Spring 5-16-2020

A Data-Driven Framework for Decision Making Under Uncertainty: Integrating Markov Decision Processes, Hidden Markov Models and Predictive Modeling

Hossein Kamalzadeh
Southern Methodist University, hkamalzadeh92@gmail.com

A DATA-DRIVEN FRAMEWORK FOR DECISION MAKING UNDER

UNCERTAINTY:

INTEGRATING MARKOV DECISION PROCESSES, HIDDEN

MARKOV MODELS AND PREDICTIVE MODELING

Approved by:

_____

Dr. Michael Hahsler

_____

Dr. Eli Olinick

_____

Dr. Halit Uster

_____

Dr. Aurelie Thiele

_____

Dr. Vishal Ahuja

_____

Dr. Jay Rosenberger

A DATA-DRIVEN FRAMEWORK FOR DECISION MAKING UNDER

UNCERTAINTY:

INTEGRATING MARKOV DECISION PROCESSES, HIDDEN

MARKOV MODELS AND PREDICTIVE MODELING

A Ph.D. Dissertation Presented to the Graduate Faculty of the

Lyle School of Engineering: Engineering Management, Information, and Systems

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Operations Research

by

Hossein Kamalzadeh

(B.S., Isfahan University of Technology, 2014)
(M.S., Amirkabir University of Technology, 2016)

May 16, 2020

Kamalzadeh, Hossein   B.S., Isfahan University of Technology, 2014

                                       M.S., Amirkabir University of Technology, 2016

A Data-Driven Framework for Decision Making Under Uncertainty:

Integrating Markov Decision Processes, Hidden

Markov Models and Predictive Modeling

   Advisor: Professor Michael Hahsler

Doctor of Philosophy conferred May, 16, 2020

Dissertation completed Jan, 20, 2020

   The problem of decision making under uncertainty can be broken down into two parts. First, how do we learn about the world? This involves the problem of modeling the system and its uncertainty. Secondly, given what we currently know about the world, how should we decide what to do, taking into account uncertainty of future events and observations that may change our conclusions. Many systems evolve over time and often the next state of the system is not known with certainty, often modeled as a probability distribution over system states. Dealing with such systems especially when we can make a decision at different points in time is difficult due to uncertainty. Making optimal decisions requires understanding the system including its characteristics, how it evolves and changes over time, and how taken actions affect the system. There are multiple dimensions to this problem, and each dimension might require its own specific method. We need a descriptive method that can summarize the system and its evolution, a predictive model that is used to extract information from the complicated systems and also a prescriptive model that works as the main decision model and incorporates the effects of actions. In this thesis I consider Partially Observable Markov Decision Process (POMDP) as the main decision-making/prescriptive model, Hidden Markov Models (HMM) as the descriptive model of system evolution, and a predictive model to create observations

for the POMDP. In this research, I develop a framework by combining these methods and demonstrate its use with two applications. I apply the proposed framework to the problem of diabetes screening and also resource allocation under uncertainty for emergency management. I demonstrate using simulation that implementing the proposed policy will bring about significant improvements in both systems compared to the existing policies.

Keywords: decision-making under uncertainty, predictive analytics, Markov decision process, hidden Markov models

TABLE OF CONTENTS

List of Tables

Chapter 1

Introduction

The problem of decision making under uncertainty can be broken down into two parts. First, how do we learn about the world? This involves both the problem of modeling our uncertainty about the world, and that of drawing conclusions from evidence and our initial information. Secondly, given what we currently know about the world, how should we decide what to do, taking into account future events and observations that may change our conclusions (Dimitrakakis and Ortner (2018)). In other words, understanding the system for which we are trying to make a decision plays an important role in making an optimal decision at any point in time. This understanding includes knowledge about the currently most likely state of the system, and how it may evolve over time. The next step is to find a method to make optimal decision.

Many systems evolve over time in a discrete manner where their status changes from one state to another. Often, the state of the system is not known with certainty. Dealing with such systems especially when we have to make a decision at some point in time is difficult due to our uncertainty about the system. Though the short term effects of the decision made now might be obvious to the decision maker, the long term consequences of such decisions are hard to estimate since the system has inherent uncertainty associated with it. In other words, the decision maker might think that the decision he is making right now is the best since it seems to have the largest short-term reward, the long-term effects on the system are less certain. This can be more complicated if the decisions can affect how the system changes from one state to another. An example is called the tiger problem (Cassandra et al. (1994)), where you are trapped in a room with two doors. Behind one of the doors is a hungry tiger waiting to eat you while behind the other is a treasure. You have no idea behind which door the tiger is. You can open either doors or listen to see if you can hear anything. Listening is

not accurate since you might hear the tiger behind the left door while it is actually behind the right one. Making the optimal decision (how long to listen before opening a door) in such a system demands understanding the uncertainty, and how the actions taken reduce uncertainty.

1.1. Contributions

Making optimal decisions under uncertainty, requires understanding the system including its characteristics, how it evolves and changes, and how the actions affect the system over time. This demands special tools and methods to deal with, and it might also require integration of multiple methods from different areas; methods that help us model the underlying states of a system, and how these states evolve into each other. A descriptive method handles modeling the system and its evolution. A prescriptive model works as the main decision model and analyses the effects of actions and decisions on the system and in long term. A predictive model is used to extract information from the system, providing the decision model with information that the actual descriptive model may not be able to provide. The major contribution of this research is the integration of these methods in a single data-driven framework and the application to several problems. An R package named 'pomdp' is also developed to support this research, enabling the user to easily define POMDP models and solve them. The manual of the package can be found in the appendix.

1.2. Methodology

The main problem this research is dealing with is the difficulty of making optimal decisions in situations that have inherent uncertainty stemming from a complex system. These systems can often be modeled as a combination of multiple states that transition into each other with certain probabilities. This requires methods that can model the progression of these systems and take into account their multi-state nature. We will later see that this is the main reason why we have chosen methods such as Partially Observable Markov Decision Process (POMDP) as the main decision model for this research. We will clarify on the methods and

techniques used in this research later in chapter 2 and discuss the relation between them.

In this research we use real data collected in a quantitative approach from the application areas studied. The data collected is analyzed from various perspectives to estimate the characteristics of the associated system it was collected from and these characteristics are later used to simulate a duplicate of the system in order to further analyze it.

This research is mainly model-based driven by idealized model (which is usually denoted as axiomatic research). The primary concern here is to obtain optimal solutions within the defined model and make sure that these solutions provide insights into the structure of the problem. Typically, axiomatic research is normative, although descriptive research, aimed at understanding the process that has been modeled, is also present. Normative research is primarily interested in developing policies, strategies, and actions, to improve over the results available in the existing literature, to find an optimal solution for a newly defined problem, or to compare various strategies for addressing a specific problem. Although in the axiomatic domain, the discussion on methodology is largely absent, the operational research approach of this research consists of a number of phases including (1) conceptualization, (2) modeling, (3) model solving, and (4) implementation (Will M. Bertrand and Fransoo (2002)).

In the conceptualization phase, we develop a conceptual model of the problem and system being studied. We make decisions about the variables that need to be included in the model, and the scope of the problem and model to be addressed. In the next phase, we actually build the quantitative model, thus defining causal relationships between the variables. After this, the model solving process takes place, in which the mathematics usually play a dominant role. Finally the results of the model are implemented, after which a new cycle can start.

1.3. Motivations and Applications

In this research, we examine two systems each having their own characteristics and behavior while we are trying to provide policies for the decision makers of each system, policies that work optimally given the uncertainty in the systems. We use the proposed

framework we talked about in the previous section of this chapter for both applications and we demonstrate how this framework and the integration of the methods we use works for both applications in providing optimal policies for decision makers in systems that have inherent uncertainty. For the first system, we focus on chronic diseases such as HIV, Diabetes, and CKD, where modeling the initial uncertainty about what stage of the disease the patient is in and what decision should be taken with respect to the patient's status taking into account the future events are the major problems. For the second system, we mainly focus on emergency management, where the uncertainty lies in which area of the city is in need of more resources in the near future. What is common among these two systems, is first, they are both systems that have states changing over time with uncertainty associated with them, and second, the actions and decisions of the decision maker affects the system and has long term effect on it. We also use the same framework we talked about to deal with each system.

Application 1: Diabetes Screening: In chapter 3, we focus on chronic diseases specifically diabetes. Type 2 diabetes (which for the sake of simplicity we call diabetes here) is a major cause of morbidity and mortality worldwide. Diabetes is the 7th leading cause of death in the U.S. and causes macro-vascular complications, including heart attacks and strokes, and micro-vascular complications including retinopathy, nephropathy, and neuropathy (Petersen (2016)). The number of people who have Diabetes worldwide was estimated to be 221 million in 2010 and is expected to increase to 300 million by 2025 (Bjork (2001)) In the U.S. 9.4% of the population (30.3 million) have diabetes, 7.2 million of which are undiagnosed. An additional 33.9% of the population (84.1 million) have prediabetes of which almost 77 million are undiagnosed (CDCP (2017)). Consequently, diabetes is a major source of medical expenditures in the form of direct medical costs including hospital inpatient care (43% of the total direct medical expenditures), prescription medication to treat the complications caused by diabetes (18%), antidiabetic agents and diabetic supplies (12%), physician office visits (9%), and nursing/residential facility stays (8%) (Petersen (2016)). In

the U.S., estimates of direct costs were increasing from \$176 billion in 2012 to \$237 billion in 2017 (American Diabetes Association (2018)). Diabetes also imposes high indirect costs due to work-related absenteeism, reduced productivity at work and home, reduced labor force participation from chronic disability and premature mortality which increased from \$69 billion in 2012 to \$90 billion in 2017 (American Diabetes Association (2018); Bjork (2001); Petersen (2016)).

Application 2: Emergency Management: In chapter 4, we focus on emergency management. According to Dallas Fire and Rescue Department, a structural fire incident needs resources from several fire stations around the city which are close to the incident location, each providing a specific type of vehicle. This means if more than a single structural fire incident happens in a small area of a city within a short period of time, no resources would be available to be dispatched to the incident. This can cause huge damages.

An important question for the Dallas Fire and Rescue Department is whether resources should be moved around in the city to cover areas where the resources are currently responding to an ongoing incidence. Every time an incident happens, resources in a particular zone of the city will be dispatched and become unavailable for several hours. If another incident happens in that zone during that time, resources from other areas of the city will need to respond which will increase response time. To mitigate such situations, we can temporarily reallocate resources.

1.4. Structure of this thesis

In chapter 2, we propose to utilize and combine three techniques and methods in a single framework to model each system using its key characteristics. In our framework we have a descriptive model that uses the characteristics of the system's evolution to model its changes over time including the inherent uncertainty in the changes. This model is learned directly from the data available from each application area. We use a prescriptive decision model, to optimize the decisions and actions the decision maker can make taking into account each

action's immediate and long term effects on the system. The decision model provides us with an optimal policy. Additional information (signals) are provided using the predictive model. The predictive model is also directly learned from the available data.

In chapter 3, we focus on chronic diseases specifically diabetes. We propose a targeted screening policy (equivalently, screening strategy) that uses all available information on individual patients to identify whom to screen (that is, which patients should receive the gold-standard test) and when to screen them; the policy is also age-specific. We develop and validate our model on a detailed and proprietary dataset – of over 12,000 patients over an 18-month period – from a large safety-net hospital and demonstrate, using a simulation analysis, that our proposed screening policy can improve patient outcomes.

In chapter 4, we focus on emergency management. We apply the proposed framework, and formulate the problem as a POMDP problem. We focus on one city zone in order to define our state space. We try to capture the availability of the resources in that zone in the near future; By implementing the proposed POMDP policy, and through simulation, we demonstrate that we can improve the average response time by a significant amount compared to existing policies.

Chapter 5 concludes this thesis. In the appendix of this thesis, details of the simulations conducted as well as the manual to the R package 'pomdp' developed to support the research are included.

Chapter 2

A Data-Driven Decision Framework

We propose to utilize and combine three techniques and methods in a single framework to create a decision framework that uses data in all main phases. In our framework we have a descriptive model that uses the characteristics of the system's evolution to model its changes over time including the inherent uncertainty in the changes. This model is learned directly from the data available from each application area. We use a prescriptive decision model, to optimize the decisions and actions the decision maker can make taking into account each action's immediate and long term effects on the system. The decision model provides us with an optimal policy. Additional information (signals) are provided using the predictive model. The predictive model is also directly learned from the available data.

Figure 2.1 represent the integration of the methods into a single decision-making framework. The vertical classification of the methods used in the framework can vary based on the application but it is strongly related to the perspective the decision maker is looking at the system from. We will later see how this classification works for each of the applications.

In this work we will use Partially Observable Markov Decision Process (POMDP) as the main decision-making/prescriptive model (Figure 2.1 upper-right box). The reason behind choosing POMDP as the main decision model, is the nature of the systems we are analyzing; the status of each system can be modeled into separate states that change over the time and these states could be the actual states of a Markov chain. We will elaborate on this later in each chapter associated with each application.

