q in $\mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$ defined as $Q = \sum_{j=1}^n q_j \otimes V_j$, with $q_j \in \mathbb{P}^m$ and $V = \{V_j\}_{j=1}^n$ is an orthogonal decomposition (OD) on $\mathbb{C}_h^{n\times n}$. It turns out that actually Q is a QH-lottery.We call it a *simple* QH-lottery. Note that a convex combination of simple QH-lotteries is a QH-lottery. However, such a combination need not be simple anymore. The next theorem isolates necessary and sufficient conditions for an element of the composite space $\mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$ to be a QH-lottery.

Theorem 4 *Let $Q \in \mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$ of the form $Q = Diag(Q_1, \ldots, Q_m)$. Then $Q \in \mathsf{QL}$ if and only if $Q_j \geq 0$ for every $j = 1, \ldots, m$ and $\sum_{j=1}^m Q_j = I_n$.*

Remark 5 *It should be observed that $Q_1, \ldots, Q_m$ are Hermitian positive semi-definite matrices that sum up to the identity operator. This is the definition of positive-operator valued measure (POVM). Therefore the generalisation of horse lotteries to the quantum setting naturally leads to POVMs.*

Remark 6 *The classical definition of horse lotteries can be recovered by the quantum one just by considering, instead of the space $\mathbb{C}_h^{n\times n}$, the space of diagonal real-valued matrices. Obviously, the composite system under consideration is $\mathbb{D}^m \times \mathbb{D}^n$. This space is isomorphic to the space $\mathscr{L}(X \times \Omega)$ of real valued functions whose domain is $X \times \Omega$, where $\Omega = \{1, \ldots, n\}$ and $X = \{1, \ldots, m\}$. By applying Definition 3, we immediately obtain that an object $Q \in \mathscr{L}(X \times \Omega)$ satisfies Property 1 if and only if $Q(\cdot, \omega)$ is a pmf on X for each $\omega \in \Omega$, meaning that Q is a (classical) horse lottery.*

3.2 Coherent preference relations

Horse lotteries are given a behavioural interpretation through a notion of preference. The idea is that Alice, who aims at receiving a prize from X, will prefer some horse lotteries over some others, depending on her knowledge about the quantum experiment, as well as on her attitude towards the prizes. We consider the following well-known axioms of rational preferences, formulated here in the quantum setting.

Definition 7 *A preference relation over quantum horse lotteries is a subset $\succ \subseteq \mathsf{QL} \times \mathsf{QL}$. It is said to be* **coherent** *if it satisfies the following axioms:*

(A.1) $(\forall P, Q, R \in \mathsf{QL}) P \not\succ P$ *and* $P \succ Q \succ R \Rightarrow P \succ R$ [strict partial order]*;*

(A.2) $(\forall P, Q \in \mathsf{QL}) P \succ Q \Leftrightarrow (\forall R \in \mathsf{QL})(\forall \alpha \in (0,1])\, \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)R$ [mixture independence].

Our approach is therefore a straightforward generalisation of the classical setting to the quantum case. In the classical axiomatisations of rational preferences, it is customary to assume that the preference relation has a best and a worst horse lottery. For us it is enough to assume that the worst

one exists and that it actually corresponds to a worst element in X.[5] Formally, we denote the last (m-th) element of X as z. By $p^z \in \mathbb{P}^m$ we denote the pmf that assigns all the mass to z. Finally, by Z we denote the QH-lottery $p_z \otimes I_n$. Notice that $Z = \sum_{i=n}^{n} p_z \otimes \Pi_i$, for every OD $\Pi = \{\Pi_i\}_{i=1}^n$, and therefore $(I_m \otimes \Pi_i)Z(I_m \otimes \Pi_i) = p_z \otimes \Pi_i$, for every $\Pi_i \in \Pi$.

Definition 8 *Let $\succ \subset \mathsf{QL} \times \mathsf{QL}$ be a preference relation. We say that $\succ$ has the worst outcome if there is $z \in X$ such that $P \succ Z$ for every $P \neq Z$.*

In what follows we assume that preference relations have such a worst outcome. The rationale is that the elements of $X \backslash \{z\}$ are actual prizes, whereas z represents the event that no prize in $X \backslash \{z\}$ is won (nothing is won). We have assumed that $m \geq 2$ precisely as a consequence of the assumption that there is the worst-outcome z among the elements of X.

The scaled differences of QL-lotteries is the set defined by

$$\mathscr{A} = \{\lambda(P-Q) \mid \lambda > 0, P, Q \in \mathsf{QL}\}, \tag{2}$$

where λ is a positive real. The set $\mathscr{A}$ constitutes a vector space.

Theorem 9 *The map $\succ \mapsto \mathscr{C} = \{\lambda(P-Q) \mid P, Q \in \mathsf{QL},\ \lambda > 0, P \succ Q\}$ determines a bijection between coherent preference relations over QL and non-pointed convex cones in $\mathscr{A}$.*

Thus, it turns out that non-pointed cones and coherent quantum preference relations are just two ways of looking at the same thing.

4. Quantum desirability vs quantum preference: two faces of the same coin

In this section we follow the same strategy as in Zaffalon and Miranda (2015) to establish an equivalence between the theories of coherent quantum preference and coherent quantum desirability. To this end, we first define the projection operator that drops the z-components from an act.

Definition 10 *The* projection operator *is the functional $\mathsf{Proj} : \mathbb{D}^m \otimes \mathbb{C}_h^{n\times n} \to \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ that takes the QL-lottery (m-block diagonal matrix) Q and returns $\mathsf{Proj}(Q) = Diag(Q_1, \ldots, Q_{m-1})$.*

In this paper, we are going to use this operator to project QH-lotteries in $\mathbb{D}^m \otimes \mathbb{C}_h^{n\times n}$into gambles on $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$. However, instead of working directly with the space QL, in what follows it will be more convenient to reason on the space $\mathscr{A}$ of scaled differences of QH-lotteries defined in (2). Note also that the restriction of Proj to $\mathscr{A}$is injective.

Based on the correspondence between cones on $\mathscr{A}$ and preference relations, it is then an easy step to show (see (Benavoli et al., 2016c, Proposition 22)) that given a coherent preference relation $\succ$, one can define a coherent set $\mathscr{K}$ of desirable gambles on $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ as $\mathscr{K} = \{\lambda\, \mathsf{Proj}(P-Q) : P \succ Q, \lambda > 0\}$ and with the property that

$$P \succ Q \Leftrightarrow \mathsf{Proj}(P-Q) \in \mathscr{K}. \tag{3}$$

One can actually verify that there is an exact correspondence between coherent sets of desirable gambles and coherent preference relations.

Theorem 11 *There is a one-to-one correspondence between coherent sets of desirable gambles over $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ and coherent preference relations over $\mathsf{QL} \times \mathsf{QL}$.*

5. The two requirements—having a worst horse lottery and a worst element in X—have been shown equivalent in Zaffalon and Miranda (2015, Proposition 8).

5. Archimedeanity and the representation theorem

Archimedeanity is an extra axiom adopted in traditional axiomatisations of rational preferences that tries to capture a form of continuity; it is such an axiom that makes it possible to have a representation of preferences in terms of expected utilities. Zaffalon and Miranda (2015, Prop. 11) (and in the quantum case, we, in (Benavoli et al., 2016c, Proposition 22)) have shown that the traditional Archimedean axiom has some drawbacks that can be fixed with a slight change in its definition. It is based on the notion of objective preference.

Definition 12 (Objective preference) *Let $P,Q \in \mathsf{QL}$. We say that P* is objectively preferred to *Q if* $\mathsf{Proj}(P-Q) \gtrsim 0$. *We denote objective preference by $P \rhd Q$.*

(Note that the definition neglects the outcome z, since it is not one any subject actually wants.)

Objective preference is a preference relation. Moreover, it is the least preference relation over $\mathsf{QL} \times \mathsf{QL}$ in the sense that it is included in any other preference relation (in this sense, we call it "objective"). Now we can directly rephrase Zaffalon and Miranda's Archimedean notion as follows for the quantum case:

(A.3) $(\forall P,Q \in \mathsf{QL})\ P \succ Q, P \not\rhd Q \Rightarrow (\exists \alpha \in (0,1))\ \alpha P + (1-\alpha)Z \succ Q$ [*Weak Archimedeanity*].

Analogously to their case, we obtain that it is equivalent to use coherent quantum sets of strictly desirable gambles in order to represent weakly Archimedean coherent preference relation on quantum horse lotteries. Recall also that a preference relation $\succ$ is said to be complete (or total) if either $P \succ Q$ or $Q \succ P$, for every $P,Q \in \mathsf{QL}$ with $P \neq Q$.

Theorem 13 *There is a one to one correspondence between coherent sets of SDG over $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ and coherent preference relations over $\mathsf{QL} \times \mathsf{QL}$ that are weakly Archimedean. Moreover, such a correspondence induces a bijection between maximal coherent sets of SDG and complete weakly Archimedean coherent preference relations.*

Based on Theorem 13, we can then obtain a representation theorem for complete weakly Archimedean coherent preference relations as follows. First of all, Theorem 2 from Benavoli et al. (2016a) can be restated in the case of quantum horse lotteries as follows.

Theorem 14 *The map that associates a maximal SDG over $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ the unique trace-one positive matrix $R \in \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ such that $Tr(G^\dagger R) \geq 0\ \forall G \in \mathscr{K}$ defines a bijective correspondence between maximal SDGs over $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ and trace-one positive matrices over $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$. Its inverse is the map $(\cdot)^\circ$ that associates each trace-one positive matrix R the maximal SDG*

$$(R)^\circ = \{\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n} \mid G \gtrsim 0\} \cup \{G \in \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n} \mid Tr(G^\dagger R) > 0\}. \tag{4}$$

All trace-one positive matrices R are of the form $R = \mathrm{Diag}\,(p_1\rho_1, \ldots, p_{m-1}\rho_{m-1})$ with $\rho_i \in \mathbb{C}_h^{n\times n}$ being density matrices and $\mathrm{Diag}(p_1, \ldots, p_{m-1}) \in \mathbb{P}^{m-1}$. Hence, applying Theorems 13 and 14 to Property 3 yields the following representation result for complete preference relations:

Corollary 15 *A relation $\succ$ over $\mathsf{QL} \times \mathsf{QL}$ is a complete weakly Archimedean coherent preference relation if and only if there is a unique trace-one positive matrix $R = Diag\,(p_1\rho_1, \ldots, p_{m-1}\rho_{m-1})$ such that*

$$P \succ Q \Leftrightarrow \Big(either\ P \rhd Q\ or\ \sum_{i=1}^{m-1} p_i Tr(P_i^\dagger \rho_i) > \sum_{i=1}^{m-1} p_i Tr(Q_i^\dagger \rho_i)\Big)\ \forall P,Q \in \mathsf{QL}. \tag{5}$$

Consistently with our generalisation of Gleason's theorem (Benavoli et al., 2016b), this result holds in any dimension (even $n = 2$), because we ask preference relations to be coherent.

6. Coherent updating and state independence

This section shows how to derive in a very simple, elegant, way a coherent rule for updating preferences. In particular our aim is to answer this question: how should Alice change her preferences in the prospect of obtaining new information in the form of an event?

We initially assume that Alice considers an event "indicated" by a certain projector $\Pi_i \in \mathbb{C}_h^{n\times n}$ in $\Pi = \{\Pi_i\}_{i=1}^n$. The information it represents is: an experiment Π is performed and the event indicated by Π_i happens.[6]

Now, assume that Alice's preferences are modelled by the coherent relation $\succ$ on QL. From Theorem 11 we can consider the coherent set $\mathscr{K}$ in $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$. Hence, we reason as in the derivation of the second axiom of QM in Benavoli et al. (2016a, Sec. V). Under the assumption that an experiment Π is performed and the event indicated by Π_i happens, Alice can focus on gambles that are contingent on $I_{m-1} \otimes \Pi_i$: these are the gambles such that "outside" $I_{m-1} \otimes \Pi_i$ no utile is received or due—status quo is maintained—; in other words, they represent gambles that are called off if the outcome of the experiment is not Π_i. Mathematically, these gambles are of the form

$$G = \begin{cases} H & \text{if } I_{m-1} \otimes \Pi_i \text{ occurs,} \\ 0 & \text{if } I_{m-1} \otimes \Pi_j \text{ occurs, with } j \neq i. \end{cases}$$

It is then clear that $H = \alpha I_{m-1} \otimes \Pi_i$ with $\alpha \in \mathbb{R}$ since $\Pi_i \Pi_j = 0$ for each $j \neq i$. In this light, we can define Alice's conditional preferences by moving to the equivalent view on gambles, restricting the attention to gambles of the form $I_{m-1} \otimes \Pi_i G I_{m-1} \otimes \Pi_i = \alpha I_{m-1} \otimes \Pi_i$ with $G \in \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$, and finally updating the preferences by looking at the corresponding preference relation.

