Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 178 (2018) 1–35

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Ambiguous partially observable Markov decision processes: Structural results and applications ☆

Soroush Saghafian

Harvard Kennedy School, Harvard University, Cambridge, MA, United States of America

Received 1 September 2016; final version received 2 March 2018; accepted 15 August 2018
Available online 20 August 2018

Abstract

Markov Decision Processes (MDPs) have been widely used as invaluable tools in dynamic decision-making, which is a central concern for economic agents operating at both the micro and macro levels. Often the decision maker's information about the state is incomplete; hence, the generalization to Partially Observable MDPs (POMDPs). Unfortunately, POMDPs may require a large state and/or action space, creating the well-known "curse of dimensionality." However, recent computational contributions and blindingly fast computers have helped to dispel this curse. This paper introduces and addresses a second curse termed "curse of ambiguity," which refers to the fact that the exact transition probabilities are often hard to quantify, and are rather ambiguous. For instance, for a monetary authority concerned with dynamically setting the inflation rate so as to control the unemployment, the dynamics of unemployment rate under any given inflation rate is often ambiguous. Similarly, in worker-job matching, the dynamics of worker-job match/proficiency level is typically ambiguous. This paper addresses the "curse of ambiguity" by developing a generalization of POMDPs termed Ambiguous POMDPs (APOMDPs), which not only allows the decision maker to take into account imperfect state information, but also tackles the inevitable ambiguity with respect to the correct probabilistic model of transitions.

Importantly, this paper extends various structural results from POMDPs to APOMDPs. These results enable the decision maker to make robust decisions. Robustness is achieved by using α-maximin expected utility (α-MEU), which (a) differentiates between ambiguity and ambiguity attitude, (b) avoids the over conservativeness of traditional maximin approaches, and (c) is found to be suitable in laboratory exper-

☆ The author is grateful (in no particular order) to Tomasz Strzalecki (Harvard), Richard Zeckhauser (Harvard), Maciej Kotowski (Harvard), Michael Veatch (Gordon College), and Hao Zhang (University of British Columbia) for valuable suggestions, comments, and discussions which helped to improve this paper. The author also thanks the editors, and the anonymous referee for their various helpful comments. This work was partially supported by the National Science Foundation under the Award Number CMMI-1562645.

E-mail address: soroush_saghafian@hks.harvard.edu.

https://doi.org/10.1016/j.jet.2018.08.006

iments in various choice behaviors including those in portfolio selection. The structural results provided also help to handle the "curse of dimensionality," since they significantly simplify the search for an optimal policy. The analysis also identifies a performance guarantee for the proposed approach by developing a bound for its maximum reward loss due to model ambiguity.

JEL classification: C61; D81; D83; D84

Keywords: POMDP; Unknown probabilities; Model ambiguity; Structural results; Control-limit policies

1. Introduction

A critical factor for economic agents operating at both the micro and macro levels is decision-making in dynamic environments. Markov Decision Processes (MDPs) have been widely used for dynamic decision-making in such environments when two main assumptions hold: (1) the state of the system is completely known/observable at each decision epoch, and (2) the (Markovian) state transitions can be probabilistically defined. Partially Observable MDPs (POMDPs) extend MDPs by relaxing the first assumption: POMDPs consider the case where the system's state is not completely observable but there exist observations/signals which yield probabilistic beliefs about the hidden state, if the second assumption above holds. However, the second assumption is unrealistic in most applications, and significantly limits the applicability of POMDPs in real-world settings.

In such settings, one might have access to some data, and to develop a POMDP, must first estimate core state and observation transition probabilities. This often comes with estimation errors and leaves the decision maker with inevitable model misspecification/ambiguity. We refer to this challenge as the *curse of ambiguity*, and address it by relaxing assumption (2) above. Hence, this paper extends POMDPs to a new dynamic decision-making framework that allows the decision maker to consider both imperfect state information and ambiguity with respect to the correct probabilistic model. We term this new framework as *Ambiguous POMDP (APOMDP).*[1]

To address the curse of ambiguity, we assume that the decision maker simultaneously faces (a) non-probabilistic ambiguity (a.k.a. *Knightian uncertainty*) about the true model, and (b) probabilistic uncertainty or risk given the true model.[2] As Arrow (1951) (p. 418) states: *"There are two types of uncertainty: one as to the hypothesis, which is expressed by saying that the hypothesis is known to belong to a certain class or model, and one as to the future events or observations given the hypothesis, which is expressed by a probability distribution."* Indeed, in this paper's framework, the decision maker is faced with Knightian uncertainty regarding the true model, while under each potential model, he has a certain probabilistic understanding of how observations and the core system state evolve over time. This draws a line between *ambiguity* (lack of

[1] To highlight the importance of considering the "curse of ambiguity," we note that the work of Savage and the applied statistical decision theory literature, which has been embraced by rational economists, suggests that probabilities should simply be estimated and that there should be no discount for ambiguity. However, the literature starting with Knight, and then dealing with the Ellsberg Paradox, and exploding on the scene with the work of Tversky and Kahneman recognizes that ambiguity plays an essential role in human decision-making.

[2] See, e.g., Stoy (2011), for an axiomatic treatment of statistical decision-making under these conditions.