We use Hidden Markov Models (HMM) as the descriptive model of the systems (Figure 2.1 upper-left box). The role of HMM is to model the evolution or dynamics of the system in a Markov chain and estimate the parameters of this Markov chain. HMM here provides the decision model with the parameters of the underlying Markov chain that is being used in
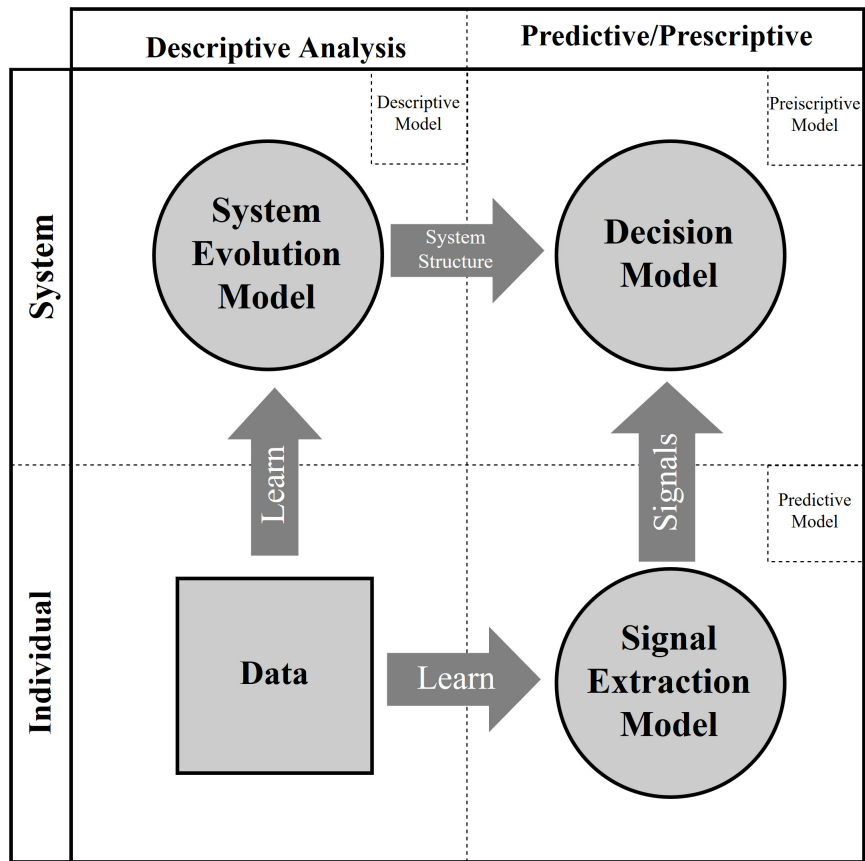
Figure 2.1: Multi-method decision-making framework to combine descriptive modeling and predictive modeling with optimization

the POMDP of the decision model. The HMM is directly learned from the historical data.

The predictive model (Figure 2.1 lower-right box) provides the POMDP model with external information in the form of observations. Below we will see how this predictive model works in combination with the decision model and how this integration works toward the contribution of this research.

2.1. Stochastic systems and their evolution

When it comes to decision making under uncertainty, understanding the system for which we are trying to make an optimal decision is of great importance. Applying rule of thumb methods is popular when it comes to decision making under uncertainty, but it typically leads

to suboptimal and often very poor decisions. Systems vary in terms of how they change over time (how their dynamics work) and where the actual inherent uncertainty comes from. The way a system changes over time and where the uncertainty comes from impact how an optimal decision should be made given the current state of the system.

The evolution of many systems over time is continuous but can be simplified into discrete time-steps with a finite number of states. Including uncertainty, such a system can be modeled as a discrete-time stochastic process with a discrete state space. Assuming that, given the current state of such a system the future state of the system is independent of the past states, the system can be modeled as a Markov chain. If we narrow down the systems we are dealing with to a system that can be modeled as a Markov chain, then a set of techniques including Markov Decision Processes (MDPs) or its generalizations such as Partially Observable Markov Decision Processes (POMDPs) can be applied to determine optimal decisions. The use of MDPs or POMDPs depends on the nature of the system and where the uncertainty comes from.

In some systems, the current state of the system cannot be observed directly and thus is unknown. Only a probabilistic belief of the current state can be constructed using observations or information coming from the system. POMDPs which are a generalization of MDPs, allow capturing this type of uncertainty regarding the observability of the current state of the Markov process. For many applications, the current state of the system is either unknown or unobservable by the decision maker and this adds to the uncertainty that lies within the system's evolution. The two major complications regarding POMDPs are due to the two types of uncertainty in these types of systems: First, the uncertainty resulting from the stochastic nature of the system evolution, and second and more importantly, the uncertainty regarding the current state of the system which has to be inferred via imperfect information. The second type of uncertainty here is formed by the relation between the underlying state of the system and the observations produced by the system revealing some information about the current state of the system.

Observations used in POMDPs can be any signal that the system emits which gives information about the actual state of the system. The nature of the observations depends on the nature of the system. In the tiger problem (Cassandra et al. (1994)) for example, the decision maker needs to decide which of two doors to open. Behind one door is treasure while behind the other is a hungry tiger. The decision maker does not know behind which door the tiger is and can only make observations by listening for tiger noises which are not perfectly accurate. The question is how often to listen for tiger noises before the decision make opens a door. The more complex the system is, the more different observations can be made about the current state of the system. An observation can be a single signal observed at a time or a combination of signals from different sources within the system. What matters is how much an observation will help the decision maker to determine the current state of the system and thus to make the best decisions. Therefore, finding accurate sources of observations from the system and choosing the best ones is a key step in modeling a POMDP and making the best decisions.

2.2. Introduction to Partially Observable Markov Decision Processes

POMDPs are generalizations of MDPs where there the state space is not completely observable to the decision maker (Drake (1962)). A discrete-time POMDP model is a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \Omega, \mathcal{O}, R, \lambda)$, where

- $\mathcal{S}$ is the set of states (s) describing the various states the system can be in,

- $\mathcal{A}$ is the set of available actions (a) the decision maker can take,

- $\mathcal{P}$ is the set of transition probabilities between the states which simply describes how the system evolves over time and is conveying part of the uncertainty in the system (stochastic dynamics of the system),

- Ω is the set of all observations (o),

- $\mathcal{O}$ is the set of observation probabilities or how the observations relate to the actual states of the system,

- R is the reward function of the model, and

- λ is a discount factor between 0 and 1.
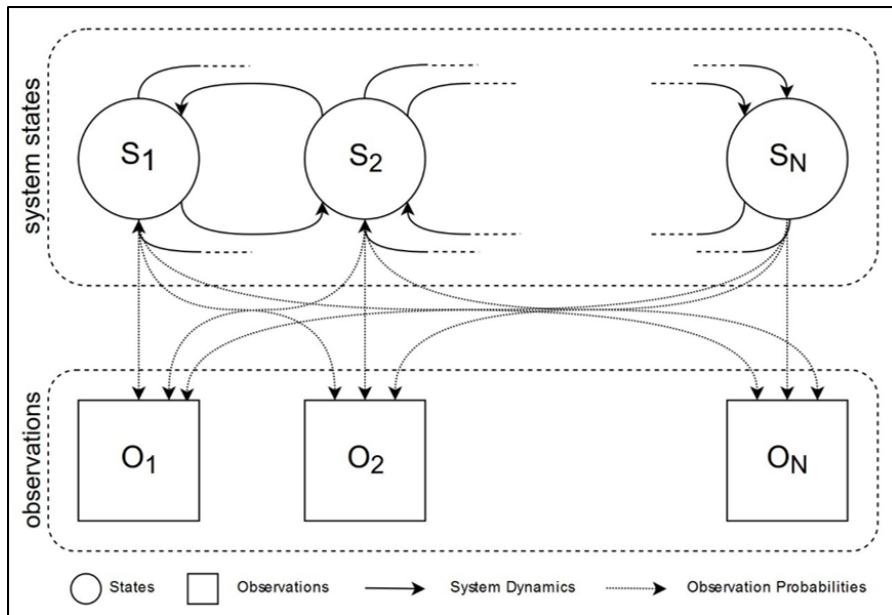
Figure 2.2: States space and observation space of a POMDP model

The state space and observation space of a POMDP model are depicted in Figure 2.2.

In a POMDP, the states, actions, and observations can be discrete. We denote them at time by s_t, a_t, and o_t respectively. Transition probabilities are action and state-dependent function: $\mathcal{P}(s_{t+1}, s_t, a_t) = pr\{s_{t+1}|\, s_t, a_t\}$. The observation probabilities are a function of state, action, and observation: $\mathcal{O}(a_t, s_{t+1}, o_t) = pr\{o_t|a_t, s_{t+1}\}$. Since the states are not directly observable, the decision maker's belief about the current state of the system is represented by a belief state π_t which is a probability distribution over all possible states. The belief state is updated using Bayes' rule every time an action is taken and an observation is observed: $\pi_{t+1}(s_{t+1}) \propto \mathcal{O}(a_t, s_{t+1}, o_t) \sum \mathcal{P}(s_{t+1}, s_t, a_t)\pi_t$. The importance of the observations and their relations with the actual states is given in the belief state update formula based on Bayes' rule where the function $\mathcal{O}$ is used. The more accurate this function is in terms of providing information about the actual state of the system, the better is the solution of the POMDP.

At each time step or decision epoch, the decision maker makes a decision and takes an action available from the action set. The decision maker takes this action based on the observation. The system then evolves into a new state, new observations are made, and the decision maker needs to take an action again. Each time an action is taken a certain amount of reward is given to the decision maker based on the given reward function $R(s_t, a_t)$ which is action and state-dependent.

In POMDPs we are trying to find a set of actions (a policy) that maximizes (minimizes) the expected total discounted rewards (costs) over an infinite horizon. Such a policy is called the optimal policy. The optimal policy $\wp^*$ is obtained by solving the Bellman Optimality Equation

$$V^*(\pi) = \max_{a \in \mathcal{A}} \left\{ R(\pi, a) + \lambda \sum_{o \in \Omega} \mathcal{O}(a, \pi, o) V^*(\pi') \right\}.$$

The optimal value can be computed by applying dynamic programming to iteratively improve the value of the function.

Since the belief space is uncountable, the above dynamic programming recursion does not translate into practical solution methodologies. Even with the finite dimensional characterization of a POMDP (finite state space, finite action space and finite observation space), determining the piecewise linear segments of the value function at each epoch is computationally expensive due to the fact that the number of piecewise linear segments can increase exponentially with the action space dimension, state space dimension and observation space dimension. Therefore, exact computation of the optimal policy is only computationally tractable for small state dimension, small action space dimension and small observation space dimension. It is shown in Papadimitriou (1987) that solving a POMDP is a PSPACE-complete[1] problem. Littman (2009) gives examples of POMDPs that exhibit this worst case behavior. It is inferred that simplifying a POMDP model in any way such

---

[1] Decision problem A is PSPACE-complete if both of the following are true (Sipser (1997)):

1. $A \in$ PSPACE (PSPACE: Decision problems solvable in polynomial space)

2. For every $X \in$ PSPACE, $X \leq_P A$.

as reducing the dimension of any of the spaces including the observation space can save significant amount of computational expenses. In the next section, we will see how this problem of interest has been studied in the literature.

2.3. Literature related to Partially Observable Markov Decision Processes

Controlling a Markov process with incomplete state information (including a partially observable state space) was first studied in Dynkin (1965). The first POMDP model was developed in Drake (1962). Other researchers at the same time developed finite horizon POMDPs in the context of stochastic control problems (Aoki (1965); Astrom (1965)). During the past years many other generalizations and versions of POMDPs have been investigated and developed by researchers including POMDPs with an uncountable action space (Sawaragi and Yoshikawa (1970)), POMDPs with Borel spaces (Rhenius (1974)), POMDPs with an arbitrary core process state space (Furukawa (1967)), non-stationary POMDPs (Hinderer (1970)), undiscounted infinite horizon POMDPs (Platzman (1980)), semi-Markov core process PODMPs (White (1975, 1976)), and so on.

There also exists a large number of papers investigating Bayesian control of the sequential decision process including Furukawa (1967); Rieder (1975); Satia and Lave (1973); Wessels (1968).

In terms of dealing with observations and the observation space, not much research has been reported. Most of the studies that utilize POMDPs to solve their problems including Ayer et al. (2012a); Cassandra (1997); Grosfeld-Nir (1996); Hauskrecht (2000); Littman (2009); Monahan (1982); Sandikci et al. (2013) simply and naively assume that a single-signal is apriori known, not considering the fact that real-world systems produce a large number of signals and that the observation space plays a significant role in POMDPs where the current state has to be inferred through observations. There exist only a few studies that deal with how to select multiple observations from a multidimensional observation space. In Hoey and Poupart (2005) authors speak of multidimensional observation spaces, how to sample from them, and how to aggregate observations in order to reduce the dimensionality of the

13

observation space. Observations can be aggregated in some cases if the policies associated with them are the same (policy-directed observation aggregation). Observations can also be aggregated in one-dimensional continuous observation spaces by discretizing the continuum into segments whose observations yield the same optimal policy (Hoey and Poupart (2005)). For multidimensional observation spaces, authors in Hoey and Poupart (2005) examine two approaches. The first approach is for observation spaces where observations are composed of conditionally independent variables. For this case, they reduce the observation space to one dimension by sequentially processing the observation variables in isolation. The second approach which is used for arbitrary multi-dimensional observations is sampling and it is proved to be an effective approximation technique for computing aggregate probabilities. The authors in Hoey and Poupart (2005) propose a dynamic partitioning technique which is integrated with point-based backups. There are several drawbacks associated with these types of methods. These methods do not work with some POMDP algorithms such as Incremental Pruning (Cassandra et al. (2013); Zhang and Liu (1996)), Witness algorithm (Kaelbling et al. (1998a)), and Bounded Policy Iteration (Poupart and Boutilier (2004)). Another drawback is that the proposed method in Hoey and Poupart (2005) deals only with POMDPs with continuous observations but discrete states. Although others have tried to reduce the dimensionality of the whole POMDP by reducing the observation space's dimensionality, they have never looked at the problem from a different perspective. All the efforts that have been made are post-POMDP dimensionality reductions. Here we apply a pre-modeling technique for observation aggregation that not only produces more meaningful and accurate observations from the system, but it also shapes both the state space and observation space in full compliance with each other.

2.4. Predictive Modeling for Observation Aggregation

Observations and their relations to the actual states in POMDPs are extremely important. Systems may produce more than one signal and these signals can be used as observations in POMDPs. The question of how to choose the signals to use as observations in a POMDP

is still open in the literature. If all signals are used as observations, we will have a large observation space which makes POMDPs very difficult to solve. Even if there is no problem with the dimensionality, determining the relationship between these many signals and actual states of the model is not always possible. As mentioned before, reducing the dimensionality of the observation space by aggregating the observations into more accurate and informative ones can save significant amount of computational expenses.

The fact that systems produce signals all the time (either continuously or discretely in time) reveals that the information from these signals gathered over time provides historical data for the system. If enough data is gathered from the sources of signals (enough signals recorded), the data can then be analyzed for further purposes using data-driven analytics techniques. One purpose is to select a subset of signals and aggregating them into a single more accurate and meaningful observation. This problem is broken down into two major steps in this research work. The first step is to select a proper subset of signals from the system (signal selection step). And the second step is to combine or aggregate the selected signals into a strong observation (signal aggregation step). These two steps are depicted in detail in Figure 2.3.