Definition 16 *Let $\succ$ be a preference relation. The relation obtained as $\succ_{\Pi_i} := \mathsf{Proj}_1^{-1}(\mathscr{K}_{\Pi_i})$, with*

$$\mathscr{K}_{\Pi_i} = \left\{ G \in \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n} \mid G \gtrsim 0 \text{ or } (I_{m-1} \otimes \Pi_i) G (I_{m-1} \otimes \Pi_i) \in \mathscr{K} \right\} \tag{6}$$

is called the **preference relation conditional** *on Π_i.*

By the same argument as in Benavoli et al. (2016a, Prop. A.6), one can prove that $\mathscr{K}_{\Pi_i}$ is a coherent set of (strictly) desirable gambles, whenever $\mathscr{K}$ is a coherent set of (strictly) desirable gambles. By Theorems 11 and 13 this yields that:

Theorem 17 *Let $\succ$ be a (weakly Archimedean) coherent preference relation. The relation $\succ_{\Pi_i}$ conditional on the event Π_i is also a (weakly Archimedean) coherent preference relation.*

Now, we rely on conditioning to introduce the concept of state-independent preferences. For this purpose, we use results in Benavoli et al. (2016a) to prove the fourth postulate of QM about composite systems. We first define the concept of epistemic irrelevance.

6. We assume that the quantum measurement device is a "perfect meter" (an ideal common assumption in QM), i.e., there are not observational errors—Alice can trust the received information.

Definition 18 *Let $\mathscr{R} \subset \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$, and let us define*

$$\begin{aligned} \mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{R}) &= \left\{G \in \mathbb{D}^{m-1} \mid G \otimes I_n \in \mathscr{R}\right\}, \\ \mathscr{R}_{|\mathbb{C}_h^{n\times n}} &= \left\{H \in \mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n} \mid H \gtrsim 0 \text{ or } (I_{m-1} \otimes \Pi_i)H(I_{m-1} \otimes \Pi_i) \in \mathscr{R}\right\}, \end{aligned}$$

An SDG $\mathscr{K}$ on $\mathbb{D}^{m-1} \otimes \mathbb{C}_h^{n\times n}$ is said to satisfy **epistemic irrelevance** *of the subsystems $\mathbb{D}^{m-1}$ to $\mathbb{C}_h^{n\times n}$ when $\mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{K}) = \mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{K}_{|\mathbb{C}_h^{n\times n}})$ for each projection measurement $\Pi = \{\Pi_i\}_{i=1}^n$.*

Let us briefly explain this definition. $\mathscr{K}$ is the SDG conditional on the event indicated by $I_{m-1} \otimes \Pi_i$, as it follows from its definition and (6). Thus, $\mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{K}) = \mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{K}_{\mathbb{C}_h^{n\times n}})$ means that Alice's marginal SDG $\mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{K})$ on the subsystem $\mathbb{D}^{m-1}$ and the marginal on $\mathbb{D}^{m-1}$ of Alice's SDG updated with the information "the event indicated by Π_i has happened", which is $\mathsf{marg}_{\mathbb{D}^{m-1}}(\mathscr{K}_{\mathbb{C}_h^{n\times n}})$, coincide. If this holds for all possible Π_i's, then any information on $\mathbb{C}_h^{n\times n}$ does not change Alice's beliefs on $\mathbb{D}^{m-1}$: this is precisely the definition of epistemic irrelevance. In case $\mathscr{K}$ is maximal and satisfies epistemic irrelevance, we have shown in Benavoli et al. (2016a, Sec. VII.c) that the representation Theorem 14 applied to such $\mathscr{K}$ defines a matrix R that factorizes as $R = p \otimes \rho$. Therefore, as in the classical framework for decision theory, the "joint" density matrix R factorizes as the product of $p \in \mathbb{P}^m$ and the density matrix $\rho \in \mathbb{C}_h^{n\times n}$. Stated otherwise, $\mathscr{K}$ models independence between utility (p) and the "probabilistic" information (ρ) associated to the quantum system. Alice's preferences are state-independent in this case.

7. Related work

Axiomatic frameworks for the theory of subjective expected utility were originally given by Savage (1954) and by Anscombe and Aumann (1963). Karni (2013) provides a recent overview of several variations and extensions of these two models.

Busemeyer and Bruza's (2012) presents an overview and many references to quantum-like approaches to cognition and decision theory. Deutsch (1999), Khrennikov (2016) and Danilov et al. (2016) are examples of other works addressing similar issues. In particular the latter proposes an axiomatisation for quantum preferences directly in the space of Hermitian matrices similar to the one presented here. However, the authors only consider what we call *simple* quantum horse lotteries. In doing so, the traditional axiom of mixture independence is formulated relative to the particular orthogonal decomposition associated to a simple lottery; an additional axiom becomes then necessary to bind lotteries based on different orthogonal decompositions. Moreover their representation theorem crucially employs Gleason's theorem and therefore only works on spaces of dimension greater than two. Because of those characteristics, it is unclear to us whether or not that axiomatisation is coherent: e.g., whether it guarantees that a subject whose quantum preferences on a space of dimension two cannot be made a sure loser, that is, shown to be irrational. The case of dimension two is particularly critical as dispersion-free probabilities—which Benavoli et al. (2016b) have shown to be incoherent—could in principle be employed.

8. Concluding remarks

In this paper, we have axiomatised rational preferences over quantum horse lotteries. Such a development is a natural follow up of our recent work (Benavoli et al., 2016a), which has shown that Quantum Mechanics is the Bayesian theory of probability over Hermitian matrices. By bridging

those rational preferences with quantum desirability, we have given a representation theorem in terms of quantum probabilities and utilities.

There are many directions that can be explored starting from those in this paper. Two of them are particularly important in our view. The first regards the full extension of our setting to partial (i.e., incomplete) preferences; this would enable it to deal with sets of quantum probabilities and utilities. The second is the definition of horse lotteries in their full generality as compound quantum lotteries.

References

F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, 34:199–2005, 1963.

A. Benavoli, A. Facchini, and M. Zaffalon. Quantum mechanics: The Bayesian theory generalized to the space of Hermitian matrices. *Physical Review A*, 94:042106, 2016a.

A. Benavoli, A. Facchini, and M. Zaffalon. A Gleason-type theorem for any dimension based on a gambling formulation of Quantum Mechanics. *ArXiv e-prints 1606.03615*, 2016b.

Alessio Benavoli, Alessandro Facchini, and Marco Zaffalon. Quantum rational preferences and desirability. *arXiv preprint arXiv:1610.06764*, 2016c.

J. R. Busemeyer and P. D. Bruza. *Quantum models of cognition and decision*. Cambridge University Press, 2012.

V. I. Danilov, A. Lambert-Mogiliansky, and V. Vergopoulos. Dynamic consistency of expected utility under non-classical(quantum) uncertainty. Preprint, May 2016. URL `https://halshs.archives-ouvertes.fr/halshs-01324046`.

D. Deutsch. Quantum theory of probability and decisions. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 455, pages 3129–3137. The Royal Society, 1999.

Edi Karni. Axiomatic foundations of expected utility and subjective probability. *Handbook of Economics of Risk and Uncertainty*, 1, 2013.

A. Khrennikov. Quantum Bayesianism as the basis of general theory of decision-making. *Philosophical Transactions of the Royal Society A*, 374(20150245), 2016.

Leonard J Savage. *The foundations of statistics*. John Wiley and Sons, New York, 1954.

P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, New York, 1991.

P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Reprinted in Williams (2007).

P. M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of Williams (1975).

M. Zaffalon and E. Miranda. Desirability and the birth of incomplete preferences. *arXiv preprint arXiv:1506.00529*, 2015.

Proceedings of Machine Learning Research 58 (2016) 97-109

Rational Beliefs Real Agents Can Have – A Logical Point of View

Marcello D'Agostino MARCELLO.DAGOSTINO@UNIMI.IT
Department of Philosophy
University of Milan
20122 Milano, Italy

Tommaso Flaminio TOMMASO.FLAMINIO@UNINSUBRIA.IT
Department of Pure and Applied Sciences
University of Insubria
21100 Varese, Italy

Hykel Hosni HYKEL.HOSNI@UNIMI.IT
Department of Philosophy
University of Milan
20122 Milano, Italy.

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

The purpose of this note is to outline a framework for uncertain reasoning which drops unrealistic assumptions about the agents' inferential capabilities. To do so, we envisage a pivotal role for the recent research programme of *depth-bounded Boolean logics* (D'Agostino et al., 2013). We suggest that this can be fruitfully extended to the representation of rational belief under uncertainty. By doing this we lay the foundations for a prescriptive account of rational belief, namely one that *realistic agents*, as opposed to idealised ones, can feasibly act upon.

Keywords: Prescriptive rationality, tractability, logic-based probability, Bayesian norms

1. Introduction and motivation

Probability is traditionally the tool of choice for the quantification of uncertainty. Since Jacob Bernoulli's 1713 *Ars Conjectandi*, a number of arguments have been put forward to the effect that departing from a probabilistic assessment of uncertainty leads to *irrational* patterns of behaviour. This contributed to linking tightly the rules of probability with the defining norms of rationality, as fixed by the well known results of de Finetti (1974); Savage (1972). Lindley (2006) and Parmigiani and Inoue (2009) provide recent introductory reviews.

Over the past few decades, however, a number of concerns have been raised against the adequacy of probability as a norm of rational reasoning and decision making. Following the lead of Ellsberg (1961), whom in turn found himself on the footsteps of Knight (1921) and Keynes (1921), many decision theorists take issue with the idea that probability provides adequate norms for rationality. This is put emphatically in the title of Gilboa et al. (2012), a paper which has circulated for almost a decade before its publication. As a result,

considerable formal and conceptual effort has gone into extending the scope of the probabilistic representation of uncertainty, as illustrated for instance by Gilboa and Marinacci (2013). Related to this, is the large family of *imprecise probability* models, and its decision-theoretic offsprings, which constitute the cutting edge of uncertain reasoning research, see e.g. Augustin et al. (2014).

One key commonality between "non Bayesian" decision theory and the imprecise probabilities approach is the fact they take issue with the identification of "rationality" and "probability" on representational grounds. For they insist on the counterintuitive consequences of assuming that the rational representation of uncertainty necessitates the Bayesian norms, and in particular that all uncertainty is to be represented probabilistically.

This note makes a case for adding a *logical* dimension to this ongoing debate. Key to this is a logical framing of probability. As recalled explicitly below, probability functions are normalised on classical tautologies. That is to say that a Bayesian agent is required to assign maximum degree of belief to *every* tautology of the propositional calculus. However classic results in computational complexity imply that the problem of deciding whether a given sentence is a tautology, exceeds, in general, what is considered to be *feasible*. Hence, probability imposes a norm of rationality which, under widely agreed hypotheses, realistic agents cannot be expected to meet. A related concern had already been put forward by Savage (1967), but this did't lead proponents of the Bayesian approach to take the issue seriously. This is precisely what the research outlined in this note aims to do.

By framing the question logically, we can offer a perspective on the problem which highlights the role of *classical* logic in determining the unwelcome features of canonical Bayesian rationality (Section 3). This suggests that a normatively reasonable account of rationality should to take a step back and rethink the logic in the first place.

The recently developed framework of Depth-Bounded Boolean logics (DBBLs) is particularly promising in this respect. By re-defining the meaning of logical connectives in terms of *information actually possessed by the agent* DBBLs give rise to a hierarchy of logics which (i) accounts for some key aspects of the asymmetry between knowledge and ignorance and (ii) provide computationally feasible approximations to classical logic. Section 4.2 reviews informally the core elements of this family of logics.

Finally, Section 5 outlines the applicability of this framework to probabilistic reasoning. In particular it points out how the hierarchy of DBBLs to can serve to define a hierarchy of prescriptively rational approximation of Bayesian rationality.

2. Bayesian rationality

In a number of areas, from Economics to the Psychology of reasoning and of course Statistics, probability has been defended as the norm of rational belief. Formally this can be seen to imply a normative role also for classical logic. So the Bayesian norms of rationality are best viewed as combination of probability and logic.

This allows us to distinguish two lines of criticisms against Bayesian rationality. First, it is often pointed out that probability washes out a natural asymmetry between knowledge and ignorance. Second, the intractability of classical logical reasoning is often suggested to deprive the normative theory of practical meaning. Both lines of criticisms can be naturally linked to the properties of *classical* logic.

2.1 Against the probability norm: the argument from information

Uncertainty has to do, of course, with not knowing, and in particular not knowing the outcome of an event of interest, or the value of a random variable. Ignorance has more subtle features, and is often thought of as our inability to quantify our own uncertainty. Knight (1921) gave this impalpable distinction an operational meaning in actuarial terms. He suggested the presence of ignorance is detected by the absence of a compete insurance market for the goods at hand. On the contrary, a complete insurance market provides an operational definition of *probabilistically quantifiable* uncertainty. Contemporary followers of Knight insist that not all uncertainty is probabilistically quantifiable and seek to introduce more general norms of rational belief and decision under "Knightian uncertainty" or "ambiguity".

A rather general form of the argument from information against Bayesian rationality is summarised by the following observation by Schmeidler (1989):

> The probability attached to an uncertain event does not reflect the heuristic amount of information that led to the assignment of that probability. For example, when the information on the occurrence of two events is symmetric they are assigned equal probabilities. If the events are complementary the probabilities will be 1/2 independent of whether the symmetric information is meager or abundant.

