Multi-Objective Decision Making

Synthesis Lectures on Artificial Intelligence and Machine Learning

Editors
Ronald J. Brachman, ***Yahoo! Labs***
Peter Stone, ***University of Texas at Austin***

Multi-Objective Decision Making
Diederik M. Roijers and Shimon Whiteson
2017

Lifelong Machine Learning
Zhiyuan Chen and Bing Liu
2016

Statistical Relational Artificial Intelligence: Logic, Probability, and Computation
Luc De Raedt, Kristian Kersting, Sriraam Natarajan, and David Poole
2016

Representing and Reasoning with Qualitative Preferences: Tools and Applications
Ganesh Ram Santhanam, Samik Basu, and Vasant Honavar
2016

Metric Learning
Aurélien Bellet, Amaury Habrard, and MarcSebban
2015

Graph-Based Semi-Supervised Learning
Amarnag Subramanya and Partha Pratim Talukdar
2014

Robot Learning from Human Teachers
Sonia Chernova and Andrea L. Thomaz
2014

General Game Playing
Michael Genesereth and Michael Thielscher
2014

Judgment Aggregation: A Primer
Davide Grossi and Gabriella Pigozzi
2014

An Introduction to Constraint-Based Temporal Reasoning
Roman Barták, Robert A. Morris, and K. Brent Venable
2014

Reasoning with Probabilistic and Deterministic Graphical Models: Exact Algorithms
Rina Dechter
2013

Introduction to Intelligent Systems in Traffic and Transportation
Ana L.C. Bazzan and Franziska Klügl
2013

A Concise Introduction to Models and Methods for Automated Planning
Hector Geffner and Blai Bonet
2013

Essential Principles for Autonomous Robotics
Henry Hexmoor
2013

Case-Based Reasoning: A Concise Introduction
Beatriz López
2013

Answer Set Solving in Practice
Martin Gebser, Roland Kaminski, Benjamin Kaufmann, and Torsten Schaub
2012

Planning with Markov Decision Processes: An AI Perspective
Mausam and Andrey Kolobov
2012

Active Learning
Burr Settles
2012

Computational Aspects of Cooperative Game Theory
Georgios Chalkiadakis, Edith Elkind, and Michael Wooldridge
2011

Representations and Techniques for 3D Object Recognition and Scene Interpretation
Derek Hoiem and Silvio Savarese
2011

A Short Introduction to Preferences: Between Artificial Intelligence and Social Choice
Francesca Rossi, Kristen Brent Venable, and Toby Walsh
2011

Human Computation
Edith Law and Luis von Ahn
2011

Trading Agents
Michael P. Wellman
2011

Visual Object Recognition
Kristen Grauman and Bastian Leibe
2011

Learning with Support Vector Machines
Colin Campbell and Yiming Ying
2011

Algorithms for Reinforcement Learning
Csaba Szepesvári
2010

Data Integration: The Relational Logic Approach
Michael Genesereth
2010

Markov Logic: An Interface Layer for Artificial Intelligence
Pedro Domingos and Daniel Lowd
2009

Introduction to Semi-Supervised Learning
XiaojinZhu and Andrew B.Goldberg
2009

Action Programming Languages
Michael Thielscher
2008

Representation Discovery using Harmonic Analysis
Sridhar Mahadevan
2008

Essentials of Game Theory: A Concise Multidisciplinary Introduction
Kevin Leyton-Brown and Yoav Shoham
2008

A Concise Introduction to Multiagent Systems and Distributed Artificial Intelligence
Nikos Vlassis
2007

Intelligent Autonomous Robotics: A Robot Soccer Case Study
Peter Stone
2007

© Springer Nature Switzerland AG 2022

Reprint of original edition © Morgan & Claypool 2017

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

Multi-Objective Decision Making

Diederik M. Roijers and Shimon Whiteson

ISBN: 978-3-031-00448-3 paperback
ISBN: 978-3-031-01576-2 ebook

DOI 10.1007/978-3-031-01576-2

A Publication in the Springer series
SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Lecture #34
Series Editors: Ronald J. Brachman, *Yahoo! Labs*
Peter Stone, *University of Texas at Austin*
Series ISSN
Print 1939-4608 Electronic 1939-4616