The first step is a feature selection problem. We try to select a subset of features (signals) that are later going to be used in a predictive model to produce more meaningful observations. From another perspective, using historical data recorded from the signals, in this step we identify which signals are giving more information about the actual states of the model. The outcome of this step would be a list of signals, sorted based on their strength in pointing to the right state of the system.

What needs to be taken into consideration in the signal selection step, is the problem of missing data. This matters because one type of signal may be accurate, but it might be harder to observe and thus not always available. We will later see in the signal aggregation step why this is important.

The signal aggregation step is implemented by a predictive model that uses the selected features from the previous step in order to provide outcomes that are more meaningful and
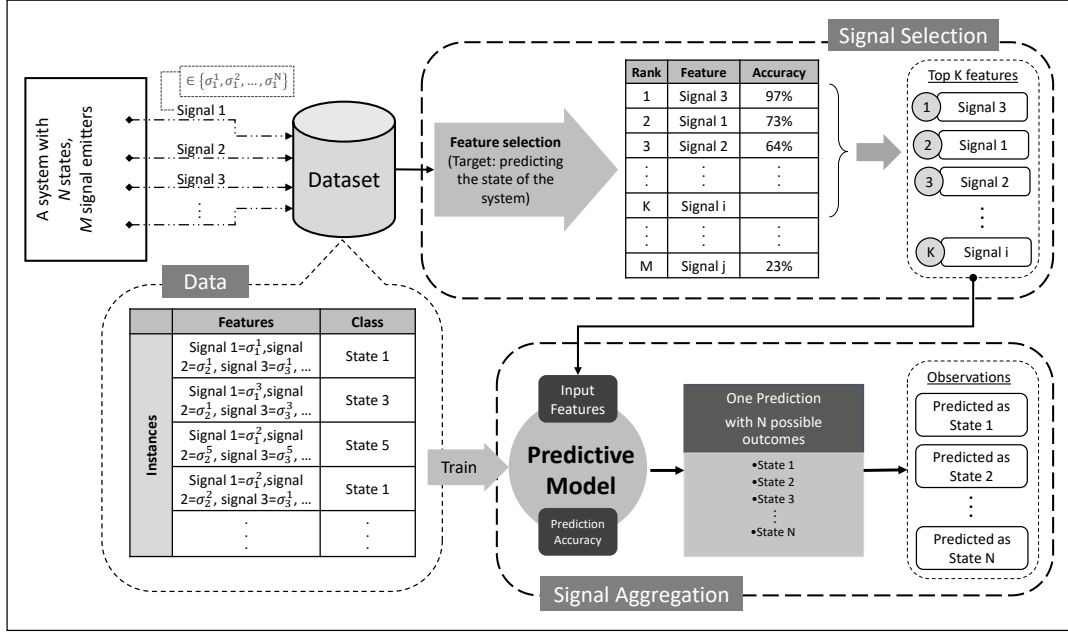
Figure 2.3: Data-driven signal selection and aggregation framework

accurate. In this step, we develop a classification model where the input is the selected signals from the signal selection step and the classes are the actual states of the model. By training this classifier we will have a predictive model that takes all the signals as the input and then predicts the state of the system. These predictions are then used in the POMDP as observations to update the belief state.

Predictive models are rarely perfect. There are always misclassification errors associated with such models. These errors are taken into consideration and used as the relationship between the predictions (that are going to be used as the observations) and the actual states of the POMDP. In another word, the accuracy of the predictive model is implemented in the POMDP as the observation probability function.

Figure 2.3, shows an example system that has a total number of N states (i.e. $card(S) = N$) and the system produces M signal at each epoch. Each signal can take N distinct values (signal $i \in \left\{\sigma_i^1, \sigma_i^2, \ldots, \sigma_i^N\right\}$). This means that in a POMDP model that takes all these observations into account, the total number of observations would be N^M (i.e.

$card(\Omega) = N^M$). Not all these N signal are perfect in pointing to the true underlying state of the system at each epoch. Some might work better than others. The strong signals are selected in the signal selection part of the framework. The predictive model then uses these signals to predict the state of the system. Although the prediction (the outcome of the predictive model) is in terms of the state of the system, it typically will not be completely accurate but can be used as an informative observation, increasing our understanding of what state the system most likely is in. Using the framework in Figure 2.3 we produce only one signal out of M signals, and this signal can take N distinct values (observations). Thus, the total number of observations will be N and therefore the size of the observation space is reduced significantly.

Chapter 3

Optimal Individualized Diabetes Screening (P1)

This chapter describes the application of the decision framework to the problem of diabetes screening. In this chapter, we provide details on the techniques used in the decision framework in Figure 3.1. We use POMDP to formulate the sequential screening decision-making problem. The model is informed by the population-specific disease progression learned from data using the HMM. The disease stages and the costs to the healthcare system and the patient are derived from the medical literature and clinical expertise. The screening decisions are highly personalized using a predictive model trained on a large set of electronic health record data. While any predictive model can be used, we apply here a logistic regression model with L1 regularization (LASSO). The solution of the POMDP given the assumptions is an optimal screening policy which can be used in clinical practice.

We propose to supplement existing guidelines with an opportunistic screening strategy that (1) incorporates all clinical information available about each patient to identify individuals at higher risks of developing prediabetes or diabetes, and (2) identifies the optimal time to perform the screening to optimize expected health outcomes and healthcare cost. Figure 3.1 shows the high-level multi-method framework proposed in this paper. We use a POMDP model (upper right) to find an optimal policy for the main decision-making problems of whom to screen and when to screen/re-screen. The transition parameters of the POMDP model are estimated using a disease progression model (upper left), a Hidden Markov Model learned from historical patient data. The observations used by the POMDP model are created via a predictive model that incorporates patient-level risk factors (lower right).
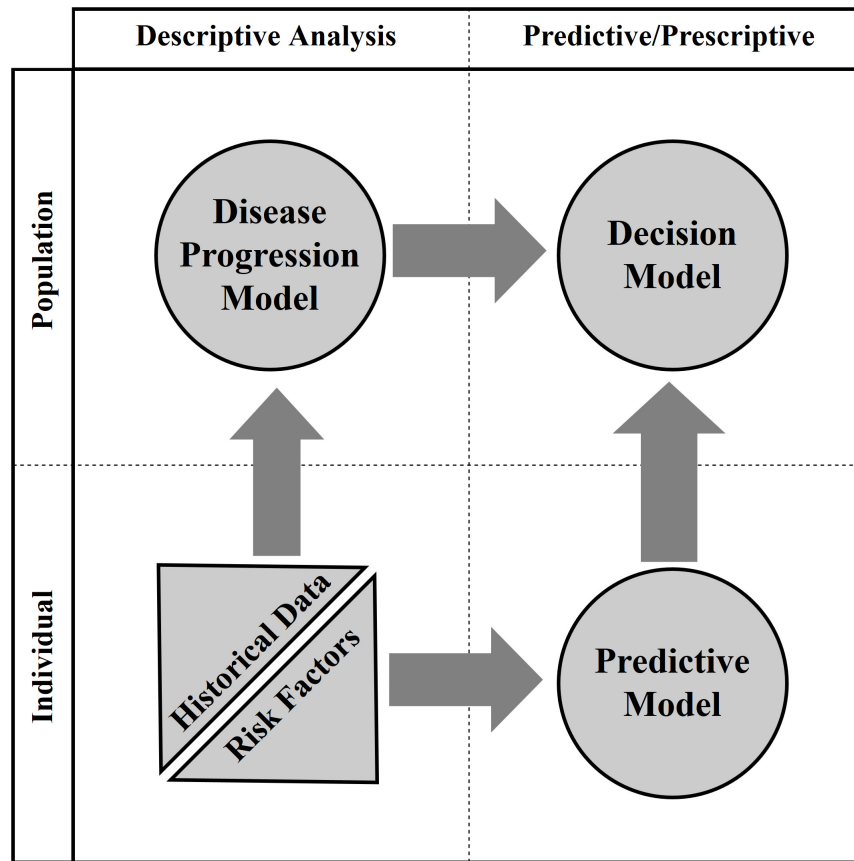
Figure 3.1: Multi-method decision-making framework to combine progression modeling and predictive modeling with optimization

3.1. Background on Diabetes

Like many other chronic diseases, Type 2 diabetes has a prolonged asymptomatic period during which early detection is possible because diabetes onset occurs on average 9-12 years before clinical diagnosis Lu et al. (2010). Diabetes risk increases across a continuum with higher glucose levels corresponding to higher risk as the glucose level is an indicator of whether the patient has diabetes. For diagnostic and treatment purposes, two key stages are characterized – prediabetes and diabetes. In the prediabetes stage, patients are asymptomatic and blood glucose is higher than normal but not high enough to be classified as diabetes. Although progression to diabetes can be reversed by lifestyle modification and interventions like bariatric surgery, many patients with prediabetes go on to develop diabetes,
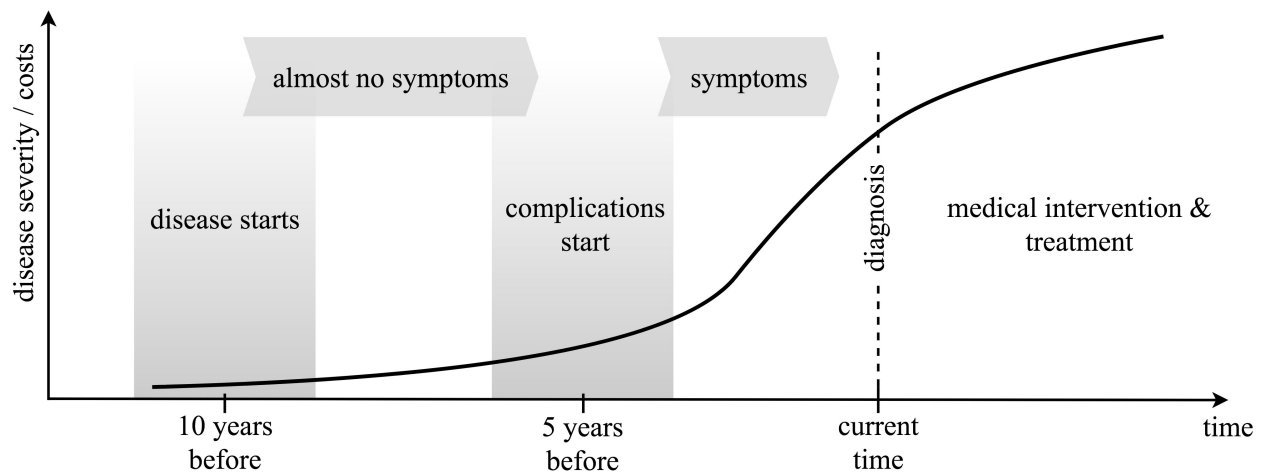
Figure 3.2: the costs associated with the disease increase very quickly as the severity of the disease increases

a chronic disease requiring medical treatment to control the disease and prevent/manage complications. Importantly, identification of patients during the prediabetes stage allows the delivery of evidence-based interventions to delay or prevent the development of diabetes prevention Program (2008); Group (2002). Thus, screening of individuals at risk for diabetes and timely surveillance of patients with prediabetes to detect the transition to diabetes is critical to improving health outcomes and reducing healthcare costs (see Figure 3.2).

Systematic diabetes screening and prevention programs can identify patients at risk for diabetes and target preventive interventions to delay or prevent the development of type 2 diabetes. The American Diabetes Association (ADA) and the US Preventative Services Task Force (USPSTF) provide physicians with guidelines for screening. These guidelines recommend screening about 70% of the population Calonge and Petitti (2008); Care (2013), which is a very expensive proposition and in many cases, operationally impractical Howard et al. (2010). The guidelines are based on only a as small number of predictors including age, body mass index, and a few risk factors. This results in a sensitivity as low as 65% for USPSTF and specificity as low as 23% and 67% for ADA and USPSTF respectively for identifying diabetes cases.

3.2. Related Literature

Figure 3.1 provides an overview of the methods used in this paper to address the problem of diabetes screening. Many of these methods have been employed independently to answer specific questions in the healthcare context but they have not been integrated to address such a decision-making problem. In the following, we review the literature related to these methods and techniques.

3.2.1. MDP for Medical Decision Making

Markov Decision Processes (MDPs) and Partially Observable MDPs (POMDPs) are methodological tool of choice to study medical decision making problems for chronic diseases such as Diabetes (Kuo et al. (1999); Santoso and Mareels (2001); Shih et al. (2007); Hoerger et al. (2004)), HIV/AIDS (Lee et al. (2014); Gafa et al. (2012); Shechter et al. (2008)), Cancers (Ayer et al. (2012b); García-Mora et al. (2010); Ahsen and Burnside (2018); Maillart et al. (2008); Chhatwal et al. (2010)) and their associated complications (Sandikci et al. (2008, 2013); Schaefer et al. (2004); Sukkar et al. (2012); Alagoz et al. (2010)). MDPs have also been used for hypertension treatment specifically for designing therapeutic regimens for patients with hypertension (Zargoush and Daskalopoulou (2018)). Also, MDPs have been used for a wide range of healthcare management problems such as dealing with emergency department congestion (Patrick (2011)). An attractive key feature of MDPs is that they can be used to deal with sequential decision-making problems in contexts with large levels of uncertainty (for example, in terms of how fast the disease progresses in a given patient population). In such settings, MDP's can be used to determine the optimal time for screening and treatment initiation (Alagoz et al. (2010)). For example, MDPs have been used to answer questions such as: optimal time to initiate antiretroviral therapy in HIV patients (Shechter et al. (2008)), optimal time for breast cancer screening in women (Maillart et al. (2008); Chhatwal et al. (2010); Ayer et al. (2012b)), or optimal time for accepting a living-donor transplant in patients suffering from end-stage liver disease (Alagoz et al. (2004, 2005); Sandikci et al. (2013, 2008)). Readers are directed to (Alagoz et al. (2010); Monahan (2008); Cassandra

(1997); Schaefer et al. (2004)) for a review of literature describing uses of MDPs in medicine.

3.2.2. HMM to Model Disease Progression

Disease progression modeling is important for disease prognosis improvement, drug development, and clinical trial design. Difficulties with modeling disease progression include progression heterogeneity (patients have different progression trajectories due to many reasons), incomplete patient records (censoring and missing information), discrete observations (disease progression is a continuous process, but patients' records of the progression are observed and recorded at discrete times with varied intervals), and irregularity of observations (due to irregular visits) (Wang et al. (2014)).