Gilboa (2009) interprets Schmeidler's observation as expressing a form of "cognitive unease", namely a feeling that the theory of subjective probability which springs naturally from Bayesian epistemology, is silent on one fundamental aspect of rationality (in its informal meaning). But why is it so? Suppose that some matter is to be decided by the toss of a coin. According to Schmeidler's line of argument, I should prefer tossing my own, rather than some one else's coin, on the basis, say of the fact that I have never observed signs of "unfairness" in my coin, whilst I just don't know anything about the stranger's coin. See also Gilboa et al. (2012); Gilboa (2009). This argument is of course reminiscent of the Ellsberg two-urns problem, which had been anticipated in Keynes (1921).

Similar considerations have been put forward in artificial intelligence and in the foundations of statistics. An early amendment of probability theory aimed at capturing the asymmetry between uncertainty and ignorance is known as the theory of Belief Functions (Shafer, 1976; Denoeux, 2016). Key to representing this asymmetry is the relaxation of the additivity axiom of probability. This in turn may lead to situations in which the *probabilistic excluded middle* does not hold. That is to say an agent could rationally assign belief less than 1 to the classical tautology $\theta \vee \neg\theta$. Indeed, as we now illustrate, the problem with normalising on tautologies is much more general.

2.2 Against the logic norm: the argument from tractability

Recall that classical propositional logic is *decidable* in the sense that for each sentence θ of the language there is an effective procedure to decide wether θ is a tautology or not. Such a procedure, however, is unlikely to be *feasible*, that is to say executable in practice. In terms of the theory of computational complexity this means that there is probably no algorithm running in polynomial time. So, a consequence of the seminal 1971 result by Stephen Cook, the tautology problem for classical logic is widely believed to be *intractable*.

If this conjecture is correct, we are faced with a serious foundational problem when imposing the normalisation of probability on tautology. For we are imposing agents constraints of rationality which they simply may never be able to satisfy.

It is remarkable that L.J. Savage had anticipated this problem with the Bayesian norms he centrally contributed to defining. To this effect he observed in Savage (1967) the following:

> A person required to risk money on a remote digit of π would have to compute that digit in order to comply fully with the theory though this would really be wasteful if the cost of computation were more than the prize involved. For the postulates of the theory imply that you should behave in accordance with the *logical implications of all that you know.* Is it possible to improve the theory in this respect, making allowance within it for the cost of thinking, or would that entail paradox [...] , as I am inclined to believe but unable to demonstrate ? If the remedy is not in changing the theory but rather in the way in which we attempt to use it, clarification is still to be desired. (Our emphasis)

Fifty years on, the difficulty pointed out by Savage failed to receive the attention it deserves. As the remainder of this note illustrates, however, framing the issue logically brings about significant improvements in our understanding of the key issues, paving the way for a tractable approximation of Bayesian rationality – or rational beliefs real agents can have.

3. Logic, algebra and probability

A well-known representation result (see, e.g. Paris (1994)) shows that every probability function arises from distributing the unit mass across the 2^n atoms of the Boolean (Lindenbaum) Algebra generated by the propositional language $L = \{p_1, \dots p_n\}$, and conversely, that a probability function on L is completely determined by the values it takes on such atoms. Such a representation makes explicit the dependence of probability on classical logic. This has important and often underappreciated consequences. Indeed logic plays a twofold role in the theory of probability. First, logic provides the language in which events –the bearers of probability– are expresses, combined and evaluated. The precise details depend on the framework. See Flaminio et al. (2014) for a characterisation of probability on classical logic, and Flaminio et al. (2015) for the general case of Dempster-Shafer belief functions on many-valued events.

In measure-theoretic presentations of probability, events are identified with subsets of the field generated by a given sample space Ω. A popular interpretation for Ω is that of the elementary outcomes of some experiment, a view endorsed by A.N. Kolmogorov, who insisted on the generality of his axiomatisation. More precisely, let $\mathcal{M} = (\Omega, \mathcal{F}, \mu)$ a *measure space* where, $\Omega = \{\omega_1, \omega_2 \dots\}$ is the set of elementary outcomes, $\mathcal{F} = 2^\Omega$ is the field of sets ($\sigma-$algebra) over Ω. We call *events* the elements of $\mathcal{F}$, and $\mu : \mathcal{F} \to [0, 1]$ a *probability measure* if it is normalised, monotone and σ-additive, i.e.

(K1) $\mu(\Omega) = 1$

(K2) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$

(K3) If $\{E\}_i$ is a countable family of pairwise disjoint events then $P(\bigcup_i E_i) = \sum_i P(E_i)$

The Stone representation theorem for Boolean algebras and the representation theorem for probability functions recalled above guarantee that the measure-theoretic axiomatisation of probability is equivalent to the logical one, which is obtained by letting a function from the language L to the real unit interval be a *probability function* if

(PL1) $\models \theta \Rightarrow P(\theta) = 1$

(PL2) $\models \neg(\theta \wedge \phi) \Rightarrow P(\theta \vee \phi) = P(\theta) + P(\phi)$.

Obvious as this logical "translation" may be, it highlights a further role for logic in the theory of probability, in addition that is to the linguistic one pointed out above. This role is best appreciated by focussing on the consequence relation $\models$ and can be naturally referred to as inferential.

In its measure-theoretic version, the normalisation axiom is quite uncontroversial. Less so, if framed in terms of classical tautologies, as in PL1. Indeed both arguments against Bayesian norms discussed informally above, emerge now formally. The first is to do with the fact that $\models$ interprets symmetrically "knowledge" and "ignorance" as captured by the fact that $\models \theta \vee \neg\theta$ is a tautology. Indeed similarly bothersome consequences follow directly from $PL1$ and $PL2$, namely

1. $P(\neg\theta) = 1 - P(\theta)$
2. $\theta \models \phi \Rightarrow P(\theta) \leq P(\phi)$

2 implies that if θ and ϕ are logically equivalent they get equal probability.

The argument from information recalled in Section 2.1 above clearly has its logical roots in the semantics of classical logic.

Similarly, the argument from tractability of Section 2.2 leads one into questioning the desirability of normalising probability on *any* classical tautology. Taken as a norm of rationality this requires agents to be capable of reasoning beyond what is widely accepted as feasible. Again, the unwelcome features of probability are rooted in classical logic.

A further, important, feature which emerges clearly in the logical presentation of probability is that uncertainty is resolved by appealing to the semantics of classical logic. This leads to the piecemeal identification of "events" with "sentences" of the logic. This identification, however, is not as natural as one may think.

On the one hand, an *event*, understood classically, either happens or not. A sentence expressing an event, on the other hand is evaluated in the binary set as follows

$$v(\theta) = \begin{cases} 1 \text{ if the event obtained} \\ 0 \text{ otherwise.} \end{cases}$$

Hence, the probability of an event $P(\theta) \in [0, 1]$ measures the agent's degree of belief that the event did or will obtain. Finding this out is, in most applications, relatively obvious. However, as pointed out in Flaminio et al. (2014), a general theory of what it means for "states of the world" to "resolve uncertainty" is far from trivial.

A more natural way of evaluating events arises taking an *information-based* interpretation of uncertainty resolution. The key difference with the previous, classical case, lies in the fact that this leads naturally to a *partial* evaluation of events, that is

$$v^i(\theta) = \begin{cases} 1 \text{ if I am informed that } \theta \\ 0 \text{ if I am informed that } \neg\theta \\ \bot \text{ if I am not informed about } \theta. \end{cases}$$

Quite obviously standard probability logic does *not* apply here, because the classical resolution of uncertainty has no way of expressing the $\bot$ condition.

As the next section shows, by looking for a logic which fixes this information asymmetry, we will also find a logic which deals successfully with the tractability problem.

4. An informational view of propositional logic

The main idea underlying the informational view of classical propositional logic is to replace the notions of "truth" and "falsity", by "informational truth" and "informational falsity", namely *holding the information* that a sentence φ is true, respectively false. Here, by saying that an agent a *holds* the information that φ is true or false we mean that this information (i) is *accepted* by a in the sense that a is ready to *act upon it*[1] (ii) it is *feasibly available* to a, in the sense that a has the means to obtain it in practice (and not only in principle); given the (probable) intractability of classical propositional logic this condition is not in general preserved by the corresponding consequence relation.

Clearly, these notions do not satisfy the informational version of the Principle of Bivalence: it may well be that for a given φ, we neither hold the information that φ is true, nor do we hold the information that φ is false. Knowledge and ignorance are not treated symmetrically under the informational semantics. However, in this paper we assume that they do satisfy the informational version of the Principle of Non-Contradiction: no agent can *actually* possess both the information that φ is true and the information that φ is false, as this could be deemed to be equivalent to possessing no definite information about φ.[2]

4.1 Informational semantics

We use the values 1 and 0 to represent, respectively, informational truth and falsity. When a sentence takes neither of these two defined values, we say that it is *informationally indeterminate*. It is technically convenient to treat informational indeterminacy as a third value that we denote by "$\bot$".[3] The three values are partially ordered by the relation $\preceq$

1. The kind of justification for this acceptance and whether or not the agent is human or artificial do not concern us here. Acceptance may include some (possibly non-conclusive) evidence that a deems sufficient for acceptance, or communication from some external source that a regards as reliable.
2. Notice that this assumption does not rule out the possibility of *hidden* inconsistencies in an agent's information state, but only of inconsistencies that can be feasibly detected by that agent. It is, however, possible to investigate paraconsistent variants of this semantics in which even this weak informational version of the Principle of Non-Contradiction is relaxed. This will be the subject of a subsequent paper.
3. This is the symbol for "undefined", the bottom element of the information ordering, not to be confused with the "falsum" logical constant.

$\wedge$	1	0	$\bot$
1	1	0	$\bot$
0	0	0	0
$\bot$	$\bot$	0	$\bot, 0$

$\vee$	1	0	$\bot$
1	1	1	1
0	1	0	$\bot$
$\bot$	1	$\bot$	$\bot, 1$

$\neg$	
1	0
0	1
$\bot$	$\bot$

Figure 1: Informational tables for the classical operators

such that $v \preceq w$ ("v is less defined than, or equal to, w") if, and only if, $v = \bot$ or $v = w$ for $v, w \in \{0, 1, \bot\}$.

Note that the old familiar truth tables for $\wedge, \vee$ and $\neg$ are still intuitively sound under this informational reinterpretation of 1 and 0. However, they are no longer exhaustive: they do not tell us what happens when one or all of the immediate constituents of a complex sentence take the value $\bot$. A remarkable consequence of this approach is that the semantics of $\vee$ and $\wedge$ becomes, as first noticed by Quine (1973, pp. 75–78), *non-deterministic*. In some cases an agent a may accept a disjunction $\varphi \vee \psi$ as true while abstaining on both components φ and ψ. To take Quine's own example, if I cannot distinguish between a mouse and chipmunk, I may still hold the information that "it is a mouse or a chipmunk" is true while holding no definite information about either of the sentences "it is a mouse" and "it is a chipmunk". In other cases, e.g. when the component sentences are "it is a mouse" and "it is in the kitchen" and I still hold no definite information about either, the most natural choice is to abstain on the disjunction. Similarly, a may reject a conjunction $\varphi \wedge \psi$ as false while abstaining on both components. To continue with Quine's example, I may hold the information that "it is a mouse and a chipmunk" is false, while holding no definite information about either of the two component sentences. But if the component sentences are "it is a mouse" and "it is in the kitchen" and I abstain on both, I will most probably abstain also on their conjunction. In fact, this phenomenon is quite common as far as the ordinary notion of information is concerned and the reader can figure out plenty of similar situations. Thus, depending on the "informational situation", when φ and ψ are both assigned the value $\bot$, the disjunction $\varphi \vee \psi$ may take the value 1 or $\bot$, and the conjunction $\varphi \wedge \psi$ may take the value 0 or $\bot$.

As a consequence of this informational interpretation, the traditional truth-tables for the $\vee$, $\wedge$ and $\neg$ should be replaced by the "informational tables" in Figure 1, where the value of a complex sentence, in some cases, is not uniquely determined by the value of its immediate components.[4] A non-deterministic table for the informational meaning of the Boolean conditional can be obtained in the obvious way, by considering $\varphi \to \psi$ as having the same meaning as $\neg\varphi \vee \psi$ (see D'Agostino, 2015, p. 82).

4. In his (Quine, 1973) Quine calls them "verdict tables" and the values are "assent", "dissent" and "abstain". This non-deterministic semantics was subsequently and independently re-proposed (with no apparent connection with the intuitive interpretation given by Quine) by Crawford and Etherington (1998) who claimed without proof that it provides a characterization of unit resolution (a tractable fragment of resolution that requires formulae to be translated into clausal form). The general theory of non-deterministic semantics for logical systems has been brought to the attention of the logical community and extensively investigated (with no special connection with tractablity) by Arnon Avron and co-authors (see Avron and Zamansky (2011) for an overview).

4.2 Depth-bounded Boolean logics

In (D'Agostino et al., 2013) and (D'Agostino, 2015) it is shown that the informational semantics outlined in the previous section provides the basis to define an infinite hierarchy of *tractable* deductive systems (with no syntactic restriction on the language adopted) whose upper limit coincides with classical propositional logic. As will be clarified in the sequel the tractability of each layer is a consequence of the shift from the classical to the informational interpretation of the logical operators (that is the same throughout the hierarchy) and on an upper bound on the nested use of "virtual information", i.e. information that the agent *does not actually hold*, in the sense specified in the previous section.