Multi-Objective Decision Making

Diederik M. Roijers
University of Oxford
Vrije Universiteit Brussel

Shimon Whiteson
University of Oxford

SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING #34

ABSTRACT

Many real-world decision problems have multiple objectives. For example, when choosing a medical treatment plan, we want to maximize the efficacy of the treatment, but also minimize the side effects. These objectives typically conflict, e.g., we can often increase the efficacy of the treatment, but at the cost of more severe side effects. In this book, we outline how to deal with multiple objectives in decision-theoretic planning and reinforcement learning algorithms. To illustrate this, we employ the popular problem classes of *multi-objective Markov decision processes* (MOMDPs) and *multi-objective coordination graphs* (MO-CoGs).

First, we discuss different use cases for multi-objective decision making, and why they often necessitate explicitly multi-objective algorithms. We advocate a *utility-based* approach to multi-objective decision making, i.e., that what constitutes an optimal solution to a multi-objective decision problem should be derived from the available information about user utility. We show how different assumptions about user utility and what types of policies are allowed lead to different solution concepts, which we outline in a taxonomy of multi-objective decision problems.

Second, we show how to create new methods for multi-objective decision making using existing single-objective methods as a basis. Focusing on planning, we describe two ways to creating multi-objective algorithms: in the inner loop approach, the inner workings of a single-objective method are adapted to work with multi-objective solution concepts; in the outer loop approach, a wrapper is created around a single-objective method that solves the multi-objective problem as a series of single-objective problems. After discussing the creation of such methods for the planning setting, we discuss how these approaches apply to the learning setting.

Next, we discuss three promising application domains for multi-objective decision making algorithms: energy, health, and infrastructure and transportation. Finally, we conclude by outlining important open problems and promising future directions.

KEYWORDS

artificial intelligence, decision theory, decision support systems, probabilistic planning, multi-agent systems, multi-objective optimization, machine learning

Contents

Preface xiii

Acknowledgments xv

Table of Abbreviations xvii

1 Introduction 1

1.1 Motivation 2

1.2 Utility-based Approach 6

2 Multi-Objective Decision Problems 9

2.1 Multiple Objectives 9

2.2 Multi-Objective Coordination 10

2.2.1 Single-Objective Coordination Graphs 11

2.2.2 Multi-Objective Coordination Graphs 12

2.3 Multi-Objective Markov Decision Processes 13

2.3.1 Single-Objective Markov Decision Processes 14

2.3.2 Multi-Objective Markov Decision Processes 16

3 Taxonomy 19

3.1 Critical Factors 19

3.1.1 Single vs. Multiple Policies 19

3.1.2 Linear vs. Monotonically Increasing Scalarization Functions 20

3.1.3 Deterministic vs. Stochastic Policies 21

3.2 Solution Concepts 22

3.2.1 Case #1: Linear Scalarization and a Single Policy 24

3.2.2 Case #2: Linear Scalarization and Multiple Policies 25

3.2.3 Case #3: Monotonically Increasing Scalarization and a Single Deterministic Policy 27

3.2.4 Case #4: Monotonically Increasing Scalarization and a Single Stochastic Policy 28

3.2.5 Case #5: Monotonically Increasing Scalarization and Multiple Deterministic Policies 28

3.2.6 Case #6: Monotonically Increasing Scalarization and Multiple Stochastic Policies . . . 31
3.3 Implications for MO-CoGs . . . 32
3.4 Approximate Solution Concepts . . . 33
3.5 Beyond the Taxonomy . . . 34

4 Inner Loop Planning . . . 37
4.1 Inner Loop Approach . . . 37
4.1.1 A Simple MO-CoG . . . 37
4.1.2 Finding a PCS . . . 39
4.1.3 Finding a CCS . . . 41
4.1.4 Design Considerations . . . 43
4.2 Inner Loop Planning for MO-CoGs . . . 43
4.2.1 Variable Elimination . . . 43
4.2.2 Transforming the MO-CoG . . . 46
4.2.3 Multi-Objective Variable Elimination . . . 47
4.2.4 Comparing PMOVE and CMOVE . . . 54
4.3 Inner Loop Planning for MOMDPs . . . 58
4.3.1 Value Iteration . . . 58
4.3.2 Multi-Objective Value Iteration . . . 59
4.3.3 Pareto vs. Convex Value Iteration . . . 59