A large portion of the literature on disease progression modeling focuses on evidence-based modeling using machine learning and statistical techniques based on observational data. A popular model is the hidden Markov model, where disease progression is modeled as a progression through a set of unobservable discrete disease states governed by transition probabilities. For example, a general hidden Markov model to estimate transition rates between states as well as the probabilities of states of misclassification is presented in Jackson et al. (2003). Another study (Liu et al. (2015)) presents an effective learning method for continuous-time HMMs by dealing with the challenges of estimating the posterior state probabilities and the computation of end-state conditional statistics. In Sukkar et al. (2011) the authors develop a six-state HMM of Alzheimer's disease which allows progression by one or two states or regression by one state using data from 595 subjects. They calculate the states transitions and conditional probabilities of being in each state using the developed model. The authors also propose an HMM for the Alzheimer's progression in another study (Sukkar et al. (2012)) with the ability to identify more granular disease stages than the three currently accepted clinical stages for Alzheimer's disease. Some studies use techniques other than HMMs to model the disease progression or obtain state transitions such as simulation (Lee et al. (2008)). Best practices on estimating the transition rates between states including techniques such as HMMs can be found in Denton (2018); Siebert et al. (2012).

3.2.3. Predictive Models in Healthcare

There is a growing number of studies using predictive models in healthcare decision making. These studies include the use of analytics in healthcare such as personalized diabetes management (Bertsimas et al. (2017)), chemotherapy regimens for cancer (Bertsimas et al. (2016)), hospital readmissions (Shams et al. (2015)), and healthcare screening decisions such as screening for Hepatocellular Carcinoma (Yuen and Lai (2003)), breast cancer screening (Maillart et al. (2008)), and HIV screening (Deo et al. (2015)). Studies on the use of predictive models for diabetes screening are reviewed in (Collins et al. (2011)) where the authors conduct a systematic review of the methodology of 39 studies and in (Jahani and Mahdavi (2016)) where the authors develop neural network models for diabetes prediction and compare with other models.

(Collins et al. (2011)) survey 39 studies with 43 risk prediction models that use 4 to 64 predictors including age, family history, body mass index (BMI), hypertension and fasting glucose. The most common modeling method among these studies is logistic regression. It is reported, that almost all reviewed studies remove incomplete cases or do not mention how missing data are treated. There are two types of predictive model in the literature, single-factor and multi-factor models. The single-factor models use common predictors such as age or BMI for which the availability in routine clinical settings is high. The drawback for single-factor predictive models is that no prediction can be made if the factor is not available for a patient. On the other side, for multi-factor models, can incorporate many factors, but since all these factors need to be available for the patient, for the sake of practicality a small number of predictors is typically preferred. Multi-factor models consider more information about the patient and therefore can provide better predictions compared to single-factor models.

The majority of the reviewed literature focuses on using a only single technique of the multi-method framework proposed in Figure 3.1. While these methods individually can be used to predict desease progression at the population level or what patients are more at risk of having undiagnosed diabetes, they only... The key contributions of this paper

is that we group all these methods and techniques together, using one to feed another, feeding all with real data to answer a question that has implications for clinical practice as well as contributions to a theoretical operations literature. Researchers have used the same techniques but independently, they have used MDPs or POMDPs to model decision making problems that concern healthcare but independent of what a specific hospital system would need or without using real data. They current state of the art is to simply assume some transition rates while we actually calculate using real data. They have used HMMs to estimate transition and progression rates for various disease but not in the context of a decision making problem. They have used data driven methods including predictive models to predict specific chronic diseases such as diabetes but never used it to feed MDPs as an individualized input for the decision making problem. ... our approach is able to answer the questions of whom to screen, when to screen them and how often rescreening should take place in an integrated, analytics-driven decision framework that takes health outcomes, healthcare cost, cohort information, and available individual patient information into account.

3.3. The Partially Observable Markov Decision Process Formulation

Partially Observable Markov Decision Processes (POMDP) are an extension of Markov Decision Processes (MDP) to make optimal decisions when the current state of the system (in our case, the true health status of the patient) is not directly observable. The method uses a probabilistic belief distribution over the unobservable states of the system which is informed by observations. These Markov models assume that the process is Markovian, i.e., that future states only depend on the current state. While this is a very strong assumption, models based on the assumption are often very useful.

The set of states for the screening decision model are healthy, prediabetes, and diabetes. The decision is whether to screen the patient, henceforth referred to as "screening" decision. We assume the following: (a) the decision-maker is the clinician who acts on behalf of the patient and the health system, (b) the screening decision for a given patient is independent

24

of other patients, (c) screening decisions are made at discrete points in time when the patient and clinician meet, and (d) patients stay in each state for at least one decision epoch.

A discrete-time POMDP model is a 7-tuple $(S, A, P, \Omega, O, R, \lambda)$, where S is the set of states, A is the set of actions, P is the set of transition probabilities between the states, Ω is the set of all observations, O is the set of observation probabilities, R is the reward function of the model and λ is discount factor. Below are the detailed description of the essential components associated with the POMDP that need to be defined in advance to model the problem Cassandra et al. (1994); Kaelbling et al. (1998b); Puterman (2005):

3.3.1. Time Horizon and Decision Epochs

We use decision epochs of one year. Decisions are made at the beginning of each period starting from the first time the patient meets the clinician. We represent the epochs with $t = 0, ..., T$. The time horizon in our problem expands from the first time the patients meets the clinician until the patient dies or reaches the age of 79.

3.3.2. State Space

The state space in our model consists of a total of 7 distinct states $S = \{H, P, D, SH, SP, SD, \Delta\}$ and includes both observable and unobservable states. The 3 unobservable states are: Healthy (H), Prediabetes (P), Diabetes (D) which are the main underlying stages of diabetes. The 3 observable states are the screened representatives of the observable states: Screened Healthy (SH), Screened Prediabetes (SP), Screened Diabetes (SD). These states are completely observable, since they are the outcome of screening. The last state is Death (Δ), which is the absorbing state.

3.3.3. Action Space

The action space, $A = \{S, N\}$, represents the decision to screen (S) or not to screen (N) a patient. We use $a_t \in A$ to denote the action that is taken at time t at each decision epoch.

3.3.4. Transition Probabilities

These probabilities indicate the probability of a patient moving from the current state (s_t) to another state (s_{t+1}), given action a_t is taken. This probability is denoted by $p\{s_{t+1} \mid s_t, a_t\}$. These transition probabilities are associated with the arcs on the Markov model underlying the POMDP (depicted in Figure 3.3). We use P to represent the set of all transition probabilities (typically one state-to-state transition matrix per action). Regression from diabetes to prediabetes or healthy states is very unlikely we therefore do not include an arc from state D to P or D to H, corresponding to a transition probability of zero.

For our model, we assume that the transition probabilities are stationary in the considered cohort. Thus, we drop the index t and use the notation $p\{s' \mid s, a\}$ to denote the "stationary" probability of transitioning to state s' given the current state is s and action a is taken. A key characteristic of the transition probabilities is that the sum of the probabilities of transitioning from the current state to all other states including the current one should be equal to 1 for each single action; that is

$$\sum_{s' \in S} p\{s' \mid s, a\} = 1, \text{ for all s and a.} \tag{3.1}$$

We have the following:

$$p\{H \mid H, N\} = 1 - \sum_{s' \in S - \{H\}} p\{s' \mid H, N\} = 1 - p\{P \mid H, N\} - p\{D \mid H, N\} - p\{\Delta \mid H, N\},$$
$$\tag{3.2}$$

Similarly, for states P and D we have:

$$p\{P \mid P, N\} = 1 - \sum_{s' \in S - \{P\}} p\{s' \mid P, N\} = 1 - p\{H \mid P, N\} - p\{D \mid P, N\} - p\{\Delta \mid P, N\}, \tag{3.3}$$

26

and

$$p\{DH \mid D, N\} = 1 - \sum_{s' \in S - \{D\}} p\{s' \mid D, N\} = 1 - p\{H \mid D, N\} - p\{P \mid D, N\} - p\{\Delta \mid D, N\}.$$

(3.4)

We assume that a positive screening result (i.e., the patient is diagnosed with prediabetes or diabetes) influences the patient. The patient will receive medical treatment or may perform lifestyle changes (e.g., diet, exercising, weight loss). We capture these effects using the factors $\beta, \gamma \in (0, 1)$ which are used to reduce the transition probabilities for the disease to progress from screened states (SP, SD) into more severe stages compared to patients in the same states but not screened.

3.3.5. Observations and Observation Probabilities

At each decision epoch, a signal/observation, $o \in \Omega$, provides information about the true underlying (unobservable) state of the patient. Depending on the nature of the problem, observations can be obtained from various sources. We propose to create these observations using a predictive model (see Section 3.4) which classifies the patients into the groups of Predicted as Healthy (PH), Predicted as Prediabetic (PP) and Predicted as Diabetic (PD). Thus, the observation space is $\Omega = \{PH, PP, PD\}$. Predictive models are usually not perfect and therefore the predictions used as observations are probabilistically connected to the unobservable states, i.e., the probability associated with predicting a specific observation $o \in \Omega$, given that the true state of the patient is s is $O(o \mid s)$ where O is the set of all observation probabilities.

3.3.6. Belief States

$\Pi(S)$ is the probability simplex over the state space S, defined as $\Pi(S) = \{\pi \in R^3 : \sum_{i=1}^{3} \pi_i = 1, \pi_i \geq 0, \forall i\}$, also called the belief space Sandikci et al. (2013); Brafman (1997); Sandikci et al. (2010). We use π_t to denote the belief state at period t which is the probability

27

distribution over the set of possible states, i.e., $\pi_t = (\pi_t(H), \pi_t(P), \pi_t(D))$.

3.3.7. Reward Functions

The POMDP maximizes expected rewards. Taking action a while being in state s will bring about an immediate reward denoted by the reward function $r(s, a)$. We use as the values of the reward function estimates that combine the patient's QALY (Quality Adjusted Life Year Neumann et al. (2014b)), the costs of prediabetes, diabetes and screening tests all measured in US dollars. We formulate each state-and-action specific reward function from the societal perspective as follows:

$$
r(s, a) = \begin{cases}
Q, & s = H, a = N \\
(1 - \alpha_P)Q, & s = P, a = N \\
(1 - \alpha_U D)Q, & s = D, a = N \\
Q, & s = SH, a = N \\
(1 - \alpha_P)Q, & s = SP, a = N \\
(1 - \alpha_D D)Q, & s = SD, a = N \\
(Q - C_S)u_r, & s = H, a = S \\
(1 - \alpha_P)Q - C_P - C_S, & s = P, a = S \\
(1 - \alpha_D D)Q - C_D - C_S, & s = D, a = S \\
(Q - C_S)u_r, & s = SH, a = S \\
(1 - \alpha_P)Q - C_P - C_S, & s = SP, a = S \\
(1 - \alpha_D D)Q - C_D - C_S, & s = SD, a = S
\end{cases}
\tag{3.5}
$$

where the terms C_s, $C_D, C_P, Q, l_e, l_d, u_r, Q$ and $\alpha_i,\ i = \{P, UD, DD, D\}$ are later described and estimated in Table 2 in section 4 alongside their values.

3.3.8. Bayesian Belief State Update and Optimality Equation

To implement learning from a new observation o, the belief state $\pi = (\pi(H), \pi(P), \pi(D))$ is updated to π' using the Bayes' rule. The updated component of π' associated with state

28

s' is given by

$$\pi'(s') = \frac{O(o' \mid s') \sum_{s \in S} p\{s' \mid s, a\}\pi(s)}{\sum_{s' \in S} O\{o' \mid s'\} \sum_{s \in S} p\{s' \mid s, a\}\pi(s)} \qquad (3.6)$$

Using belief states, the POMDP can be reformulated as a continuous state MDP and the optimal solution is the result of solving the Bellman optimality equations Puterman (2005):

$$\nu(s, \pi) = max_a \{r(s, a) + \lambda \sum_{j} \sum_{s'} \sum_{o'} p\{s' \mid s, a\}O\{o' \mid s'\}\nu(s', \pi')\} \qquad (3.7)$$

where $\lambda \in [0, 1)$ is the discount rate. The result will be the optimal screening program suggested by the model (see Section 5).
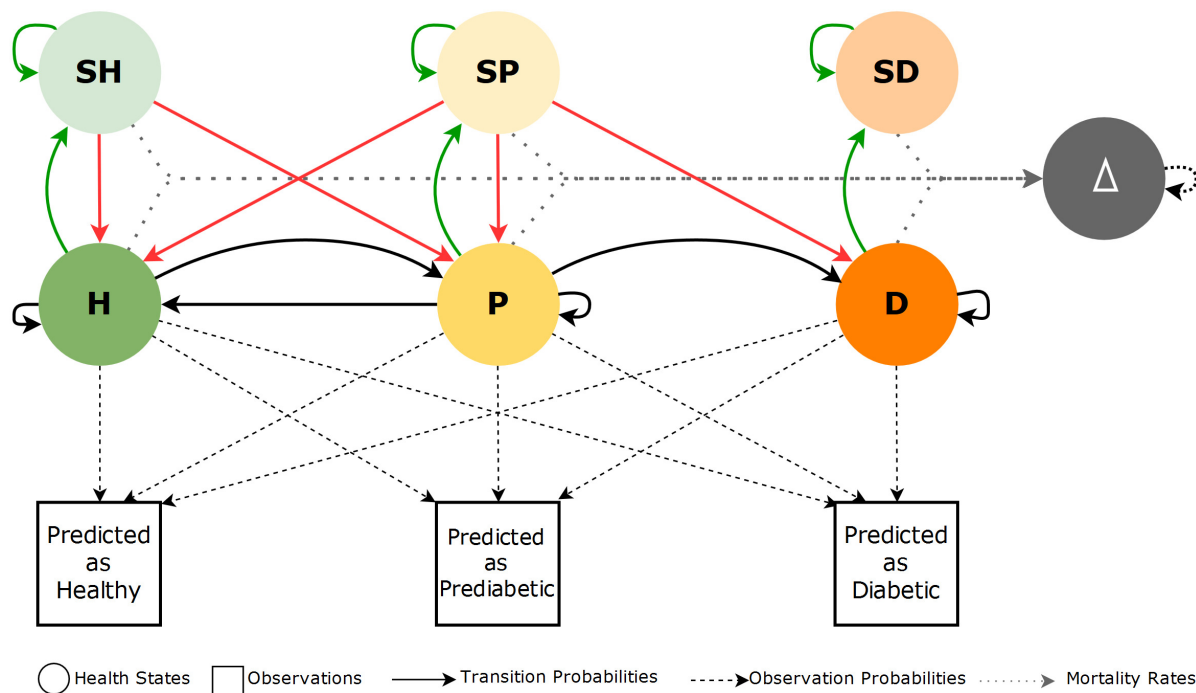
Figure 3.3: Underlying health states and observations of our POMDP model and the transitions among them. Only possible transitions are shown and those, which are not likely such as the transition from Diabetic to Healthy, are not depicted. Black arcs correspond to the natural progression of disease, green arcs correspond to the screening decisions, and red arcs correspond to reversion from screend states to uncontrolled ones.