Definition 1 *A* 0*-depth information state is a valuation* V *of the formulae in* L *that agrees with the informational tables.*

Note that, given the non-determinism of the informational tables, the valuation V is not uniquely determined by an assigment of values to the atomic sentences. For example the valuation V_1 that assigns $\bot$ to both p and q and $\bot$ to $p \vee q$ is as admissible as the valuation V_2 that still assigns $\bot$ to both p and q, but 1 to $p \vee q$. Let S_0 be the set of all 0-depth information states.

Definition 2 *We say that* φ *is a 0-depth consequence of a finite set* Γ *of sentences, and write* $\Gamma \models_0 \varphi$*, when*

$$(\forall V \in S_0)\, V(\Gamma) = 1 \Longrightarrow V(A) = 1.$$

We also say that Γ *is* 0-depth inconsistent*, and write* $\Gamma \models_0$ *if there is no* $V \in S_0$ *such that* $V(\Gamma) = 1$.

It is not difficult to verify that $\models_0$ is a Tarskian consequence relation, i.e., it satisfies reflexivity, monotonicity, transitivity and substitution invariance.

In fact, it can be shown that we do not need to consider valuations of the whole language L but can restrict our attention to the subformulae of the formulae that occur as premises and conclusion of the inference under consideration. Let us call *search space* any finite set Λ of formulae that is closed under subformulae, i.e., if φ is a subformula of a formula in Λ, $\varphi \in \Lambda$.

Definition 3 *A 0-depth information state over a search space* Λ *is a valuation* V *of* Λ *that agrees with the informational tables.*

Let S_0^Λ be the set of all 0-depth information states over a search space Λ. Given a finite set Δ of formulae, let us write Δ^* to denote the search space consisting of all the subformulae of the formulae in Δ. Then, it can be shown that:

Theorem 4 $\Gamma \models_0 \varphi$ *if and only if* $(\forall V \in S_0^{(\Gamma \cup \{\varphi\})^*})\, V(\Gamma) = 1 \Longrightarrow V(A) = 1$. *Moroever,* $\Gamma \models_0$ *if and only if there is no* $V \in S_0^{(\Gamma \cup \{\varphi\})^*}$ *such that* $V(\Gamma) = 1$.

On the basis of the above result, in (D'Agostino et al., 2013) it is shown that $\models_0$ is *tractable*:

Theorem 5 *Whether or not* $\Gamma \models_0 \varphi$ *(*Γ *is* 0*-depth inconsistent) can be decided in time* $O(n^2)$ *where* n *is the total number of occurrences of symbols in* $\Gamma \cup \{\varphi\}$ *(in* Γ*).*

A simple proof system that is sound and complete with respect to $\models$ is shown in (D'Agostino et al., 2013; D'Agostino, 2015) in the form of a set of introduction and elimination rules (in the fashion of Natural Deduction) that are only based on *actual information*, i.e., information that is held by an agent, with no need for *virtual information*, i.e., simulating information that does not belong to the current information state, as happens in case-reasoning or in some ways of establishing a conditional (as in the introduction rule for the conditional in Gentzen-style natural deduction).

The subsequent layers of the hierarchy depend on fixing an upper bound on the depth at which the nested use of virtual information is allowed.

Let $\sqsubseteq$ be the partial ordering of 0-depth information states (over a given search space) defined as follows: $V \sqsubseteq V'$ if and only if V' is a refinement of V or is equal to V, that is, for every formula φ in the domain of V and V', $V(\varphi) \neq \bot$ implies that $V'(\varphi) = V(\varphi)$.

Definition 6 *Let* V *be a 0-depth information state over a search space* Λ*.*

- $V \Vdash_0 \varphi$ *if and only if* $V(\varphi) = 1$
- $V \Vdash_{k+1} \varphi$ *if and only if*

$$(\exists \psi \in \Lambda)(\forall V' \in S_0^{\Lambda})\, V \sqsubseteq V' \text{ and } V'(\psi) \neq \bot \Longrightarrow V' \Vdash_k \varphi.$$

Here $\Vdash_j$, with $j \in \mathbb{N}$, is a kind of "forcing" relation and the shift from one level of depth to the next is determined by simulating refinements of the current information state in which the value of some $\psi \in \Lambda$ is defined (either 1 or 0) and checking that in either case the value of φ is forced to be 1 at the immediately lower depth. Such use of a definite value for ψ, that is not even implicitly contained in the current information state V of the agent, is what we call *virtual information.*

Definition 7 *A* k*-depth information state over a search space* Λ *is a valuation* V *of* Λ *that agrees with the informational tables and is closed under the forcing relation* $\Vdash_k$*.*

Let S_k^{Λ} be the set of all k-depth information states over Λ.

Definition 8 *We say that* φ *is a* k*-depth consequence of* Γ*, and write* $\Gamma \models_k \varphi$ *if*

$$(\forall V \in S_k^{(\Gamma \cup \{\varphi\})^*})\, V(\Gamma) = 1 \Longrightarrow V(\varphi) = 1.$$

We also say that Γ *is* k*-depth inconsistent, and write* $\Gamma \models_k$ *, if there no* $V \in S_k^{(\Gamma \cup \{\varphi\})^*}$ *such that* $V(\Gamma) = 1$*.*

It can also be shown that $\Gamma \models_k \varphi$ if and only if there is a finite sequence $\psi_1, \ldots, \psi_n$ such that $\psi_n = \varphi$ and for every element ψ_i of the sequence, either (i) ψ_i is a formula in Γ or (ii) $V \Vdash_k \psi_i$ for all $V \in S_0^{(\Gamma \cup \{\varphi\})^*}$ such that $V\{\psi_1, \ldots, \psi_{i-1}\} = 1$.

Unlike $\models_0$, $\models_k$ is not a Tarskian consequence relation, but gets very close to being such, for $\models_k$ satisfies reflexivity, monotonicity, substitution invariance and the following restricted

version of transitivity in which the "cut formula" is required to belong to the search space defined by the deduction problem under consideration.

$$(\forall \psi \in (\Gamma \cup \{\varphi\})^*)\ \Gamma \models_k \psi \text{ and } \Delta, \psi \models_k \varphi \Longrightarrow \Gamma, \Delta \models_k \varphi. \qquad \text{(Bounded Transitivity)}$$

In (D'Agostino et al., 2013) it is shown that $\models_k$ is tractable for every fixed k.

Theorem 9 *Whether or not $\Gamma \models_k \varphi$ (Γ is k-depth inconsistent) can be decided in time $O(n^{2k+2})$, where n is the total number of occurrences of symbols in $\Gamma \cup \{\varphi\}(\Gamma)$.*

Observe that, by definition, if $\Gamma \models_j \varphi$ (Γ is j-depth inconsistent), then $\Gamma \models_k \varphi$ (Γ is k-depth inconsistent) for every $k > j$. Classical proposition logic is the limit of the sequence of the depth-bounded consequence relations $\models_k$ as $k \to \infty$.

A proof system for each of the k-depth approximations is obtained by adding to the introducton and elimination rules for $\models_0$ a single structural rule that reflects the use of virtual information in Definition 6, and bounding the depth at which nested applications of this rule are allowed (see (D'Agostino et al., 2013; D'Agostino, 2015) for the details and a discussion of related work).

5. Towards a prescriptive theory of Bayesian rationality

Let us briefly recap. By framing probability logically we are able to locate the source of a number of important criticisms which are commonly held up against Bayesian rationality in classical logic. The theory of Depth-Bounded Boolean logics meets some of those objections, and gives us an informational semantics leading to a hierarchy of tractable approximations of classical logic. The logical axiomatisation of probability recalled above naturally suggests to investigate which notion of rational belief is yielded once $\models$ is replaced with $\models_k$ in PL1-PL2 above.

This gives us a natural desideratum, namely to construct a family of rational belief measures B_i from L to $[0, 1]$, $i \in \mathbb{N}$ acting as the analogues of probability functions on Depth-bounded logics. Since DBLs coincide, in the limit, with classical propositional logic, our desideratum is then the construction of a hierarchy of belief measures $B_0, \ldots, B_k \ldots$ which asymptotically coincides with probability, i.e. such that for all sentences θ, $B_\infty(\theta) = P(\theta)$.

Each element in the resulting hierarchy would then be a natural candidate to providing a logically rigorous account of a *prescriptive* model of rational belief, in the sense of Bell et al. (1988): every agent whose deductive capabilities are bounded by $\models_k$ must quantify, on pain of irrationality, uncertainty according to B_k.

There is an obvious link between the interpretation of disjunction given by the non-deterministic informational semantics discussed in Section 4.1 and the behaviour of this logical connective in quantum logic. As is well-known, in quantum logic a proposition θ can be represented as a closed subspace M_θ of the Hilbert space $\mathcal{H}$ under consideration. The disjunction $\varphi \vee \psi$ is not represented by the union of M_φ and M_ψ, for in general the union of two closed subspaces is not a closed subspace, but by $M_\varphi \sqcup M_\psi$, i.e. as the smallest closed subspace including both M_φ and M_ψ. So, as is the case for the informational interpretation of disjunction given by the non-deterministic semantics discussed above, a disjunction $\varphi \vee \psi$ in quantum logic may be true even if neither of the disjuncts are true, since $M_\varphi \sqcup M_\psi$ may contain vectors that are not contained in $M_\varphi \cup M_\psi$. On this point see

(Aerts, 2000) and (Dalla Chiara et al, 2004). The negative part of the analogy concerns the behaviour of conjunction which in quantum logic is interpreted as $M_\varphi \cap M_\psi$, so that if a conjunction is false, at least one of the two conjuncts must be false, which departs from the informational intepretation of this operator given by our non-deterministic table. We also point out that this connection between the non-deterministic semantics of Depth-bounded Boolean Logics and the semantics of Quantum Logic opens to a natural parallel between our desideratum and quantum probabilities. This is reinforced by recent experimental findings in the cognitive sciences (Pothos and Busemeyer, 2013; Oaksford, 2014) suggesting that some features of Bayesian quantum probability (Pitowsky, 2003) provide accurate descriptions of experimental subjects.

The key step towards achieving our goal will be of course to define the sense in which we take any B_k to be a *rational* belief measure. The task, as it can be easily figured out, is far from trivial. Though encouraging, our preliminary results suggest that much work is still to be done in this direction. At the same time they suggest that the consequences of such a fully-fledged framework will be far reaching, as it will provide significant steps towards identifying norms of rationality *realistic agents can abide to*.

Aknowledgements. The authors would like to thank the two referees for the careful reading, improvement suggestions and encouraging remarks.

References

D. Aerts, E. D'Onts, and L. Gabora. Why the disjunction in quantum logic is not classical. *Foundations of Physics*, 30:1473–1480, 2000.

T. Augustin, F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes, editors. *Introduction to Imprecise Probabilities*. Wiley, 2014.

A. Avron and A. Zamansky. Non-deterministic semantics for logical systems. In D.M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 16, pages 227–304. Springer Verlag, 2nd edition, 2011.

D.E. Bell, H. Raiffa, and A. Tversky. *Decision making: Descriptive, normative, and prescriptive interactions*. Cambridge University Press, 1988.

J.M. Crawford and D.W. Etherington. A non-deterministic semantics for tractable inference. In *AAAI/IAAI*, pages 286–291, 1998.

M.L. Dalla Chiara, R. Giuntini and R. Greechie. *Reasoning in Quantum Theory*. Springer-Science+Business Media, B.V. 2004.

M. D'Agostino. An informational view of classical logic. *Theoretical Computer Science*, 606:79–97, 2015. doi: http://dx.doi.org/10.1016/j.tcs.2015.06.057.

M. D'Agostino, M. Finger, and D. Gabbay. Semantics and proof-theory of depth bounded Boolean logics. *Theoretical Computer Science*, 480:43–68, 2013.

B. de Finetti. *Theory of Probability, Vol 1*. John Wiley and Sons, 1974.

Thierry Denoeux. 40 years of DempsterShafer theory *International Journal of Approximate Reasoning* 79- 1-6, 2016.

D. Ellsberg. Risk, Ambiguity and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

T. Flaminio, L. Godo, and H. Hosni. On the logical structure of de Finetti's notion of event. *Journal of Applied Logic*, 12(3):279–301, 2014.

T. Flaminio, L. Godo, and H. Hosni. Coherence in the aggregate: A betting method for belief functions on many-valued events. *International Journal of Approximate Reasoning*, 58:71–86, 2015.

I. Gilboa. *Theory of Decision under Uncertainty*. Cambridge University Press, 2009.

I. Gilboa and M. Marinacci. Ambiguity and the Bayesian Paradigm. In D. Acemoglu, M. Arellano, and E. Dekel, editors, *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society*. Cambridge University Press, 2013.

I. Gilboa, A. Postlewaite, and D. Schmeidler. Rationality of belief or: Why Savage's axioms are neither necessary nor sufficient for rationality. *Synthese*, 187(1):11–31, 2012.

J.M Keynes. *Treatease on Probability*. Harper & Row, 1921.