5 Outer Loop Planning . . . 63
5.1 Outer Loop Approach . . . 63
5.2 Scalarized Value Functions . . . 65
5.2.1 The Relationship with POMDPs . . . 67
5.3 Optimistic Linear Support . . . 67
5.4 Analysis . . . 75
5.5 Approximate Single-Objective Solvers . . . 76
5.6 Value Reuse . . . 80
5.7 Comparing an Inner and Outer Loop Method . . . 82
5.7.1 Theoretical Comparison . . . 82
5.7.2 Empirical Comparison . . . 83
5.8 Outer Loop Methods for PCS Planning . . . 84

6 Learning . . . 87
6.1 Offline MORL . . . 88
6.2 Online MORL . . . 89

7 Applications **91**
7.1 Energy 91
7.2 Health 92
7.3 Infrastructure and Transportation 92

8 Conclusions and Future Work **95**
8.1 Conclusions 95
8.2 Future Work 97
8.2.1 Scalarization of Expectation vs. Expectation of Scalarization 97
8.2.2 Other Decision Problems 99
8.2.3 Users in the Loop 99

Bibliography **101**

Authors' Biographies **111**

Preface

Many real-world decision problems have multiple, possibly conflicting, objectives. For example, an autonomous vehicle typically wants to minimize both travel time and fuel costs, while maximizing safety; when seeking medical treatment, we want to maximize the probability of being cured, but minimize the severity of the side-effects, etcetera.

Although interest in multi-objective decision making has grown in recent years, the majority of decision-theoretic research still assumes only a single objective. In this book, we argue that multi-objective methods are underrepresented and present three scenarios to justify the need for explicitly multi-objective approaches. Key to these scenarios is that, although the utility the user derives from a policy—which is what we ultimately aim to optimize—is scalar, it is sometimes impossible, undesirable, or infeasible to formulate the problem as single-objective at the moment when the policies need to be planned or learned. We also present the case for a *utility-based* view of multi-objective decision making, i.e., that the appropriate multi-objective solution concept should be derived from what we know about the user's utility function.

This book is based on our research activities over the years. In particular, the survey we wrote together with Peter Vamplew and Richard Dazeley [Roijers et al., 2013a] forms the basis of how we organize concepts in multi-objective decision making. Furthermore, we use insights from our work on multi-objective planning over the years, particularly in the context of the PhD research of the first author [Roijers, 2016]. Another important source for writing this book were the lectures we gave on the topic at the University of Amsterdam, and the tutorials we did at the IJCAI-2015 and ICAPS-2016 conferences, as well as the EASSS-2016 summer school.

Aim and Readership This book aims to provide a structured introduction to the field of multi-objective decision making, and to make the differences with single-objective decision theory clear. We hope that, after reading this book, the reader will be equipped to conduct research in multi-objective decision-theory or apply multi-objective methods in practice.

We expect our readers to have a basic understanding of decision theory, at a graduate or undergraduate level. In order to remain accessible to a wide range of readers, we provide intuitive explanations and examples of key concepts before formalizing them. In some cases, we omit detailed proofs of theorems in order to better focus on the intuition behind and implications of these theorems. In such cases, we provide references to the detailed proofs.

Outline This book is structured as follows. In Chapter 1, we motivate multi-objective decision making by providing examples of multi-objective decision problems and scenarios that require explicitly multi-objective solution methods. In Chapter 2, we introduce two popular classes of decision problems that we use throughout the book to illustrate specific algorithms and general

theoretical results. In Chapter 3, we present a taxonomy of solution concepts for multi-objective decision problems. Using this taxonomy, we discuss different solution methods. First, we assume that the model of the environment is known to the agents, leading to a *planning* setting. In Chapters 4 and 5, we discuss two different approaches for finding a *coverage set* using planning algorithms. In Chapter 6, we remove the assumption that the agents are given a model of the environment, and consider cases where they must learn about the environment through interaction. Finally, we discuss several illustrating applications in Chapter 6, followed by conclusions and future work in Chapter 8.

Diederik M. Roijers and Shimon Whiteson
April 2017

Acknowledgments

This book is based on our research on multi-objective decision making over the years. During this research, we collaborated with people whose input has been essential to our understanding of the field. We would like to thank several of them explicitly.