3.4. Hidden Markov Models

A significant issue with using models that are based on transitions between unobservable states is how to estimate transition probabilities reliably from data. In our case, the states H, P, D are not directly observable, but the POMDP models needs transition probabilities between these states. The available data are provided by patient histories where at some point in time a diagnosis of prediabetes or diabetes is made, typically via a HbA1c screening lab test. We assume that up to the point in time when the diagnosis is made no significant medical intervention is performed and that the lab test reveals the true state (with some error). Transition probabilities between the unobservable states can be estimated from such data using a Hidden Markov Model (HMM), where the word hidden is used here for the fact that the true disease states are unobservable.

A HMM is a sequence of random variables X_t for time $t = \{1, 2, ...\}$ representing the hidden state with 3 possible values H, P, D and a sequence of associated random variables Y_t whose realizations of the 3 possible values SH, SP, SD represent observations. There are two types of parameters associated with HMMs: the transition probabilities between two unobservable states given by the transition matrix

$$M = \{m_i j\} = P(X_t = j \mid X_{t-1} = i), \tag{3.8}$$

and the probabilities that indicate the likelihoods that a certain hidden state will lead to a specific observation in the form the emission probability matrix

$$N = \{n_j(y_t)\} = P(Y_t = y_t \mid X_t = j). \tag{3.9}$$

The initial state distribution for $t = 1$ is defined as $q_i = P(X_1 = i)$. The aim is to estimate the parameters of the hidden Markov chain, $\sigma = (M, N, q)$ from observational data. The standard estimation procedure for HMMs is the Baum-Welch algorithm which utilizes the Expectation–Maximization iterative algorithm in order to find the maximum likelihood estimate of the parameters of the model given a set of historical observations Huang et al.

(2001). The transition matrix M provides a data-driven estimate for the transitions between the unobservable states in the POMDP specific to the cohort under consideration.

3.5. The Predictive Risk Model

Predictive risk models are powerful tools that can contribute to the decision-making process especially in the field of medical decision making. PRMs are usually multivariable, using several patient risk factors that are used to predict an outcome such as patient's status. These models can be utilized in many different ways including identifying those who are at an increased risk of having an undiagnosed condition to target healthcare interventions or lifestyle changes to.

Instead of using different risk factors directly in the update of the belief state in the POMDP, we propose to use a predictive risk model (PRM) to generate personalized predictions (used as observations) for the POMDP. Using a PRM offers many attractive features including a wide selection of available classification methods, a simple and efficient learning process, the possibility of data-driven feature selection, and the availability of methods that deal with missing data. These are very important advantages for working with electronic health record data, where the amount of information available for each patient can vary substantially.

The PRM model is used to predict one of the $K = 3$ values for the response variable $G = H, P, D$ using a feature vector x. Here we consider multinomial regression, an extension of logistic regression for a response variable with multiple levels. The probability of value k is predicted by

$$P(G = k \mid X = x) = \frac{e^{\beta_{0k} + \beta_k^T x}}{\sum_{l=1}^{K} e^{\beta_{0l} + \beta_l^T x}} \tag{3.10}$$

and the value with the highest probability is used as the prediction. The parameters are estimated from N observations y_i using

$$min_{\beta_0,\beta} \frac{1}{N} \sum_{i=1}^{N} l(y_i, \beta_{0k} + \beta_k^T x_i) + \lambda \parallel \beta \parallel_1, \qquad (3.11)$$

where the function l calculates the negative log-likelihood contribution of observation i, and the last term is used for $L1$ regularization.

Predictive models make classification errors. For example, a healthy patient may be misclassified as having prediabetes. These errors can be assessed using standard cross-validation techniques and are typically summarized in a so-called confusion matrix. Since we use predictions as observations o and the correct classification is given by the unobservable state s, the confusion matrix can be used as an estimate for the observation matrix O.

3.6. Parameter Estimation

In this section, we will first describe the data used in this research, and then we provide explanations on how we estimated each set of parameters using the techniques previously introduced and described in this chapter.

3.6.1. Data Description

The data used in this part of the research comes from the Electronic Health Record (EHR) of a large, integrated safety-net health system. Our cohort consists of patients from the Parkland Health & Hospital System, who are at risk for diabetes but have not been diagnosed with diabetes at the time of cohort entry. The cohort period is 2010 to 2014, during which individuals in the cohort may be diagnosed with diabetes. We retain patients, who have been diagnosed with diabetes, follow them over time, noting that additional information has been collected on them after their diabetes diagnosis. The cohort includes established primary care patients with an index visit occurring between January 1, 2012, and June 30, 2013, and 2 or more completed outpatient visits between the index visit and December 31, 2014. Patients are between 18 and 64 years of age at cohort entry. We exclude prevalent diabetes

Table 3.1: Key characteristics of the cohort studied

		Entire Cohort (N=12071)	Normal Glycemia (N=4883)	Diabetes (N=1314)
Age, years (SD)		47.49 (10.5)	45.18 (11.1)	50.02 (8.9)
Female, %		69.9	69.8	68.2
Race/ethnicity	Non-Hispanic White	13.3	15.2	10.2
	Black	39.8	35.2	44.7
	Hispanic	42	45	40.4
	Other	4.9	4.7	4.7
Education, years, mean (SD)		8.73 (3.3)	9.06 (3.2)	8.15 (3.4)
BMI, kg/m2, mean (SD)		31.37 (7.4)	29.74 (6.8)	35.2 (8.1)
Primary payer, %	Charity	40	39	41.7
	Private	13.2	13.1	11.8
	Medicare/Medicaid	26.7	25	31.2
	Self-pay	20	22.6	15.3
Lab values, mean (SD)	Random Glucose	97.48 (17)	93.16 (12.9)	112.4 (27.3)
	HDL-C	51.74 (15.5)	53.65 (15.5)	47.57 (13.8)
	LDL-C	193.52 (38.6)	190.95 (38)	195.87 (39.4)
	Triglycerides	146.35 (99.4)	135.52 (86.6)	173.62 (136.3)
	Systolic BP	129.11 (15.7)	126.01 (15.6)	135.44 (15.5)
	White Blood Count	7.39 (2.7)	7.34 (2.7)	7.71 (2.2)
	Ferritin	140.06 (322.1)	140.26 (360.3)	150.95 (312.7)
Tobacco User, %	Yes	12.2	12.8	12
	Never	69.5	71.3	66.1
	Passive	1.8	1.9	1.7
	Quit	16.4	14	20.1
Alcohol User, %		10	9.9	10.2
Family history DM, %		71.1	74.8	62.1
Hypertension, %		46.9	38	62.6
CHF, %		2.3	1.7	3.8
Cardiovascular Disease, %		22.8	17.8	29.5
Medication use, %	Steroids	18	18.6	17.6
	Anti-hypertensives	45.5	37.4	61.5

and gestational diabetes. We excluded patients diagnosed with diabetes and prediabetes on or 18 months before the index visit using ICD-9-CM encounter codes, problem list diagnoses, and laboratory results (A1c, fasting glucose, oral glucose tolerance tests) meeting diagnostic criteria . Table 3.1 provides summary statistics.

We estimate various parameters of our model using the data described in Table 3.1. We reiterate that our goal is to provide an age-specific screening policy; some parameters such as mortality rates are estimated for various age ranges.

3.6.2. Estimating Transition Probabilities

We estimate the transition probabilities for the POMDP, using patients' historical data (screening results from the EHR) as an input for the HMM. Screening results can be subject

to error and thus the true health status of the patients is not observed directly but through realizations or observations which are the LAB results. These trajectories are all discrete observations recorded irregularly in time with varied intervals (due to irregular visits) and are not always complete. The trajectory (P,*,*,*,*,P,*,*,D,*,D) can be an example of a sequence of observations for a single patient during 11 years where * are representatives of missing values (the years at which the patient did not show up or was not screened).

HMM uses the above-mentioned trajectories as an input for an iterative algorithm and tries to find the transition probabilities that best fit the input sequences. The outcome of HMM will be a transition probability matrix that best fits our data. This is shown in matrix P.

$$
p = \begin{pmatrix}
0.9438 & 0.048 & 0 & 0.0082 \\
0.0328 & 0.9242 & 0.0348 & 0.0082 \\
0 & 0 & 0.9916 & 0.0084 \\
0 & 0 & 0 & 1
\end{pmatrix} \tag{3.12}
$$

Note that it is impossible to go directly from Healthy to Diabetic since all the patients will experience prediabetes by definition before getting into diabetes. There is also no regression from diabetes to prediabetes and as soon as the patient gets into diabetes will have to stay there.

The three elements of the matrix P that cannot be estimated from our data are age-specific mortality rates for healthy, prediabetes and diabetes states. We obtain this from the National Center for Health Statistics that reports the age-adjusted death rate of 0.0072 (2016 numbers) for standard population considering all possible causes of death Kochanek et al. (2016) . We also estimate the age-specific death rates for age groups starting from 15 as well as the mortality rate for diabetes for the total population.

In order to estimate the age-specific mortality rates we define three random variables X, Y, and Z, where X is the random variable indicating the age of the patients in the cohort, Y is the random variable representing the age at the death of the patients, and Z is the difference

between these two random variables, i.e. $Z = Y - X$. Based on Kochanek et al. (2016) the life expectancy for the U.S. population in 2016 is 78.6, i.e. $E[Y] = 78.6$. Also from our dataset we have $E[X] = 47.5$ (Table 3.1). Thus $E[Z] = E[Y - X] = E[Y] - E[X] = 78.6 - 47.5 = 31.1$ would be the expected remaining life for the whole cohort. The quantity of interest is actually $p\{Z = 1\} = P_1$ for each specific age that translates into the mortality rate per each year assuming the mortality rate is stationary. Thus, we have:

$$p\{Z = n\} = (1 - P_1)^{n-1} P_1, \tag{3.13}$$

assuming each year is independent of other years. (3.13) is the probability distribution function (pdf) of the Geometric distribution where Z is the number of independent trials until a failure occurs (in this case, death). Thus, based on the Geometric distribution, $E[Z] = 1/P_1$. This way we can calculate the death rate per year P_1 for each age assuming $E[Y] = 78.6$ using $E[Z] = E[Y - X] = E[Y] - E[X] = E[Y] - X$. We calculate this rate for healthy, prediabetes and diabetes patients and show them with $P_{(}1, H)$, $P_{(}1, P)$, and $P_{(}1, D)$ respectively. It is worth highlighting that the life expectancy, $E[Y]$, varies for patients with diabetes. Based on a 2010 report by the Diabetes UK, type 2 diabetes reduces the lifespan by 10 years Key statistics on UK (2010). Another study claims that for people over 55, type 2 diabetes reduces lifespan for an average of 6 years for women and 5 years for men Loukine et al. (2012).

3.6.3. Estimating Observation Probabilities

We extracted more than 40 features from the electronic health records and addressed missing data using multiple imputation de Goeij et al. (2013). All features are scaled to z-scores and multinomial regression with L1 regularization is applied. Table 3.2 gives the parameters for the strongest 10 predictors converted to odds-ratios for the class diabetes against health and prediabetes. The AUC column shows the area under the receiver operating characteristic curve (diabetes against healthy and prediabetes) achieved by adding more and more features.

Table 3.2: Top 10 features of the proposed regularized multinominal regression model

Feature	OR	AUC
MEAN RANDOM BLOOD GLUCOSE LEVEL	1.67	65.53%
BMI	1.40	68.50%
SYSTOLIC BP	1.14	71.17%
HYPERTENSION	1.04	72.10%
FAMILY HISTORY	1.19	72.10%
HIGH DENSITY CHOLESTEROL	0.85	72.60%
AGE	1.19	72.87%
BLOOD PREASURE MEDICATION	1.06	72.87%
CHOLESTEROL MEDICATION	1.09	73.15%
CHOLESTEROL HDL RATIO	1.02	73.42%

The observation matrix O needed by the POMDP is estimated using the risk model's confusion matrix obtained via ten-fold cross-validation. For example, the observation probability that a healthy patient will be classified as having prediabetes $O(P, H)$ is the estimated classification error of the model. The estimated observation matrix is given as follows.

$$O(o \mid s) = \begin{pmatrix} 0.8 & 0.15 & 0.05 \\ 0.15 & 0.7 & 0.15 \\ 0.05 & 0.25 & 0.7 \end{pmatrix} \qquad (3.14)$$

The observation matrix for a perfect predictive model would have a probability of 1 along the diagonal of the matrix and zero otherwise.

3.6.4. Estimating Rewards

Table 3.3 lists the values for these parameters and the cost of a diabetes screening test in the U.S. All reward parameters and constants are annual and are based on estimates found in the literature (see column source in the table). All costs are estimated from the societal perspective.