F.H. Knight. *Risk, uncertainty and profit*. Beard Books Inc, 1921.

D. V. Lindley. *Understanding uncertainty*. John Wiley and Sons, 2006.

M. Oaksford. Normativity, interpretation, and Bayesian models. *Frontiers in Psychology*, 5, 15, 2014.

J.B. Paris. *The uncertain reasoner's companion: A mathematical perspective*. Cambridge University Press, 1994.

G. Parmigiani and L. Inoue. *Decision Theory: Principles and Approaches*. Wiley, 2009.

I. Pitowsky. Betting on the outcomes of measurements: A Bayesian theory of quantum probability. *Studies in History and Philosophy of Science Part B - Studies in History and Philosophy of Modern Physics*, 34(3), 395414, 2003.

E. M. Pothos, and J. R. Busemeyer. Can quantum probability provide a new direction for cognitive modeling? *Behavioral and Brain Sciences*, 36(3), 255274,2013.

W.V.O. Quine. *The Roots of Reference*. Open Court, 1973.

L.J. Savage. Difficulties in the theory of personal probability. *Philosophy of Science*, 34(4): 305–310, 1967.

L.J. Savage. *The Foundations of Statistics*. Dover, 2nd edition, 1972.

D. Schmeidler. Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57(3):571–587, 1989.

G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.

Proceedings of Machine Learning Research 58 (2016) 111-127

Hindsight Bias Impedes Learning

Shaudi Mahdavi SHAUDI@WHARTON.UPENN.EDU
Wharton School, University of Pennsylvania, Philadelphia, PA 19148.

M. Amin Rahimian MOHAR@SEAS.UPENN.EDU
Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19148.

Editor: Tatiana V. Guy, Miroslav Kárný, David Rios-Insua, David H. Wolpert

Abstract

We propose a model that addresses an open question in the cognitive science literature: How can we rigorously model the cognitive bias known as hindsight bias such that we fully account for critical experimental results? Though hindsight bias has been studied extensively, prior work has failed to produce a consensus theoretical model sufficiently general to account for several key experimental results, or to fully demonstrate how hindsight impedes our ability to learn the truth in a repeated decision or social network setting. We present a model in which agents aim to learn the quality of their signals through repeated interactions with their environment. Our results indicate that agents who are subject to hindsight bias will always believe themselves to be high-type "experts" regardless of whether they are actually high- or low-type.

Keywords: Hindsight Bias, Bounded Rationality, Sequential Decision-making, Non-Bayesian Social Learning

1. Introduction

Hindsight bias, also known as the curse of knowledge, is a pervasive cognitive bias that captures the difficulty of faithfully reconstructing one's foresight knowledge in retrospect, after one has learned actual outcome or correct answer of a decision-making task (Fischhoff and Beyth, 1975; Camerer et al., 1989). Hindsight bias has been shown to contaminate judgment and decision-making on tasks ranging from predicting outcomes to arithmetic calculation to knowledge retrieval, and in people of all ages, skill and knowledge levels, intelligence, and cultures (Pohl et al., 2002; Bernstein et al., 2011; Pohl et al., 2010; Coolin et al., 2014).

Hindsight bias has been observed in many domains, in both the laboratory and the real world. In politics, pundits and laypeople alike commonly forget their pre-election predictions after a surprising election result, and report feeling that they had in fact anticipated the actual outcome all along. In healthcare settings, diagnosing physicians often experience exaggerated feelings of having been able to anticipate patient outcomes after the fact, or cannot objectively retroactively assess colleagues' diagnostic competence in cases where they begin knowing the patient's fate. Analogously, lawyers commonly claim to have known the outcome all along in the aftermath of a verdict, regardless of what they actually predicted before and during the trial. In gambling, sports bettors' mistaken recollection of having

predicted shocking game results has been shown to give some gamblers unearned confidence to take increasingly risky future bets. Similar examples abound in educational, research, business, financial, and scientific decision making (Fischhoff and Beyth, 1975; Fischhoff, 2007; Fiedler et al., 2015; Arkes, 2013; Harley, 2007). As we discuss in this paper, the effects of hindsight bias, such as overconfidence, misattribution of blame to external factors, and normalized perception of unlikely tail events, can gravely distort heuristics and future decisions.

Ubiquity aside, hindsight bias is a compelling topic for study because of the striking difficulty of eliminating hindsight bias from people's decision-making. Indeed, Fischhoff (1975) and many others have deemed the so-called curse of knowledge an inevitability. This supposed inevitability bears out intuition: Once an outcome or answer is revealed, a person's outlook, and modes of interpretation, and judgment can be fundamentally changed in complex ways. Just as knowledge cannot be willed away, in the absence of costly, scrupulous bookmarking of knowledge and thoughts at various points in time, it can be extremely challenging to faithfully reconstruct one's prior, naive memory, judgments, and thought processes from a new position of information privilege.

We contribute to the substantial literature on hindsight bias by introducing a new, more complex, more general model of hindsight biased decision-making than currently exists. Using this model, we examine biased decision-making in two novel settings: when people can make infinitely many repeated judgments, and when they make decisions while communicating with a large social network of heterogeneous neighbors.

1.1 What is Hindsight Bias?

First documented by Fischhoff and Beyth (1975), hindsight bias formalizes the well-known "I knew it all along" distortion of cognition and information processing. This highly pervasive bias clouds a decision maker's ability to recall her prior expectations for an outcome after learning new information such as when the true state of the world, or to reliably anticipate the decisions of people she knows are not privy some knowledge she is. Outcomes and answers that seemed impossible to anticipate can suddenly seem obvious once the truth is revealed, because these truths either distort original memories, bias the reconstruction of foresight (pre-answer) knowledge, or both.

Interestingly, unless the revealed outcome or correct answer is subsequently debunked, most laboratory interventions attempting to mitigate hindsight bias have failed. In fact, several experiments have shown that even informing people they are prone to such bias, and instructing them to avoid allowing it to contaminate their answers during a range of decision-making tasks, fails to prevent hindsight from contaminating their answers (Fischhoff, 1976; Pohl and Hell, 1996; Bernstein et al., 2011).

Hindsight bias is generally identified and measured within two main experimental paradigms. The first, known as memory design, involves asking participants questions about a relatively obscure topic. These questions may involve judging the probability of a certain outcome, or answering a factual question (often with a numerical answer), such as the number of satellites of a planet. To assess hindsight bias, researchers first pose such a question to participants, reveal the answer or true outcome, and then again ask what the participants' original answer was. Hindsight bias is observed when participants' post-revelation answer

moves away from the original response and towards the correct answer. The number of times and distance a subject's answer move towards the correct answer determine the magnitude of her bias. The second, hypothetical design, involves first informing experimental subjects of the answer to a relatively obscure question, then asking how they would have answered before learning the answer, or how a hypothetical naïve peer would have answered the same question. In this setup, hindsight bias would be expected to cause subjects to report hypothetical answers closer to the correct answer when they are informed of the answer than when they are not (Coolin et al., 2016; Leary, 1982)

Hindsight bias affects judgment in a broad range of settings. For example, Biais and Weber (2009) study hindsight bias exhibited by stock market traders. Traders need to update their risk assessments when volatility is stochastic, using realized returns. On observing unexpectedly positive or negative returns, rational agents should raise their volatility estimates. Hindsight biased traders will form inaccurate beliefs compared to fully rational agents, because in retrospect, they perceive their prior beliefs as having accounted for this tail event. In doing so, they fail to consider such returns were unexpected, and thus underestimate variances. This failure to correctly assimilate information naturally leads to inferior trades, and thus suboptimal financial performance, as they fail to hedge appropriately, fail to reject analyses that are challenged by the facts, and fail to differentiate between their information and that of their peers.

Next, consider the decisions in the context of medicine (Danz et al., 2015). First, we note that after a tumor has been diagnosed, in hindsight, 90 percent of lung cancers and 70 percent of breast cancers that had previously not been recognized are observed on radiographs (Berlin, 2003). Suppose that in the past, a trainee doctor had seen a radiograph of an patient but overlooked the tumor. Suppose also, as in the situation posed above, that the tumor was discovered later. A supervising doctor not involved with the diagnosis, who, armed with the knowledge of the existence of this tumor, is hindsight biased, unfairly evaluates the subordinate's ability to diagnose cancer. In other words, being informed about the location of the tumor, the supervisor overestimates the likelihood that she would herself have found the tumor initially. Consequently, she might underestimate the abilities of the young trainee doctor, against the backdrop of her own inflated sense of ability (Madarász, 2011).

1.2 Origins of Hindsight Bias

Despite a substantial amount of research devoted to the topic, theoretical modeling work is sparse, and no consensus model of hindsight bias has been proposed – likely because generalizing across the incredible heterogeneity of settings and tasks addressed by experimental work on the subject appears daunting. The prevailing theories of hindsight bias can be hierarchically grouped into three categories of theoretical constructs (Roese and Vohs, 2012; Blank et al., 2008; Nestler et al., 2008; Bernstein et al., 2015).

The first theoretical umbrella relates to distorted memories, which can cause people to incorrectly recall that they had correctly guessed the actual outcome or answer. Within this category, the first memory-based theory is *automatic assimilation*, or the idea that the brain encodes the correct answer automatically, and with negligible effort (Fischhoff, 1982; Bernstein et al., 2012). Automatic assimilation is consistent with the observation that informing participants of their hindsight bias does not reduce bias of their responses. The

next theory is the *trace-strength hypothesis*, which asserts that memory traces, or records, for the original and correct answers possess different strengths, or degrees of accessibility, that determine the magnitude of the hindsight bias (Hell et al., 1988; Wood, 1978; Harley et al., 2004). The third memory construct is *recollection bias*, which describes the phenomenon in which learning the correct answer makes it more difficult to recall the original answer – either because the new memory overwrites or becomes more accessible than the previous memory (Erdfelder and Buchner, 1998; Erdfelder et al., 2007).

The second theoretical umbrella relates to the *reconstruction bias*, which arises when people attempt to reconstruct their naive foresight knowledge by some process that is, to some extent, biased or weighted toward the correct answer or revealed outcome (Schwarz and Stahlberg, 2003). The first theoretical construct in this category is *anchoring and adjustment*, in which participants anchor their response to correct answer, and adjust (perhaps incompletely) toward their original answer to a degree reflecting the magnitude of their bias (Epley and Gilovich, 2006). The second of these theories is *biased memory search*, according to which the correct answer corrupts memory recall by guiding the memory search toward content related to the correct answer (Hoffrage et al., 2000). The final construct, *metacognitive regulation theory*, asserts that adjustment toward the correct answer depends on one's confidence in the original and recalled answers, and on the extent of surprise caused by the newly-learned information (Müller and Stahlberg, 2007).

The final theoretical umbrella is *motivational theory*. In essence, this theory is rooted in the idea that people tend to misremember their original judgments or answers such that they are perceived as wiser or smarter than they actually believe themselves to be. Motivational theory presupposes that humans are motivated to maintain a sense of self-worth, which they derive from markers of intelligence such as decision-making aptitude, level of knowledge, and fidelity of memory. Three theories comprise this category. The first is *retroactive pessimism*, which asserts that people protect their sense of self-worth in the face of negative outcomes by convincing themselves a negative outcome was more likely than they believed it to be in foresight, and perhaps inevitable (Tykocinski and Steinberg, 2005). The second construct is *defensive processing*. According to this theory, people self-protectively perceive negative outcomes as less-predictable in hindsight to absolve themselves of responsibility and guilt for having made bad decisions that led to the outcome in question (Louie, 1999; Mark and Mellor, 1991). The last, *motivated sense-making* combines the two preceding theories, such that people use external and internal factors, depending on the nature and degree of surprise the experience, to explain discrepancies in expectation versus outcome (Pezzo, 2011).

1.3 Motivation and Contributions to the Literature

In this paper, we present a novel model of hindsight-biased learning that is both more generalizable and more complex than past models. First, thanks to its relative complexity, unlike existing models, our model clearly captures elements of all three aforementioned theoretical constructs. Our model accounts for faulty memories as well as faulty memory reconstruction (through anchoring and adjustment) as previous models have done, but with an additional element of randomness that accounts for motivational theory. By introducing this randomness, we account for the tension between a person's drive to produce correct answers to feel competent (motivational theory), and the orthogonal need to stress-test

hypotheses and admit heuristic weakness, in order to be able to improve future decision-making.

Second, we incorporate elements that allows us to distinguish among agents heterogeneous in multiple dimensions. We define classes of good decision-making and bad decision-making agents as a broad proxy for inherent differences in skill, domain knowledge, age, intelligence, memory capacity, that influence decision-makers' susceptibility to hindsight bias in the real world, and can account for result discrepancies in experiments administered to different populations of subjects (Roese and Vohs, 2012). We further accommodate subject heterogeneity by distinguishing between signal reception and signal perception. Unlike previous models, we distinguish between decision-makers not only in terms of what signals they receive, but also in how they perceive those signals – an additional layer of complexity that lends itself to increased generalizability.

Thanks to our modeling choice, we can contribute the first study of hindsight biased decision outcomes under infinitely repeated decisions, and in an arbitrarily large networked setting with social learning. More specifically, unlike previous models, which typically consider decision-making as a two-step process, we study the cumulative effects of hindsight bias over repeated decisions, such as those of a stock trader who makes investment decisions every day. We also study the implications of communications with a network of heterogeneous agents for hindsight-biased learning.