Together with Peter Vamplew and Richard Dazeley we wrote our 2013 survey article on multi-objective sequential decision making. The discussions we had about the nature of multi-objective decision problems were vital in shaping our ideas about this field, and lay the foundation for how we view multi-objective decision problems.

In the past few years, one of our main collaborators (and Diederik's other PhD supervisor), has been Frans A. Oliehoek. Together, we developed many algorithms for multi-objective decision making, including the CMOVE and OLS algorithms that we discuss in Chapters 4 and 5. Frans's vast expertise on partially observable decision problems and limitless capacity for generating new ideas have been invaluable to our work in the field of multi-objective decision making.

Together with Joris Scharpff, Matthijs Spaan, and Mathijs de Weerdt, we worked on the traffic network maintenance planning problem (which we discuss in Section 7.3), and in this context improved upon the original OLS algorithm (Chapter 5). We enjoyed this productive collaboration.

We would also like to thank our other past and present co-authors and collaborators who we have worked with on multi-objective decision making problems: Alexander Ihler, João Messias, Maarten van Someren, Chiel Kooijman, Maarten Inja, Maarten de Waard, Luisa Zintgraf, Timon Kanters, Philipp Beau, Richard Pronk, Carla Groenland, Elise van der Pol, Joost van Doorn, Daan Odijk, Maarten de Rijke, Ayumi Igarashi, Hossam Mossalam, and Yannis Assael.

Finally, we would like to thank several people with whom we had interesting discussions about multi-objective decision making over the years: Ann Nowé, Kristof van Moffaert, Tim Brys, Abdel-Illah Mouaddib, Paul Weng, Grégory Bonnet, Rina Dechter, Radu Marinescu, Shlomo Zilberstein, Kyle Wray, Patrice Perny, Paolo Viappiani, Pascal Poupart, Max Welling, Karl Tuyls, Francesco Delle Fave, Joris Mooij, Reyhan Aydoğan, and many others.

Diederik M. Roijers and Shimon Whiteson
April 2017

Table of Abbreviations

Abbreviation	Full Name	Location
AOLS	approximate optimistic linear support	Algorithm 5.10, Section 5.5
CCS	convex coverage set	Definition 3.7, Section 3.2.2
CH	convex hull	Definition 3.6, Section 3.2.2
CHVI	convex hull value iteration	Section 4.3.2
CLS	Cheng's linear support	Section 5.3
CMOVE	multi-objective variable elimination	Section 4.2.3
CoG	coordination graph	Definition 2.4, Section 2.2.1
CS	coverage set	Definition 3.5, Section 3.2
f	scalarization function	Definition 1.1, Section 1.1
MDP	Markov decision process	Definition 2.6, Section 2.3.1
MO-CoG	multi-objective coordination graph	Definition 2.5, Section 2.2.2
MODP	multi-objective decision problem	Definition 2.2, Section 2.1
MOMDP	multi-objective Markov decision process	Definition 2.8, Section 2.3.2
MORL	multi-objective reinforcement learning	Chapter 6
MOVE	multi-objective variable elimination	Algorithm 4.5, Section 4.2.3
MOVI	multi-objective value iteration	Section 4.3.2
OLS	optimistic linear support	Algorithm 5.8, Section 5.3
OLS-R	optimistic linear support with reuse	Algorithm 5.11, Section 5.6
PMOVI	Pareto multi-objective value iteration	Section 4.3.2
PCS	Pareto coverage set	Definition 3.11, Section 3.2.4
PMOVE	Pareto multi-objective variable elimination	Section 4.2.3
POMDP	partially observable Markov decision process	Section 5.2.1
PF	Pareto front	Definition 3.10, Section 3.2.4
SODP	single-objective decision problem	Definition 2.1, Section 2.1
U	undominated set	Definition 3.4, Section 3.2
VE	variable elimination	Algorithm 4.4, Section 4.2.1
VELS	variable elimination linear support	Section 5.7
VI	value iteration	Section 4.3.1
$\mathbf{V}^{\pi}$	value vector of a policy π	Definition 2.2, Section 2.1
Π	a set of allowed policies	Definition 2.1, Section 2.1
$\succ_P$	Pareto dominance relation	Definition 3.3, Section 3.1.2