Table 3.3: Parameters associated with the reward function of the POMDP model

Parameter	Description	Source	Value
C_s	Cost of a diabetes screening test	Chatterjee et al. (2013); Zhang et al. (2003); O'connor et al. (2001); Kahn et al. (2010)	$8,346 [1]
Q	Quality-Adjusted Life Year in U.S. dollars	Neumann et al. (2014a)	$50,000
C_D	Direct medical costs per year for a new-onset diabetes	Chatterjee et al. (2013)	$4,174
C_P	Incremental direct medical costs per year for a patient with prediabetes	Chatterjee et al. (2013)	$1,316
α_P	Annual utility decrease of living with prediabetes	Ackermann et al. (2009); Neumann et al. (2014a)	0.16
α_{UD}	Annual utility decrease of living with undiagnosed diabetes	Ackermann et al. (2009); Zhang et al. (2012a); Bahia et al. (2017); Zhang et al. (2012b)	0.2
$\alpha_D D$	Annual utility decrease of living with diagnosed diabetes	Ackermann et al. (2009); Zhang et al. (2012a); Bahia et al. (2017); Zhang et al. (2012b)	0.18
m_T	Age-Adjusted mortality rate in U.S. in 2016	Kochanek et al. (2016); Murphy et al. (2015)	0.0084
m_D	Age-adjusted mortality rate for Diabetes in 2016	Kochanek et al. (2016); Murphy et al. (2015)	0.00021
l_e	Life expectancy for the U.S. population in 2016	Kochanek et al. (2016)	78.7
l_d	Lifespan decrement due to Diabetes	Loukine et al. (2012)	5
u_r	Uptake rate of Diabetes screening [2]	Khunti et al. (2015); Park et al. (2008); Orton et al. (2013); Davies (1999); Eborall et al. (2012)	0.644

3.7. Optimal Screening Policy

There are various algorithms to solve POMDP problems. Details on how to solve POMDPs and a survey on POMDPs solution methods are beyond the scope of this research and can be found in Lovejoy (1991a,b). In this paper we will apply a popular grid-based approximation Hauskrecht (2011); Ahuja and Birge (2018) called the finite grid method Sandikci et al. (2013); Cassandra and Rocco (1998).

When the solution of the POMDP problem converges, then we can create a finite state controller from the value function's partitioning of the belief space. Using this controller, the decision maker can execute the optimal policy without the need to track the actual belief states. The controller is a graph where nodes are representatives of the belief states and arcs represent updates of the belief state due to new observations. As an example of this graph, the optimal policy for patients of age 55, solved with a coarse grid is depicted in Figure 3.4. We use a coarse grid, since it results in a smaller number of belief states and a graph that is easier to visualize and interpret. The initial node is determined by the prior belief about the health status of the patient. For example, assuming that the prevalence of diabetes and prediabetes among the patients in our cohort is 10 percent and 20 percent, respectively, the

---

[3]The authors of this paper have also developed an R (Team (2018)) package called "pomdp" which provides an interface to pomdp-solve, a solver for Partially Observable Markov Decision Processes (POMDP) Kamalzadeh and Hahsler (2019)
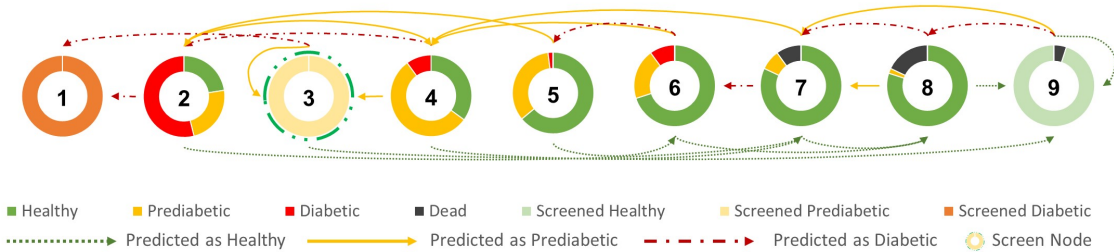
Figure 3.4: Policy graph for patients of age 55 solved with a coarse grid

initial belief state will be $\pi_1 = (\pi_1(H) = 0.7, \pi_1(P) = 0.2, \pi_1(D) = 0.1)$ or node 6 in Figure 3.4. In other words, the decision maker will assume, based on the existing prevalence rates (and with no other information) that the patient is in node 6 when she shows up for the first time. As additional information on a patient becomes available (e.g., blood pressure, BMI or symptoms of a DM complication), the predictive model will create a prediction which is used as an observation to update the beliefs, or in other words, the state of the world (represented by nodes) changes. We have arranged the graph such that the belief about the decease severity increases from right to left and Predictions as Healthy (PH) move the patient to the left, while Predictions as Diabetic (PD) move the patient to the right.

The only belief state where the optimal action is screening is state number 3. From there the patient can go to the best state 9 (screened healthy), state 1 (screened diabetic), or stay in state 3 (screened prediabetic). This implies that the optimal choice for patients of age 55 that are screened as prediabetic is to rescreen them in the next period since they remain in the screening state.

By using a finer grid, we can create decision graphs with many more belief states, however, the visualization of the decision graph becomes more and more difficult to read. We can visualize the belief space as a ternary plot and place a belief states in that space. Figure 3.5

(a) shows how a single belief state is located in the plot. Figure 3.5 (b) shows all the belief states using a very fine grid for patient of age 55. Belief states where the optimal action is screening are colored red. It can be seen that the decision boundary between screening and not screening states can be approximated by a straight line through the belief space. The patient should be screened whenever the belief about the patient's health falls below the line. A patient can be assigned to a belief state that indicates screening because the physician makes that determination during the first encounter or because the belief state for an existing patient is updated due to high-risk observations created by the PRM.

Since prevalence of diabetes is age dependent, also the optimal screening policy is age dependent. We calculate the optimal policy graphs for different age groups, find the linear separation between screening and not screening states and just place the linear separation lines in the Figure 3.5 (c). As the patient's age increases the decision boundaries move upper from the triangle base. This indicates that the model is trying to reduce the risks as the patients get older. For example, it is not optimal to screen a 60 years old patient with a belief state of (40, 50, 10), but the same patient should be screened the next year. It shows that as the age increases, the model moves the screening thresholds in a way that even patients with lower risks get screened. For ages above 71, the policy the model provides is to screen at each period (always screen).
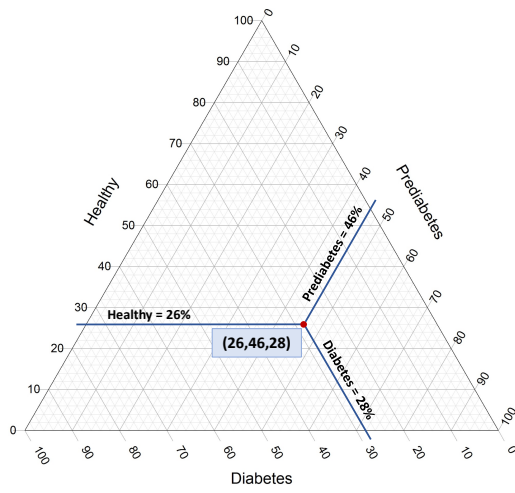
(a) how to read the graph
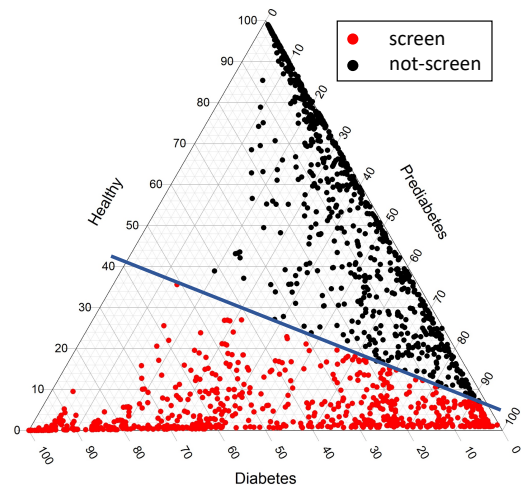
(b) decision boundary for age 55
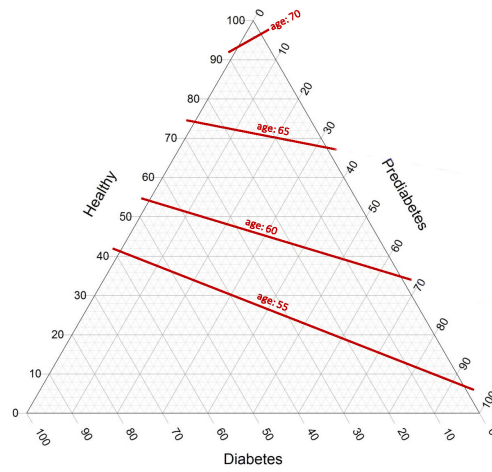
(c) decision boundaries for various ages

Figure 3.5: Ternary plots representing the decisions associated with each belief state.

3.8. Policy Implications And Evaluations

3.8.1. Simulation Model for Guidelines Evaluation

There exists a vast literature on comparing different screening policies using cost-effectiveness analysis Howard et al. (2010); Kahn et al. (2010); Chen et al. (2001); Chatterjee et al. (2013); O'connor et al. (2001); Hoerger et al. (2004). Most of the studies simulate a cohort of patients with specific parameters provided to evaluate the cost-effectiveness of either mass screening or opportunistic screening for type 2 diabetes. We apply here simulation to evaluate the effectiveness of the proposed screening policy and to compare it with existing screening guidelines. To simulate the natural progression of diabetes and its complications including retinopathy, nephropathy, and neuropathy we use the Markov models shown in Figure 8 and developed in Chen et al. (2001). Each node in the graphs represents one stage of disease, and the stages are arranged from least to highest severity as one goes from left to right. The costs associated with each stage of thee complications are taken from Howard et al. (2010); Chen et al. (2001).

The simulation consists of a hypothetical cohort of 50000 patients whose characteristics have been described previously in Table 3.1. Parameters such as transition probabilities in disease Markov models, utilities and costs, incidence, prevalence and mortality rates of each state of the disease progression models used in this paper are all taken from Chen et al. (2001); Kahn et al. (2010); Howard et al. (2010). We simulate seven different scenarios representing different screening policies including the proposed screening policy. Patients leave the simulation when they die or when they reach the maximum life expectancy. For the opportunistic screening policy, the patient's chance of getting screened or diagnosed with prediabetes, diabetes or its complications is limited by the number of times they visit the physician, either randomly for a blood test or due to observing a symptom of a complication, as in the case of existing guidelines. We also assume that early detection and treatment of prediabetes can lead a patient back to a healthy state again and reduces the chances of progressing into diabetes. Also, early detection and treatment of diabetes can reduce
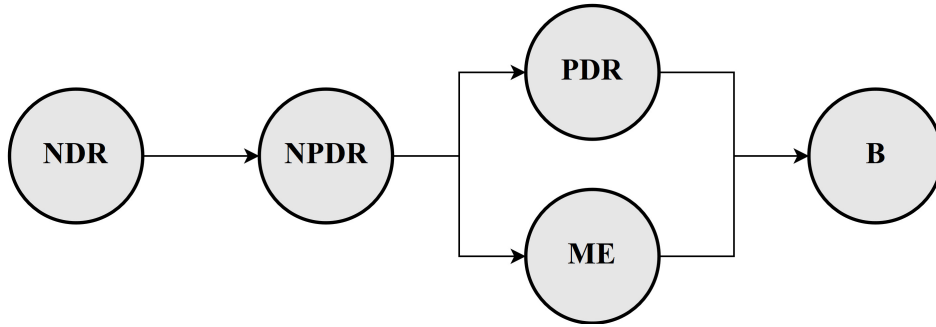
(a) Retinopathy: {NDR: No Diabetic Retinopathy, NPDR: Non-proliferative Diabetic Retinopathy, PDR: Proliferative Diabetic Retinopathy, ME: Macular Edema, B: Blindness}
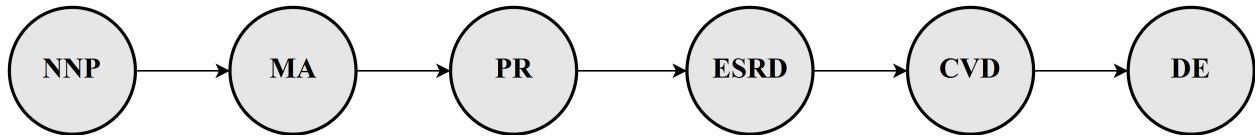
(b) Nephropathy: {NNP: No Nephropathy, MA: Microalbuminuria, PR: Proteinuria, ESRD: End Stage Renal Disease, CVD: Cardo Vascular Disease, DE: Death}
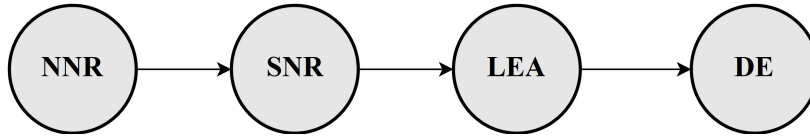
(c) Neuropathy: {NNR: No Neuropathy, SNR: Symptomatic Neuropathy, LEA: Lower Extremity Amputation}

Figure 3.6: Markov models for natural disease progression of diabetes complications

the patient's chances of developing complications or progressing into more severe stages of complications.

A detailed description of the conducted simulation can be found in the appendix.

3.8.2. Guidelines Evaluation

To compare the outcomes of the different simulated scenarios and thus the efficacy of the proposed policy with other screening policies, we report the metrics used in the literature. These metrics include ICER (incremental cost-effectiveness ratio), years gained, QALYs gained, diagnosis lead time, macrovascular events prevented, microvascular events prevented, and deaths prevented Chen et al. (2001); Kahn et al. (2010). To calculate the ICER, each systematic policy is compared with an opportunistic screening policy. Table 3.4 reports the outcome measures for the different screening policies.

Table 3.4: comparison between various screening guidelines in terms of cost-effectiveness, years and QALYs gained, diagnosis lead time and events prevented (from 50 replications)

Screening Policy [a]	ICER (cost per QALY, $US) [b] (SD)	Years Gained [c] (SD)	QALYs gained [b,c] (SD)	Diagnosis lead time [d,c] (SD)	Macrovascular events prevented [c] (SD)	Microvascular events prevented [e] (SD)	Deaths prevented [e] (SD)
SP30-3	$27,042 (1268)	0.75 (0.04)	2.04 (0.05)	19 (0.2)	22 (1.6)	207 (4)	48 (2)
SP45-1	$37,366 (1755)	0.62 (0.04)	1.18 (0.03)	14 (0.1)	21 (1.5)	178 (4)	45 (2)
SP45-3	$31,155 (1791)	0.61 (0.04)	0.96 (0.03)	11 (0.1)	20 (1.4)	165 (4)	44 (2)
SP45-5	$29,644 (2175)	0.60 (0.04)	0.86 (0.03)	9 (0.1)	20 (1.5)	157 (4)	44 (2)
SP60-3	$32,201 (2966)	0.59 (0.04)	0.60 (0.03)	6 (0.1)	19 (1.4)	142 (4)	42 (2)
SP30-1 (Maximum screening)	$36,801 (1233)	0.83 (0.05)	2.63 (0.05)	25 (0.2)	23 (1.5)	229 (4)	50 (2)
SP-POMDP	$20,426 (1339)	0.81 (0.04)	2.06 (0.05)	18 (0.2)	23 (1.5)	219 (5)	49 (2)

[a]Screening policies are abbreviated as SP AGE-REPEAT, where AGE is the age at the first screening and then the screening is repeated every REPEAT years until the patient reaches age 79.