2. Modeling Hindsight Bias

There exist several models of hindsight bias in the literature. One of the most prominent is that of Madarász (2011). Here, an agent has a signal realization s, but does not correctly remember it. Instead, her memory is biased towards the state of the world revealed to be true in the meantime. Unlike the approaches used in the earlier models of Camerer et al. (1989) and Biais and Weber (2009) where the parameter $\lambda \in (0, 1)$ is used to form the linear combination of the true signal and the true state of the world, Madarász (2011) regards the parameter λ as the weight of the two beliefs. This is a subtle but highly critical distinction. A λ-hindsight biased agent, in forming her posterior belief, weights the signal realization from the first period with $(1 - \lambda)$ and the true state of the world with λ. Here, parameter λ denotes the strength of the hindsight bias. If we denote the posterior beliefs of the agent by $\pi_+(\cdot)$ and the truth state by ω, the biased agent's posterior belief becomes $\pi_+(\cdot) = \lambda\pi_+(\cdot|\omega) + (1 - \lambda)\pi_+(\cdot|s)$. This is a deviation from the Bayesian posterior $\pi_+(\cdot|s)$, which updates the prior belief, denoted by $\pi(\cdot)$, given the observation s according to the Bayes rule. Another prominent model introduced by Biais and Weber (2009) defines the hindsight bias by forming the reconstructed prior mean of the biased agent as a weighted average of the true prior mean and of the realization of the random variable.

In contrast to the above authors' approaches, we model the cognition of the hindsight-biased agent as a random process, such that: with probability $\lambda \in (0, 1]$ the biased agent believes she has observed the correct signal (coinciding with the truth), irrespective of of her actual observation. We thus distinguish between the agents' perceptions r_t and observation s_t by modeling both as random processes, which may or may not agree with the true state ω_t.

The observations s_t are a sequence of independent random variables[1], and the perceptions r_t are such that $r_t = s_t$ with probability $1 - \lambda$ and $r_t = \omega_t$ (the truth state) with probability λ; we thus investigate the dynamic effects as the agents receives a sequence of observations, while updating their beliefs based on their perceptions:

$$r_t = \begin{cases} \omega_t & \text{with probability} \quad \lambda, \\ \\ s_t & \text{with probability} \quad (1-\lambda). \end{cases} \tag{1}$$

The above equation captures the fact that the agent has a limited memory of her past signal. The first term represents the potentially false memory from projecting her knowledge of the state of the world onto her past signal, while the second represents a factual memory of the signal observed in the past. Stated differently, in this formulation, the agent's recollection of the signal realization, given in the second term, is corrupted by the agent's memory limitation and cognitive bias, which leads her to project her current information ω_t onto her past signal s_t.

We can also justify the model and the bias parameter λ from a motivational theory perspective: Every time that our decision-maker makes a mistake, she pays an emotional (fixed) cost of believing herself mistaken. On the other hand, every time she remembers her choice correctly, she benefits by learning about the quality of her choice (informational benefit), so λ could balance the expected (emotional) cost of making mistakes against the informational advantage of new correct samples. For if λ were such that the expected cost of mistake – in other words, the expected negative emotions caused by perception of a mistake – exceeded the informational benefits of a correct sample, the agent would be always flipping her choice to avoid the negative emotions; while if the expected informational benefits of a new sample were more valuable to her, then she would always be forcing herself to remember her samples correctly so that she could learn from them. The fact that the agent randomizes between the two strategies implies that the bias parameter λ should equalize the expected costs and benefits of the incoming samples. It is worthwhile to highlight two main assumptions that underlie equation (1): first the bias parameter λ is fixed over time (time-invariance); secondly, the agent is unaware of her hindsight bias and the underlying parameter λ as well as the stimuli random process s_t are not observable to her.

3. How does Hindsight bias limit rational learning?

Inspired by Danz et al. (2015), we assume that agents are heterogeneous in some characteristics that govern their decision-making abilities. We model this as follows: we assume nature draws the type of the agent (at random), and the agent's type can be either high (i.e., good) denoted by $\bar{\theta}$ or low (i.e., bad), denoted by $\underline{\theta}$. We further assume the prior probability of being a high-type agent is $\pi_0(\bar{\theta})$. The agent does not know her own type, but does have a prior $\pi_0(\cdot)$ about the type distributions. The latter represents the agent's knowledge of the distribution of the types before she makes any personal observations about her own type;

1. The independence assumption models repeated interactions of the agent with her environment, for example examinations of radiographs of different patients by a doctor. This assumption may be relaxed to some extent as long as the arguments pertaining to the application of the law of large numbers to the sum of random variables in the appendix remain valid.

for example, $\pi_0(\bar{\theta})$ can be induced by the prior knowledge about the fraction of radiologists who have superior skills in diagnostic radiography. At the beginning of each period t, nature draws the state of the world ω_t, which can be thought to represent a property of widespread interest such as the health status of a patient. States of the world are drawn independently across periods, and the two possible realizations of the state of the world, 0 and 1, are equally likely. The latter assumption is only for the sake of clarity and the ensuing analysis does not rely on this assumption because the inferences of the agent are based only on the agreements or disagreements between the perceived random process r_t and the truth states ω_t; such agreements and disagreements are in turn dependent on the received signals s_t through (1). The quality of the signal, i.e., the probability ρ that the signal corresponds to the true state of the world, depends on the type θ of the agent: A high type ($\bar{\theta}$) agent receives signals that are correct with probability $\bar{\rho}$ whereas a low type ($\underline{\theta}$)'s signals are only correct with probability $\underline{\rho}$, where $0.5 \leq \underline{\rho} < \bar{\rho} \leq 1$.

We begin with the assumption of independence across time. At every t, we have ω_t which takes either of the two values $\{0, 1\}$ with equal probability. Thus in lack of any side information, s_t and r_t also takes values $\{0, 1\}$ with equal probabilities. Here, the learning goal is to determine whether the agent's type, denoted as θ, is low ($\underline{\theta}$) or high ($\bar{\theta}$); symbolically, whether $\theta = \underline{\theta}$ or $\theta = \bar{\theta}$.

Concretely, in a real-world setting, a high-type agent is a person with the capacity to make better inferences about the state of the world than her low-type counterparts. For modeling purposes, we could use the type parameter θ to more generally capture heterogeneity of any of a vast number of qualities that influence judgment and decision-making – for example, domain knowledge, skill, or expertise; memory size or reliability; memory retrieval; intelligence, attention span; arithmetic aptitude; age to the extent that it influences cognitive processes; confidence; and many others.

3.1 Description of the Dynamic Model

At every round state ω_t is drawn and a signal s_t is observed. The likelihood of the signal given the state is determined by the agent's type, which could be high or low. We denote the observed signals of low and high type agents by $\underline{s}_t$ and $\bar{s}_t$, respectively. The signal of a high type agent is distributed according to

$$\bar{s}_t = \begin{cases} \omega_t & \text{with probability} \quad \bar{\rho} \\ 1 - \omega_t & \text{with probability} \quad 1 - \bar{\rho}. \end{cases}$$

while the likelihood of the signals for low type agent is given by

$$\underline{s}_t = \begin{cases} \omega_t & \text{with probability} \quad \underline{\rho}, \\ 1 - \omega_t & \text{with probability} \quad 1 - \underline{\rho}. \end{cases}$$

The signal perceived by the agent differs from her actual observations due to her bias. We denote the perceived signal of a high type agent by $\bar{r}_t$; similarly for a low type agent, we use $\underline{r}_t$.

The high and low types will subsequently update their beliefs based on their (biased) perceptions rather than their actual observations. Hence for a high type we have $\pi_t(\theta) = P\{\theta|\bar{r}_1, \bar{r}_2, ..., \bar{r}_t\}, \theta \in \{\bar{\theta}, \underline{\theta}\}$, and for a low type we obtain $\pi_t(\theta) = P\{\theta|\underline{r}_1, \underline{r}_2, ..., \underline{r}_t\}, \theta \in \{\bar{\theta}, \underline{\theta}\}$; i.e. the belief at time t is the conditional probability of each state given the perceptions made up until time t. In the case that there is no hindsight bias and the agent is fully rational, the perceptions correspond to the actual observations, and we have:

$$\bar{r}_t = \bar{s}_t \ \ \forall\, t \ \rightarrow \bar{r}_t = \bar{s}_t = \begin{cases} \omega_t & \text{with probability} \quad \bar{\rho} \\ \\ 1-\omega_t & \text{with probability} \quad 1-\bar{\rho}. \end{cases}$$

for a rational high type agent and

$$\underline{r}_t = \underline{s}_t \ \ \forall\, t \ \rightarrow \underline{r}_t = \underline{s}_t = \begin{cases} \omega_t & \text{with probability} \quad \underline{\rho} \\ \\ 1-\omega_t & \text{with probability} \quad 1-\underline{\rho}. \end{cases}$$

for a rational low type agent. For such a fully Bayesian agent, we have that

$$\pi_t(\theta = \bar{\theta}) \underset{a.s.}{\longrightarrow} \mathbb{1}(\theta = \bar{\theta}) \quad \text{as} \quad t \rightarrow \infty, \tag{2}$$

similarly for $\theta = \underline{\theta}$. Thus rational agents always learn their true types (with probability one).

On the other hand, a biased agent would, with probability $0 < \lambda < 1$, perceive her observed signal as the truth ω_t that is revealed to her, after her experiment with her environment (leading to her observation). Similarly, with probability $1-\lambda$ she will perceive her signal as it was, i.e. $\bar{s}_t$ for a high type and $\underline{s}_t$ for a low type. Consequently, we can derive the likelihoods of the perceived signals conditional on the state ω_t for low and high types as follows:

$$\bar{r}_t = \begin{cases} \omega_t & \text{with probability} \quad \lambda + (1-\lambda)\bar{\rho}, \\ \\ 1-\omega_t & \text{with probability} \quad (1-\lambda)(1-\bar{\rho}), \end{cases} \tag{3}$$

for a high type, and

$$\underline{r}_t = \begin{cases} \omega_t & \text{with probability} \quad \lambda + (1-\lambda)\underline{\rho}, \\ \\ 1-\omega_t & \text{with probability} \quad (1-\lambda)(1-\underline{\rho}), \end{cases}$$

for a low type. Now, we ask whether, following experimentation with their environment and after updating their beliefs based on their perceptions according to $\pi_t(\theta) = P\{\theta|\bar{r}_1, \bar{r}_2, ..., \bar{r}_t\}, \theta \in \{\bar{\theta}, \underline{\theta}\}$, and $\pi_t(\theta) = P\{\theta|\underline{r}_1, \underline{r}_2, ..., \underline{r}_t\}, \theta \in \{\bar{\theta}, \underline{\theta}\}$: can agents correctly learn their types, despite their hindsight bias?

Recall from (2) that rational agents learn their types, almost surely; in fact, one can show that this learning occurs at a rate that is asymptotically exponentially fast, cf. Section III.A of (Rahimian and Jadbabaie, 2015). The situation for hindsight biased agents is different:

Theorem 1 (Bayesian Learning with Hindsight) *Regardless of type, an agent with hindsight believes herself to be of a high type. In particular, low-type hindsight biased agents always mislearn their true states.*

Theorem 1 establishes that low type agents may come to incorrectly believe themselves to be of a higher type due to their hindsight following their experimentations with their environment. As illustrated by this case, hindsight bias prevents rational processing, and therefore, learning from past data. As Fischhoff (1982) wrote, hindsight bias prevents agents from rejecting their hypotheses about the world. Biased agents' memories and reconstruction processes are contaminated by the actual outcome information. Therefore, they underestimate past surprises by mistakenly factoring current information into the original expectation about the state of the world. As a result, agents use weak tests to evaluate hypotheses, and fail to properly update original hypotheses that they might have about the state of the world during experimentation.

In practical terms, the inability to be surprised, learn from the past and and reject hypotheses can be very destructive. As an illustration, stock traders burdened by such limitations will fail to recognize and revise incorrect hypotheses about the market, and thus fail to cut losses, sell investments, or avoid risky trades when a fully rational agent would. Also, hindsight-biased investors might incorrectly (irrationally) incorporate new informational content like signals from sources like earnings announcement or macroeconomic news, contaminating their memory reconstruction, giving them a falsely optimistic view of the optimality and value of their trading heuristics.

3.2 How is social learning affected by the hindsight bias?

Next, we consider the dynamic model of Section 3 in a network setting. Suppose that there are n agents. There is a type variable that is common to all agents, and determining this is the objective of the learning process. This type could encapsulate any number of characteristics of this society, such as the performance of the economy: if the agents are high type, then they live in a good economy; whereas if the agents are low-type, then they live in a bad economy. In this case, based on their experiences in their environment and on their interactions with each other, agents form an opinion about the state of their economy. Starting from a prior $\pi_0(\bar{\theta})$ at beginning of each round, neighboring agents communicate with each other and update their beliefs following a log-linear update with positive weights that sum to one, cf. (Rahimian and Jababaie, 2015) and Section 4 of (Rahimian and Jadbabaie, 2016) for behavioral and cognitive foundations of log-linear belief updates. After this initial communication they interact with their environment in the following manner: at each time t, nature draws a binary state $\omega_{i,t}$ for each of the n agents. Each agent observes a signal ($\bar{s}_{i,t}$ or $\underline{s}_{i,t}$ depending on society's type) conditional on her state and after the state is revealed to her, she changes her perception of her observation into a biased version ($\bar{r}_{i,t}$ or $\underline{r}_{i,t}$ depending on society's type) and she then updates her belief about her type based on her perception and using the Bayes rule. At the initiation of each future epoch, the agents communicate their beliefs following the same log-linear rule and interact with their environment in the manner described above. The states $\omega_{i,t}$, signals $s_{i,t}$, and perceptions $r_{i,t}$ encode the essential aspects of the agents' interactions with their environment: for example in a good economy,

each agent $i \in [n]$ can make better judgments about her state $\omega_{i,t}$ and thus makes decisions that lead to increased prosperity.

Our next theorem addresses the asymptotic outcome of the social learning with hindsight; in particular, whether the agents reach consensus and if they do, is it a consensus on truth or not. What if different agents have different λ_i? How does the heterogeneity in individual biases affect the learning outcome? All these questions are addressed by the following theorem.

Theorem 2 (Social Learning with Hindsight) *Consider a society of agents with a common type θ, which could be high or low, and at every epoch they communicate their beliefs about this common type. In addition, the agents also have a sequence of individual interactions with their environment and make private observations accordingly. If agents are subject to hindsight bias with parameter λ_i for each agent i, then regardless of type, a society of agents with hindsight believe themselves to be of a high type. In particular, a society of low-type hindsight biased agents always mis-learn their true types. This is true even if all but one agents are rational, as long as $\sum_{i=1}^{n} \lambda_i > 0$ and the agents communicate in a connected social network some agents' hindsight prevents all of them from correctly learning their true types.*

Theorem 2 reinforces the conclusions of Theorem 1 in that the low types agents' inference is compromised by their hindsight, causing them to incorrectly believe themselves to be of a higher type. Furthermore, the proof of Theorem 2 demonstrates that the hindsight bias for the network as a whole is simply the average hindsight over all agents in the network. Hence, interacting with biased agents contaminates the opinion of unbiased ones and subjects them to the same kinds of bias. Putting highly biased agents into such a setting with a large number of less-biased or unbiased agents would be an effective strategy for mitigating hindsight bias within a networked setting.

4. Conclusions

We have modeled the effect of hindsight bias on the inferences of a rational observer and also on the social learning outcomes in a network of arbitrary size. We introduced a more generalizable, more nuanced model that can capture multiple dimensions of agent heterogeneity, and has a demonstrable connection to each of three major hindsight bias theoretical constructs – both of which previous models lack. Our results also help verify the intuition that people become overconfident because they do not or cannot faithfully retain their past track record of wrong predictions, or correctly reconstruct earlier thought processes.

Acknowledgments

The authors gratefully acknowledge the editors of the NIPS 2016 Workshop Imperfect Decision Makers: Admitting Real-World Rationality for their comments on an earlier version of this paper.

Appendix A. Proofs of the Main Results

A.1 Proof of Theorem 1

For a high-type hindsight biased agent, we have, for the probability it puts on being high-type:

$$\pi_t(\bar{\theta}) = \mathbb{1}(\bar{r}_t = \omega_t)\frac{\pi_{t-1}(\bar{\theta})\bar{\rho}}{\pi_{t-1}(\bar{\theta})\bar{\rho} + \pi_{t-1}(\underline{\theta})\underline{\rho}} + \mathbb{1}(\bar{r}_t \neq \omega_t)\frac{\pi_{t-1}(\bar{\theta})(1-\bar{\rho})}{\pi_{t-1}(\bar{\theta})(1-\bar{\rho}) + \pi_{t-1}(\underline{\theta})(1-\underline{\rho})}.$$

and analogously, for the probability the same agent puts on being low-type,

$$\pi_t(\underline{\theta}) = \mathbb{1}(\bar{r}_t = \omega_t)\frac{\pi_{t-1}(\underline{\theta})\underline{\rho}}{\pi_{t-1}(\bar{\theta})\bar{\rho} + \pi_{t-1}(\underline{\theta})\underline{\rho}} + \mathbb{1}(\bar{r}_t \neq \omega_t)\frac{\pi_{t-1}(\underline{\theta})(1-\underline{\rho})}{\pi_{t-1}(\bar{\theta})(1-\bar{\rho}) + \pi_{t-1}(\underline{\theta})(1-\underline{\rho})}$$

whence we calculate the following ratio:

$$\frac{\pi_t(\bar{\theta})}{\pi_t(\underline{\theta})} = \begin{cases} \frac{\pi_{t-1}(\bar{\theta})}{\pi_{t-1}(\underline{\theta})}\frac{\bar{\rho}}{\underline{\rho}}, \text{ if } r_t = \omega_t \\ \\ \frac{\pi_{t-1}(\bar{\theta})}{\pi_{t-1}(\underline{\theta})}\frac{(1-\bar{\rho})}{(1-\underline{\rho})}, \text{ if } r_t \neq \omega_t \end{cases} \tag{4}$$

It follows that we have the recursive relation

$$\frac{\pi_t(\bar{\theta})}{\pi_t(\underline{\theta})} = \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_t=\omega_t)}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_t\neq\omega_t)}\frac{\pi_{t-1}(\bar{\theta})}{\pi_{t-1}(\underline{\theta})}$$

and hence, iterating on t, we easily see that:

$$\begin{aligned}\frac{\pi_t(\bar{\theta})}{\pi_t(\underline{\theta})} &= \prod_{\tau=1}^{t}\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_t=\omega_t)}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_t\neq\omega_t)}\frac{\pi_0(\bar{\theta})}{\pi_0(\underline{\theta})} \\ &= \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\sum_{\tau=1}^{t}\mathbb{1}(\bar{r}_t=\omega_t)}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{\sum_{\tau=1}^{t}\mathbb{1}(\bar{r}_t\neq\omega_t)}\frac{\pi_0(\bar{\theta})}{\pi_0(\underline{\theta})}.\end{aligned} \tag{5}$$

Now, it follows easily from the law of large numbers that: $\sum_{\tau=1}^{t}\mathbb{1}(\bar{r}_t = \omega_t) \underset{a.s.}{\longrightarrow} t(\lambda+(1-\lambda)\bar{\rho}))$ and $\sum_{\tau=1}^{t}\mathbb{1}(\bar{r}_t = \omega_t) \underset{a.s.}{\longrightarrow} t((1-\lambda)(1-\bar{\rho}))$. Finally, using the latter two limits in the above equation, we obtain, for a high-type agent:

$$\frac{\pi_t(\bar{\theta})}{\pi_t(\underline{\theta})} \underset{a.s.}{\longrightarrow} \left(\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{(\lambda+(1-\lambda)\bar{\rho})}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\lambda)(1-\bar{\rho})}\right)^{t}\frac{\pi_0(\bar{\theta})}{\pi_0(\underline{\theta})}.$$

The base of the exponential term for all range of parameters $0.5 < \underline{\rho} < \bar{\rho} < 1$ and $0 < \lambda < 1$ is always strictly greater than one[2], indicating that the belief ratio $\pi_t(\bar{\theta})/\pi_t(\underline{\theta})$ for high type

2. To see why, consider the first-order partial derivative of the function

$$f(\lambda, \bar{\rho}, \underline{\rho}) = \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{(1-\lambda)\bar{\rho}+\lambda}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\lambda)(1-\bar{\rho})}, \text{ with respect to } \lambda\text{:}$$

agent always converges almost surely to $+\infty$ leading to her learning her true type which is high. However, the case for a low type agent is different, following the same steps as above and using $\underline{r}_t$ instead of $\bar{r}_t$ for the perceived signals of a low type agent we obtain the following almost sure limit for the belief ratio of a low type agent:

$$\frac{\pi_t(\bar{\theta})}{\pi_t(\underline{\theta})} \xrightarrow[a.s.]{} \left(\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{(\lambda+(1-\lambda)\underline{\rho})}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\lambda)(1-\underline{\rho})}\right)^t \frac{\pi_0(\bar{\theta})}{\pi_0(\underline{\theta})}$$

If a low type agent is to learn her type correctly then we need the above belief ratio to converge to 0 almost surely or equivalently we need the parameters $\underline{\rho}$, $\bar{\rho}$ and λ be such that

$$\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{(\lambda+(1-\lambda)\underline{\rho})}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\lambda)(1-\underline{\rho})} < 1, \tag{6}$$

There are no parameters $0.5 < \underline{\rho} < \bar{\rho} < 1$ and $0 < \lambda < 1$ for which the above inequality is not violated[3], causing a low type agent to always incorrectly believe she is of a higher type; thence, the ability of a low type agent to learn her type is compromised by her hindsight.

$$\frac{\partial f(\lambda,\bar{\rho},\underline{\rho})}{\partial\lambda} = (\bar{\rho}-1)\left(\frac{\bar{\rho}-1}{\underline{\rho}-1}\right)^{(\lambda-1)(\bar{\rho}-1)}\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\lambda(-\bar{\rho})+\lambda+\bar{\rho}}\left(\log\left(\frac{\bar{\rho}-1}{\underline{\rho}-1}\right)-\log\left(\frac{\bar{\rho}}{\underline{\rho}}\right)\right).$$

The latter is positive everywhere and therefore increasing monotonically over the domain, which implies that for any fixed choice of $\bar{\rho},\underline{\rho}$, $f(\lambda,\bar{\rho},\underline{\rho})$ is minimized at $\lambda^* = 0$. Now, we seek values of $\bar{\rho},\underline{\rho}$ that minimize

$$f(\lambda^*,\bar{\rho},\underline{\rho}) = \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\bar{\rho}}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\bar{\rho})}, \text{ to satisfy the condition } f(\lambda^*,\bar{\rho},\underline{\rho}) < 1.$$

Given our assumption of positive $\bar{\rho}$ and $\underline{\rho}$, we have

$$\frac{\partial f(\lambda^*,\bar{\rho},\underline{\rho})}{\partial\underline{\rho}} = \frac{(\bar{\rho}-1)\bar{\rho}(\bar{\rho}-\underline{\rho})\left(\frac{\bar{\rho}-1}{\underline{\rho}-1}\right)^{-\bar{\rho}}\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\bar{\rho}-1}}{(\underline{\rho}-1)^2\underline{\rho}^2} < 0,$$

which is monotonic decreasing in $\underline{\rho}$, which implies that for any fixed $\bar{\rho}$, $f(\lambda^*,\bar{\rho},\underline{\rho})$ is minimized by maximal $\underline{\rho}$. Hence, over the domain of f,

$$f(\lambda,\bar{\rho},\underline{\rho}) \geq f(\lambda^*,\bar{\rho},\underline{\rho}) \geq \bar{\rho}\left(\frac{\bar{\rho}}{\bar{\rho}}\right)^{\bar{\rho}}\left(\frac{1-\bar{\rho}}{1-\bar{\rho}}\right)^{(1-\bar{\rho})} = 1 \implies \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{(1-\lambda)\bar{\rho}+\lambda}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\lambda)(1-\bar{\rho})} \geq 1,$$

and hence a high type agent always learns her type correctly.