[b]All costs and QALYs are discounted at 3% per year.

[c] per patient

[d]Mean lead time in diagnosis gained by each screening strategy compared with opportunistic screening.

[e]per 1000 patients

Table 3.4 shows that the proposed policy, SP-POMDP, performs better than all other policies in every metric except for the maximum screening policy, SP30-1. Compared to opportunistic-screening, SP-POMDP diagnoses prediabetes and diabetes patients on average 18 years earlier while this is only outperformed by the maximum screening policy. In terms of macrovascular events, microvascular events and deaths prevented, although the maximum screening policy produces slightly better simulation results compared to SP-POMDP, there is not a significant difference between POMDP and the maximum screening guideline. SP-POMDP achieves very similar outcomes to SP30-1, but reduces the cost per gained QALY significantly from more than $36,000 to less than $20,500.
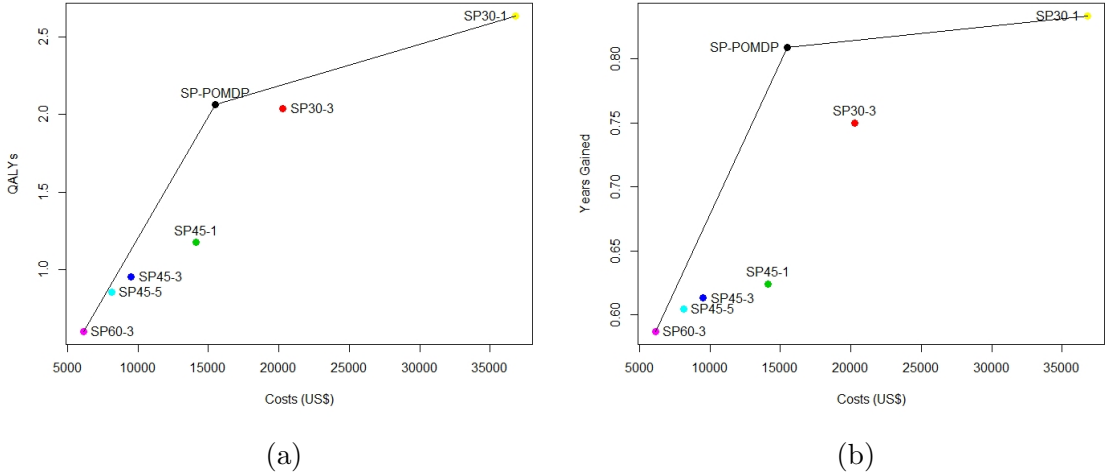
Figure 3.7: (a) Costs per QALY and (b) cost per years gained of seven screening guidelines compared with opportunistic screening in terms of QALYs and years of life gained. The efficient frontier is shown as a line.

Figure 3.7 compares the screening policies in Table 3 using the notion of the efficient frontier. Figure 3.7 shows that the efficient frontier is comprised of SP30-1 (maximum screening), SP-POMDP, and SP60-3, indicating that all other policies are inefficient. The SP-POMDP is slightly inferior to SP30-1 but it takes significantly less resources, less time from the patients, less lab tests, and cause less stress for patients.

44

3.8.3. Sensitivity Analysis of the Simulation Model

To evaluate the sensitivity of our model to the parameters introduced in Table 3.3, we performed two different sets of sensitivity analyses. The first set is designed to evaluate the sensitivity to the cost parameters including the cost of the diabetes screening test C_s, the annual utility decrease of living with prediabetes α_P and the annual utility decrease of living with diabetes $\alpha_U D$, $\alpha_D D$. The second set is designed to evaluate the sensitivity to the probabilities including transition probabilities P and observation probabilities O. The same output measures as in Table 3.4 are used to evaluate the sensitivity of the model to these parameters. Table 3.5 shows the result of the performed sensitivity analysis.

Table 3.5 shows that the proposed SP-POMDP model is robust with respect to most parameters. The model is sensitive to the costs of screening which affects cost-effectiveness, QALYs gained, and diagnosis lead time. As we expect, the higher the cost of screening, the less cost-effective (higher ICER) the policy would be and the opposite. As the cost of screening increases the proposed model move toward postponing the screening as much as possible which results in lower QALYs gained and an increase in diagnosis lead time.

The value of a QALY affects the policy with lower values of a QALY representing a lower rewards in the POMDP reward structure which will result in postpone screening to accumulating more costs and gaining fewer QALYs as well as decreasing the diagnosis lead time. On the other hand, higher values of QALY translates into higher rewards in the POMDP model which in turn translates into more cost-effective policies with more QALYs gained and longer diagnosis lead time.

Table 3.5 also shows that higher screening uptake rates also result in more cost-effective policies with more QALYs gained and longer diagnosis lead time. This is because higher uptake rates will result in higher rewards for screening healthier people, and then policies that screen people earlier (longer diagnosis lead time), thus bring about more gained QALYs. Lower transition probabilities between health states means patients take more time to develop diabetes. This results in policy graphs with more nodes and screenings can be postponed longer without producing higher cost and shorter diagnosis lead times. The

45

Table 3.5: Sensitivity analysis for varying each paramter by +/-20%

Parameters		ICER ($US)	Years Gained	QALYs gained	Diagnosis lead time	Macrovascular events prevented	Microvascular events prevented	Deaths prevented
Reference		$20,426	0.81	2.06	18	23	219	49
Costs of screening (C_s)	-20%	$17,751	0.82	2.57	22	23	228	49
	20%	$28,669	0.8	0.92	5	22	202	48
QALY	-20%	$30,412	0.8	0.91	5	21	199	48
	20%	$17,895	0.81	2.53	22	23	224	50
Treatment costs (C_P, C_D)	-20%	$20,878	0.8	2.06	18	23	219	49
	20%	$20,336	0.82	2.06	18	23	216	50
Disutilities ($\alpha_P, \alpha_{UD}, \alpha_{DD}$)	-20%	$20,500	0.83	2.07	18	23	216	50
	20%	$20,416	0.79	2.05	18	23	219	48
Uptake rate of screening (u_r)	-20%	$20,782	0.83	2.07	18	23	216	50
	20%	$18,136	0.81	2.54	22	23	226	50
Transition Probabilities [a]	-20%	$28,449	0.8	0.91	5	23	199	49
	20%	$18,355	0.82	2.56	23	23	226	49
Observation Probabilities [b]	-20%	$20,254	0.8	2	17	22	215	49
	20%	$17,826	0.81	2.59	23	23	230	49

[a] All transition probabilities except the recurring ones
[b] 20% more accurate predictive model and 20% less accurate predictive model

46

opposite happens on the hand when transition probabilities increase. We also see more cost-effective policies with more QALYs gained as well as longer diagnosis lead times when a more accurate predictive model (better observation probabilities) is available.

Chapter 4

Resource Allocation Under Uncertainty for Emergency Situations (P2)

The Emergency Response Bureau and Special Operations of the Dallas Fire and Rescue Department (DFRD) encompasses two operational divisions. One of these divisions, Emergency Response, is responsible for the day-to-day operations involving normal fire suppression and emergency first responder calls. The Emergency Response Division provides the residents of Dallas with fire suppression and protection, emergency rescue capabilities, and emergency medical first responder services. Customer Service is DFRD's primary goal for the citizens of Dallas, to be obtained through providing safety, mitigating emergency situations, and reducing loss of any kind.

An important question for the DFRD is whether resources should be moved around in the city to cover areas where the resources are currently responding to an incidence. Every time an incident happens resources in a particular zone of the city will be dispatched and become unavailable for several hours. If another incident happens in that zone during that time, resources from other areas of the city will need to respond which will increase response time. In this chapter I will describe the application of the decision framework to the problem of finding an optimal resource allocation policy.

4.1. The need for a resource reallocation policy

A Battalion is a combination of several fire stations working together to deal with situations. They are spread out in the city in a way that each Battalion covers a specific area of the city. The area that we define as city zone, includes several fire stations all working together under the same Battalion. Battalions are responsible to respond to incidents

happening in the city zone under their control. The DFRD battalions, the territory under their observance and their fire stations are shown in Figure 4.1.

Although each battalion is supposed to respond to the incidents happening in its territory, there are situations in which on battalion does not have enough resources to respond to a new incident since all or the majority of the resources of the zone are already dealing with other incidents. Using the data available from the DFRD, we can analyse if there are situations which one zone does not have enough resources to respond to incidents happening in it and asks for extra resources from other zones (battalions) of the city. We visualize some examples in Figure 4.1.

To be able to see these situations, we have focused only on one zone (battalion) of the city of Dallas and this is the downtown area to which battalion 1 responds.

The small purple dots that form lines from other zones to zone 1 are emergency vehicle locations during the dispatch process from their original stations in other zones to incident locations in zone 1 (light pink area in downtown). We can notice that there are many instances of requesting resources from other battalions for incidents happening in zone 1 (Table 4.1).

From Zone (Battalion)	Total	Per Week	Average Response Time
3	13675	131	6.1
6	2663	25	6.76
9	1930	19	6.85
8	1682	16	6.83
5	792	8	6.58
4	285	3	9.11
7	181	2	7.4
2	103	1	6.95

Table 4.1: Dispatches from other zones to zone 1 from 2015 to 2017

These happen at the moment the incidents happen and the time it takes for the resources to get to the incidents' locations from other zones are significantly larger than usual (Table 4.1).
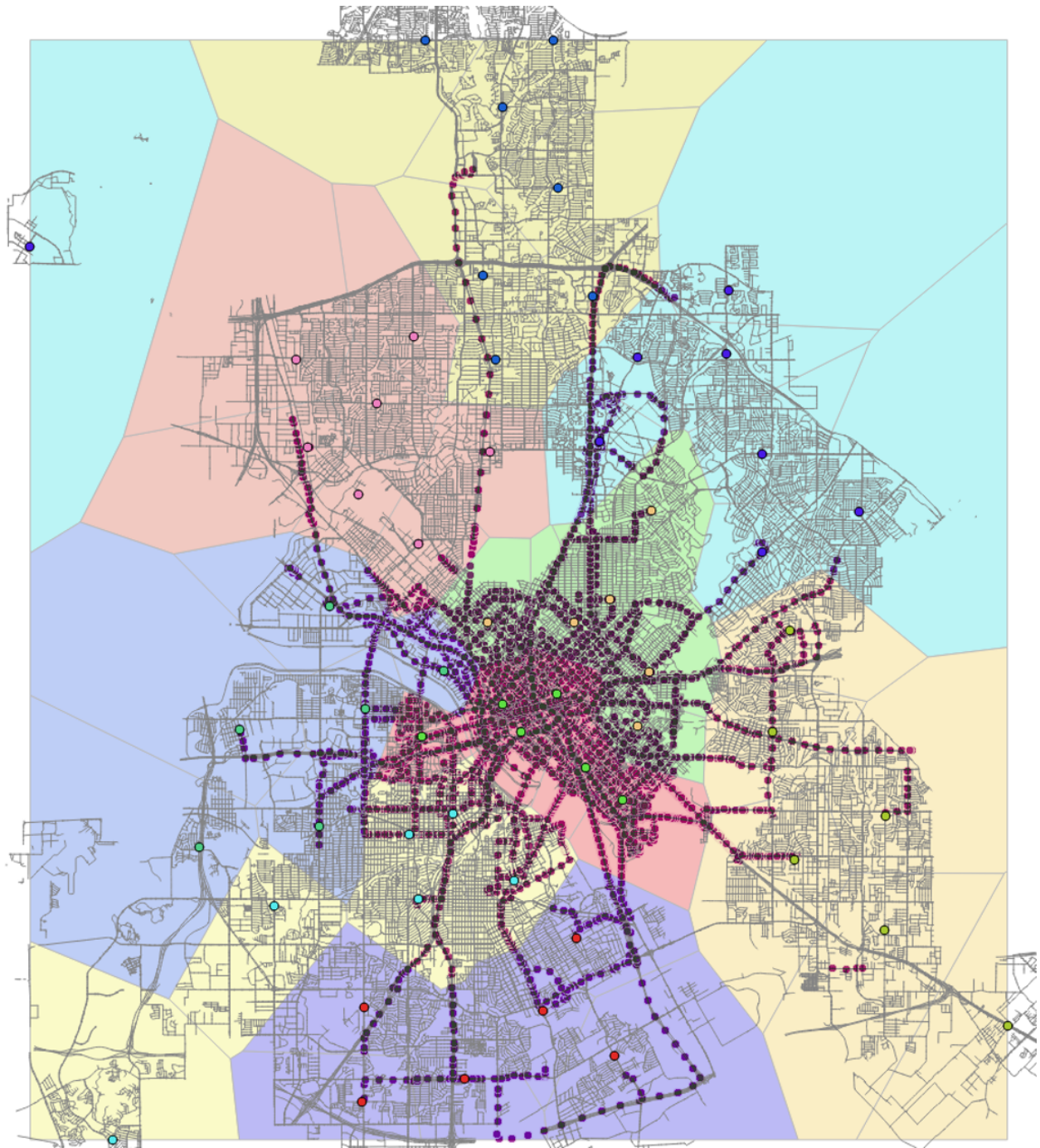
Figure 4.1: AVL of vehicles from other zones responding to incidents in zone 1

The average response times shown in Table 4.1 are way above the limit DFRD is trying to keep its response times under (6 minutes) and in emergency situation not just minutes but even seconds count since the lives of people are in danger. With a proper reallocation policy that is implemented in advance, these response times can be cut shorter.

4.2. The Partially Observable Markov Decision Process Formulation

In order to apply the framework proposed in chapter 2, we need to formulate the problem as a POMDP problem. To do so we focus on one city zone in order to define our state space. We try to capture the availability of the resources in that zone in the near future; i.e. the states of our model are trying to describe the status of a particular zone in the near future. This near future depends on how long in advance we want to be aware of the future and also the risk associated with the problem. In this research we define near future as 15 minutes from the current time. By defining what would be the status of a zone in the near future as the state of the model, we add unobservability to the model and since only in the future we will for sure know the answer to this question, the state space is always unobservable to us.