3. In more details (6) is satisfied, only if

$$\lambda < \left(\frac{\log(\frac{\underline{\rho}-1}{\bar{\rho}-1})}{\log(\frac{\underline{\rho}-1}{\bar{\rho}-1})+\log(\frac{\bar{\rho}}{\underline{\rho}})} - 2\underline{\rho}\right)(1-\underline{\rho})^{-1}, \text{ which can never be since } \lambda > 0,\ 2\underline{\rho} > 1 \text{ and}$$

$$\frac{\log(\frac{\underline{\rho}-1}{\bar{\rho}-1})}{\log(\frac{\underline{\rho}-1}{\bar{\rho}-1})+\log(\frac{\bar{\rho}}{\underline{\rho}})} < 1, \text{ for all } 0.5 < \underline{\rho} < \bar{\rho} < 1.$$

A.2 Proof of Theorem 2

Let $\mathcal{G} = ([n], \mathcal{E}, T)$, $[n] = \{1, \ldots, n\}$ be the weighted graph of social network influences, with $T = [T_{ij}]_{i,j\in[n]}$ being the matrix of influence weight and $T_{i,j} \geq 0$ being the influence of agent j on i, and $\mathcal{E} \subset [n] \times [n]$ is the set of edges corresponding to the non-zero entries of matrix T. We assume T to be a doubly stochastic matrix. At every round t, let $\pi_{i,t}(\bar{\theta})$ and $\pi_{i,t}(\underline{\theta})$ be the beliefs of agent i on the false and true states after communicating with her neighbors and experimenting with her environment at the particular round. The neighborhood of agent i is the set of all agents who have a strictly positive influence on her: $\mathcal{N}(i) = \{j : T_{i,j} > 0\}$. At all epochs t, define $\hat{\pi}_{i,t}(\bar{\theta})/\hat{\pi}_{i,t}(\underline{\theta})$ to be the updated belief of agent i after communicating with her neighbors but before experimenting with her environment, i.e. before observing $\bar{s}_{i,t}$ or perceiving $\bar{r}_{i,t}$ for a high type society (we consider the case of high-type agents and the low-type case is similar). The log-linear belief updates prescribe the updated belief of the agent before experimenting with her environment $\hat{\pi}_i(\bar{\theta})/\hat{\pi}(\underline{\theta})$ in terms of the reported beliefs of her neighbors:

$$\frac{\hat{\pi}_{i,t}(\bar{\theta})}{\hat{\pi}_{i,t}(\underline{\theta})} = \prod_{j\in\mathcal{N}(i)} \left(\frac{\pi_{j,t-1}(\bar{\theta})}{\pi_{j,t-1}(\underline{\theta})}\right)^{T_{ij}}. \tag{7}$$

Consider the agent in a society of high type agents, after communicating with her neighbors the agent engages with her environment and (4) still holds true in her updating of her beliefs following her observation of the private signal $\bar{s}_{i,t}$ and after her true state $\omega_{i,t}$ us revealed to her, subjecting her to a hind sight bias. Following (5) and (7) we can write:

$$\begin{aligned}\frac{\pi_{i,t}(\bar{\theta})}{\pi_{i,t}(\underline{\theta})} &= \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_{i,t}=\omega_{i,t})} \left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_{i,t}\neq\omega_{i,t})} \frac{\hat{\pi}_{i,t}(\bar{\theta})}{\hat{\pi}_{i,t}(\underline{\theta})} \\ &= \left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_{i,t}=\omega_{i,t})} \left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{\mathbb{1}(\bar{r}_{i,t}\neq\omega_{i,t})} \prod_{j\in\mathcal{N}(i)} \left(\frac{\pi_{j,t-1}(\bar{\theta})}{\pi_{j,t-1}(\underline{\theta})}\right)^{T_{ij}}\end{aligned} \tag{8}$$

Define the log of the belief ratios for every agent as $\phi_{i,t} = \log(\pi_{i,t}(\bar{\theta})/\pi_{i,t}(\underline{\theta}))$, and concatenate the log of the belief ratios for all agents into a single vector ϕ_t defined as $\phi_t = (\phi_{1,t}, \ldots, \phi_{n,t})$. Similarly, define the log likelihood ratio of the perceived signals as

$$\bar{\ell}_{i,t} = \mathbb{1}(\bar{r}_{i,t} = \omega_{i,t}) \log(\bar{\rho}/\underline{\rho}) + \mathbb{1}(\bar{r}_{i,t} \neq \omega_{i,t}) \log((1-\bar{\rho})/(1-\underline{\rho})) \tag{9}$$

and let $\bar{\ell}_t = (\bar{\ell}_{1,t}, \ldots, \bar{\ell}_{n,t})$ be their concatenation. Taking logarithms of both sides in (8) and using the concatenations ϕ_t and $\bar{\ell}_t$, we obtain the following linearized and vectorized update:

$$\phi_t = T\phi_{t-1} + \bar{\ell}_t = \sum_{\tau=1}^{t} T^{t-\tau}\bar{\ell}_\tau + \phi_0.$$

Next note that for a doubly stochastic matrix $\lim_{\tau\to\infty} T^\tau = (1/n)\mathbb{1}\mathbb{1}^T$. Hence, we can invoke the Cesàro mean together with the strong law applied to the i.i.d. sequence $\{\bar{\ell}_t, t \in \mathbb{N}_0\}$ to conclude that (Rahimian and Jababaie, 2015),

$$\frac{1}{t}\phi_t \to \frac{1}{n}\mathbb{1}\mathbb{1}^T\mathbb{E}\{\bar{\ell}_t\}$$

or equivalently that $\phi_{i,t} \to (t/n)\sum_{i=1}^{n}\mathbb{E}\{\bar{\ell}_{i,t}\}$ with probability one, as $t \to \infty$. Next using (3) and (9) we can write

$$\mathbb{E}\{\bar{\ell}_{i,t}\} = (\lambda_i + (1-\lambda_i)\bar{\rho})\log\left(\bar{\rho}/\underline{\rho}\right) + (1-\lambda_i)(1-\bar{\rho})\log\left((1-\bar{\rho})/(1-\underline{\rho})\right),$$

and the conclusion for high type agents follows because $\mathbb{E}\{\bar{\ell}_{i,t}\} > 0$ for all agents $i \in [n]$ with $0.5 < \underline{\rho} < \bar{\rho} < 1$ and $0 < \lambda_i < 1$; and in particular: $(1/n)\sum_{i=1}^{n}\mathbb{E}\{\bar{\ell}_{i,t}\} > 0$ implying that $\phi_{i,t} \to +\infty$ leading to all high type agents to learn their true type correctly. The situation is distinctively different in a society of low type agents. Repeating the above analysis using $\underline{r}_{i,t}$ instead of $\bar{r}_{i,t}$ we obtain that for all low type agents $i \in [n]$: $\phi_{i,t} \to (t/n)\sum_{i=1}^{n}\mathbb{E}\{\underline{\ell}_{i,t}\}$ with probability one, as $t \to \infty$, where:

$$\mathbb{E}\{\underline{\ell}_{i,t}\} = (\lambda_i + (1-\lambda_i)\underline{\rho})\log\left(\bar{\rho}/\underline{\rho}\right) + (1-\lambda_i)(1-\underline{\rho})\log\left((1-\bar{\rho})/(1-\underline{\rho})\right).$$

Defining $\lambda_{avg} = (1/n)\sum_{i=1}^{n}\lambda_i$ we get that if

$$(\lambda_{avg} + (1-\lambda_{avg})\underline{\rho})\log\left(\bar{\rho}/\underline{\rho}\right) + (1-\lambda_{avg})(1-\underline{\rho})\log\left((1-\bar{\rho})/(1-\underline{\rho})\right) > 0$$

or equivalently if

$$\left(\frac{\bar{\rho}}{\underline{\rho}}\right)^{(\lambda_{avg}+(1-\lambda_{avg})\underline{\rho})}\left(\frac{1-\bar{\rho}}{1-\underline{\rho}}\right)^{(1-\lambda_{avg})(1-\underline{\rho})} > 1, \tag{10}$$

then the society of high type agents reach a consensus on a false state; incorrectly believing themselves to be of a higher type. An analysis identical to the one in footnote 3 in Appendix A.1 reveals that (10) is in fact always true for any choice of parameters $0 < \lambda_{avg} < 1$, and $0.5 < \underline{\rho} < \bar{\rho} < 1$.

References

Hal R Arkes. The consequences of the hindsight bias in medical decision making. *Current Directions in Psychological Science*, 22(5):356–360, 2013.

Leonard Berlin. Statement of leonard berlin, m.d., to the u.s. senate committee on health, education labor and pensions: Mammography quality standards act reauthorization. 2003. Available online.

Daniel M Bernstein, Edgar Erdfelder, Andrew N Meltzoff, William Peria, and Geoffrey R Loftus. Hindsight bias from 3 to 95 years of age. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2):378, 2011.

Daniel M Bernstein, Alexander Maurice Wilson, Nicole LM Pernat, and Louise R Meilleur. Auditory hindsight bias. *Psychonomic bulletin and review*, 19(4):588–593, 2012.

Daniel M Bernstein, André Aßfalg, R Kumar, and Rakefet Ackerman. Looking backward and forward on hindsight bias. *Handbook of metamemory*, (November), 2015. ISSN 00010782. doi: 10.1145/1516046.1516071.

Bruno Biais and Martin Weber. Hindsight bias, risk perception, and investment performance. *Management Science*, 55(6):1018–1029, 2009.

Hartmut Blank, Steffen Nestler, Gernot von Collani, and Volkhard Fischer. How many hindsight biases are there? *Cognition*, 106(3):1408–1440, 2008.

Colin Camerer, George Loewenstein, and Martin Weber. The curse of knowledge in economic settings: An experimental analysis. *The Journal of Political Economy*, pages 1232–1254, 1989.

A Coolin, DM Bernstein, AE Thornton, and WL Thornton. Inhibition and episodic memory impact age differences in hindsight bias. *Experimental Aging Research*, 40:357–374, 2014.

Alisha Coolin, Edgar Erdfelder, Daniel M Bernstein, Allen E Thornton, and Wendy Loken Thornton. Inhibitory control underlies individual differences in older adults' hindsight bias. *Psychology and aging*, 31(3):224, 2016.

David Danz, Dorothea Kübler, and Julia Schmid. On the failure of hindsight-biased principals to delegate optimally. *Management Science*, 2015.

Nicholas Epley and Thomas Gilovich. The anchoring-and-adjustment heuristic why the adjustments are insufficient. *Psychological science*, 17(4):311–318, 2006.

Edgar Erdfelder and Axel Buchner. Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2):387, 1998.

Edgar Erdfelder, Martin Brandt, and Arndt Bröder. Recollection biases in hindsight judgments. *Social Cognition*, 25(1):114, 2007.

K Fiedler, M von Sydow, MW Eysenck, and D Groome. Heuristics and biases: beyond tversky and kahneman's (1974) judgment under uncertainty. *Cognitive Psychology: Revisiting the Classical Studies, eds Eysenck MW, Groome D., editors.(Los Angeles, US: Sage*, pages 146–161, 2015.

Baruch Fischhoff. Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human perception and performance*, 1(3):288, 1975.

Baruch Fischhoff. The perceived informativeness of factual information. Technical report, DTIC Document, 1976.

Baruch Fischhoff. For those condemned to study the past: Heuristics and biases in hindsight. *D. Kahneman, P. Slovic, A. Tversky, eds. Judgment Under Uncertainty: Heuristics and Biases.*, page 332–351, 1982.

Baruch Fischhoff. An early history of hindsight research. 2007.

Baruch Fischhoff and Ruth Beyth. I knew it would happen: Remembered probabilities of once—future things. *Organizational Behavior and Human Performance*, 13(1):1–16, 1975.

Erin M Harley. Hindsight bias in legal decision making. *Social Cognition*, 25(1):48, 2007.

Erin M Harley, Keri A Carlsen, and Geoffrey R Loftus. The" saw-it-all-along" effect: demonstrations of visual hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5):960, 2004.

Wolfgang Hell, Gerd Gigerenzer, Siegfried Gauggel, Maria Mall, and Michael Müller. Hindsight bias: An interaction of automatic and motivational factors? *Memory & Cognition*, 16(6):533–538, 1988.

Ulrich Hoffrage, Ralph Hertwig, and Gerd Gigerenzer. Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3):566, 2000.

Mark R Leary. Hindsight distortion and the 1980 presidential election. *Personality and Social Psychology Bulletin*, 8(2):257–263, 1982.

Therese A Louie. Decision makers' hindsight bias after receiving favorable and unfavorable feedback. *Journal of Applied Psychology*, 84(1):29, 1999.

Kristóf Madarász. Information projection: Model and applications. *The Review of Economic Studies*, 2011.

Melvin M Mark and Steven Mellor. Effect of self-relevance of an event on hindsight bias: The foreseeability of a layoff. *Journal of Applied Psychology*, 76(4):569, 1991.

Patrick A Müller and Dagmar Stahlberg. The role of surprise in hindsight bias: A metacognitive model of reduced and reversed hindsight bias. *Social Cognition*, 25(1):165, 2007.

Steffen Nestler, Hartmut Blank, and Gernot von Collani. Hindsight bias doesn't always come easy: causal models, cognitive effort, and creeping determinism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(5):1043, 2008.

Mark V Pezzo. Hindsight bias: A primer for motivational researchers. *Social and Personality Psychology Compass*, 5(9):665–678, 2011.

Rüdiger F Pohl and Wolfgang Hell. No reduction in hindsight bias after complete information and repeated testing. *Organizational Behavior and Human Decision Processes*, 67(1): 49–58, 1996.

Rüdiger F Pohl, Michael Bender, and Gregor Lachmann. Hindsight bias around the world. *Experimental psychology*, 49(4):270, 2002.

Rüdiger F Pohl, Ute J Bayen, and Claudia Martin. A multiprocess account of hindsight bias in children. *Developmental Psychology*, 46(5):1268, 2010.

M. A. Rahimian and A. Jababaie. Learning without recall: A case for log-linear learning. *IFAC-PapersOnLine*, 48(22):46–51, 2015.

M. A. Rahimian and A. Jadbabaie. Learning without recall in directed circles and rooted trees. In *2015 American Control Conference (ACC)*, pages 4222–4227, July 2015. doi: 10.1109/ACC.2015.7171992.

M. A. Rahimian and A. Jadbabaie. Bayesian heuristics for group decisions. *arXiv:1611.01006*, November 2016.

Neal J Roese and Kathleen D Vohs. Hindsight bias. *Perspectives on Psychological Science*, 7(5):411–426, 2012.

Stefan Schwarz and Dagmar Stahlberg. Strength of hindsight bias as a consequence of meta-cognitions. *Memory*, 11(4-5):395–410, 2003.

Orit E Tykocinski and Noa Steinberg. Coping with disappointing outcomes: Retroactive pessimism and motivated inhibition of counterfactuals. *Journal of Experimental Social Psychology*, 41(5):551–558, 2005.

Gordon Wood. The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, 4(2):345, 1978.