The two main variables that affect the state of the system are: the availability of the resources in the near future, and the road condition in the near future. The model has therefore a two-dimensional state space where one dimension is the remaining system capacity, and the other is the road condition. We have no control over the road conditions but the system capacity is fully under our control. Below, we define and introduce all the components of the POMDP model.

4.2.1. State Space

The state sapce of the POMDP has two dimensions. One dimension is the remaining capacity of the system (the proportion of the resources available and ready for dispatch). We categorize this dimension into 3 categories called, low capacity, medium capacity, and high capacity; which translates into what proportion of the resources is available or would be available within a short period of time from now (near future). Thus $S_{cap} = \{low, med, high\}$.

The other dimension of the state space represents the road condition. Here we only assume traffic, but other conditions such as weather conditions can be used as well. We categorize this state into normal hours or no major traffic on the road, and rush hour or major traffic on the road; the latter corresponds to higher travel times. Thus $S_{road} = \{no - traffic, traffic\}$.

Considering these two dimensions for the state space, the state space will have a total of six different states represented in Table 4.2 and Figure 4.2.

State space (S)		Remaining System Capacity		
		High	Med	Low
Road Condition	No traffic	$s_{N,H}$	$s_{N,M}$	$s_{N,L}$
	Traffic	$s_{T,H}$	$s_{T,M}$	$s_{T,L}$

Table 4.2: POMDP model state space and dimensions
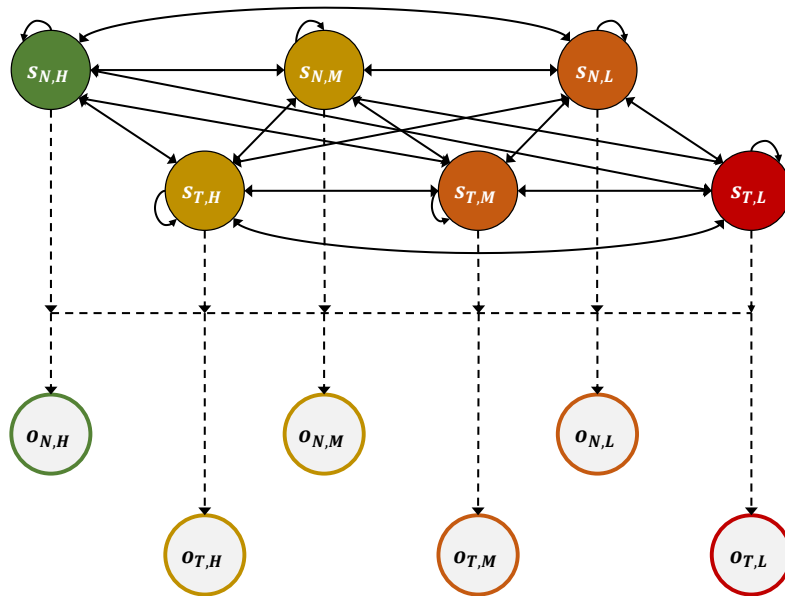
Figure 4.2: POMDP model, state space and observation space with arcs

4.2.2. Action Space

There are two available actions for each state of the system: one is to (1) Ask for extra resources from other zones that can provide the requested resource and the other one (2) is not to ask for any extra resources (the second action is simply not doing anything until the next epoch). Thus the action set would be $A = \{a_{ask}, a_{nothing}\}$.

4.2.3. Observations

We define one observation for each state of the system. These observations would be produced using a predictive model from a combination of signals that come directly from the system. Each observation points to one state of the system. The set of observation then would be $\Omega = \{o_{N,H}, o_{N,M}, o_{N,L}, o_{T,H}, o_{T,M}, o_{T,L}\}$.

4.2.4. Cost Structure

The costs in this POMDP model associated with each action and state are in the form of travel times. Thus $c(s, a)$ is the average travel time of emergency vehicles given the system is in state s and action a is taken.

4.3. Parameter Estimation

In this section we use the same methods previously introduced in chapter 2 to estimate all the parameters associated with the POMDP model.

4.3.1. Data Description

The data used in this application of the research is provided by the Dallas Fire and Rescue Department. The data mainly consists of the Automated Vehicle Locations (AVL) of the emergency vehicles used by DFRD, information on incidents responded to by the DFRD and the status of the vehicles used. Below is a brief description of each of these data components.

AVL data: The AVL data consists of the locations of the vehicles dispatched to the incidents recorded at various points in time during the response time. The AVL data contains:

- Incident number: a unique ID assigned to each incident. This column indicates what incident the vehicle was dispatched for.

- Radio name: a unique ID that indicates the type of the vehicle dispatched.

- Date and Time: the exact date and time of the recorded vehicle location.

- Latitude and Longitude: the exact coordinates of the vehicle location at the recorded moment.

- Heading: a number between 0 and 360 indicating the heading of the vehicle at the moment.

- Speed: the speed of the vehicle in miles per hour at the recorded moment.

Status data: The status data contains information about the exact date and time of the changes in the status of the vehicles dispatched to the incidents. The status data contains:

- Incident number: a unique ID assigned to each incident. This column indicates what incident the vehicle was dispatched for.

- Radio name: a unique ID that indicates the type of the vehicle dispatched.

- Assigned: date and time of the moment the vehicle was assigned.

- Enroute: date and time of the moment the vehicle started moving.

- Arrived: date and time the vehicle arrived to the incident location.

- Cleared: date and time the vehicle cleared the situation.

Incidents data: The status data contains information about the exact location, date and time of the incidents occurred. The incidents data contains:

- Incident number: a unique ID assigned to each incident.

- Response date and time: the date and time of the occurrence of the incident.

- Address: the address where the incident happened.

- Postal code.

- Longitude and Latitude: the coordinates of the location of the incident.

Table 4.3 provides some basic information on the data provided by the DFRD.

Distribution of the incident length defined as the time between assigning a vehicle and clearing the situation is depicted in Figure 4.3. Although the majority of the incidents get cleared in less than an hour, there are still significant number of incidents that take longer than that. These incidents will keep the system busy for longer than normal incidents and thus bring down the remaining capacity of the system.

The data can also provide us with the distribution of the remaining capacity of the system. Figure 4.4 shows this distribution for zone 1 of the city. Although the majority of the time, the system has high remaining capacity, still there are quite significant number of times where the remaining capacity of the system hits lower than usual. As we can see the distribution has a long tail on the left side indicating that the remaining capacity can get very low at some times.

4.3.2. Estimating Transition Probabilities

To estimate the transition probabilities, we need to use data to estimate the system state at each point in time. since the state space is 2 dimensional, we need to first estimate each dimension and then combine the two dimensions to get the actual state of the system.

For the road condition dimension of the state space, we use the rush hours from Dallas city. If the system is in rush hours then the road condition is experiencing traffic and otherwise no traffic.

Table 4.3: Basic information on the data provided by the DFRD

AVL data		Status data		Incidents data	
Data time span	1 month	Data time span	25 months	Data time span	25 months
Number of records	1,048,575	Number of records	931,940	Number of records	535,724
Number of incidents	20414 (680 per day)	Number of incidents	538,170 (717 per day)	Number of incidents	535,724
Vehicles used	149	Vehicles used	309		
Time between AVLs	20 seconds	Average time from Assigned to Enroute	1.01 minutes	Average time between incidents	14.85 minutes
Average speed	26.5 mph	Average time from Enroute to Arrived (Response time)	5.06 minutes	Zip code with the highest incident rate	75216 (Downtown area)
		Average time from Arrived to Cleared (Clearance time)	23.17 minutes		
		Average time from Assigned to Cleared	29.2 minutes		
Average number of times each vehicle is used	30 (once per day)	Average number of times each vehicle is used	3015 (4 times a day)		
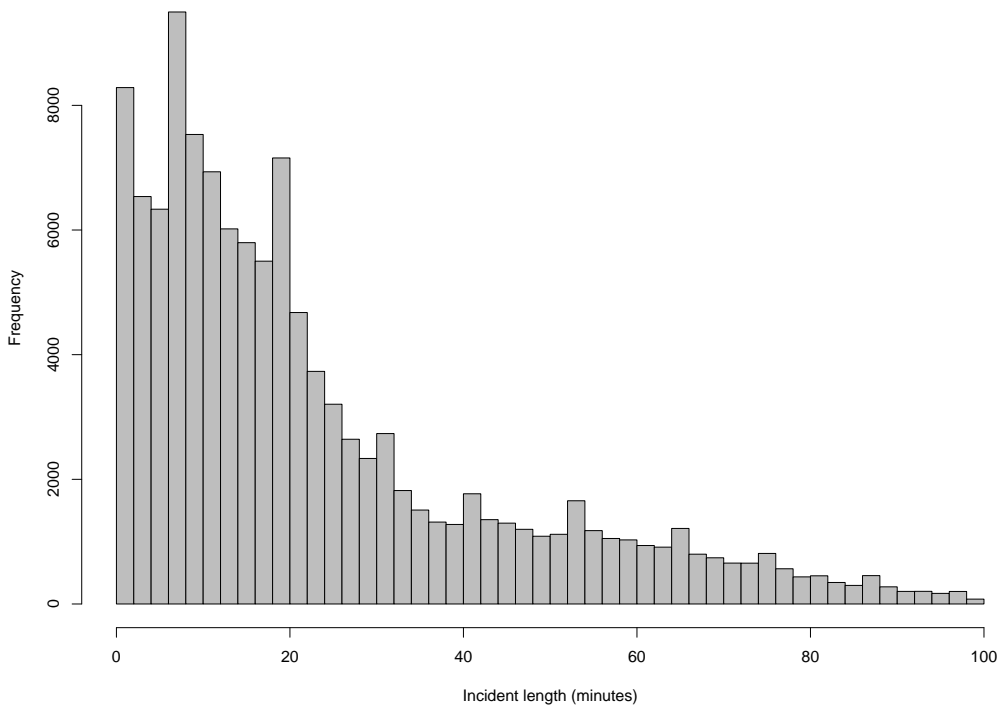
Figure 4.3: Incident length distribution for zone 1 (minutes)

For the remaining system capacity we use the number of emergency vehicles that are available at the moment.

Given the above information we can calculate the transition probabilities. Tables 4.4 and 4.5 represent the transition probabilities for action a_{ask} and action $a_{nothing}$ respectively.

4.3.3. Estimating observation probabilities

Same as the previous application, we develop a predictive model to produce both observations and observation probabilities. The outcome of the predictive model would be one of the observation defined in the previous section which belongs to the set $\Omega = \{o_{N,H}, o_{N,M}, o_{N,L}, o_{T,H}, o_{T,M}, o_{T,L}\}$.
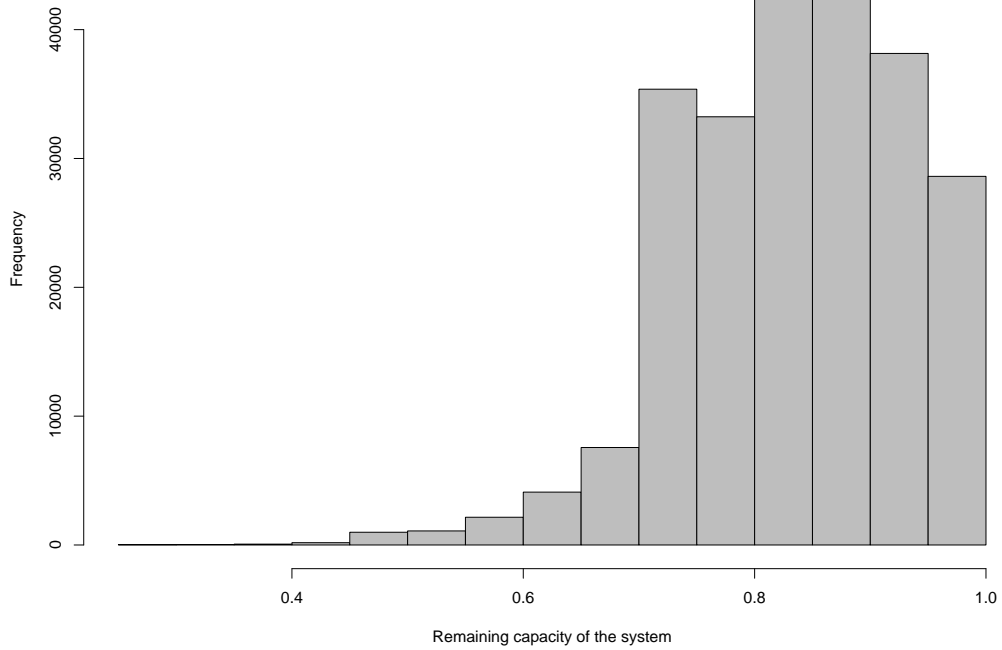
Figure 4.4: Distribution of the remaining capacity of the system for zone 1

The prediction is the state of the system and the independent variables are as follow:

- Current remaining capacity of the system:

- Current road condition

- Day of the week

- Time of the day

- Current month

- Number of incidents happened in the last epoch

Given the above features and target variable we then use Multinomial Logistic Regression to find the relationship between the independent and target variable. The predictive model

Transition Probabilities (a_{ask})		Ending state					
		$s_{N,H}$	$s_{N,M}$	$s_{N,L}$	$s_{T,H}$	$s_{T,M}$	$s_{T,L}$
Starting state	$s_{N,H}$	0.99	0	0	0.01	0	0
	$s_{N,M}$	0.99	0	0	0.01	0	0
	$s_{N,L}$	0.99	0	0	0.01	0	0
	$s_{T,H}$	0.03	0	0	0.97	0	0
	$s_{T,M}$	0.02	0	0	0.98	0	0
	$s_{T,L}$	0.006	0	0	0.994	0	0

Table 4.4: Transition probabilities for action a_{ask}

Transition Probabilities (a_2)		Ending state					
		$s_{N,H}$	$s_{N,M}$	$s_{N,L}$	$s_{T,H}$	$s_{T,M}$	$s_{T,L}$
Starting state	$s_{N,H}$	0.92	0.065	0.004	0.01	0.001	0.00
	$s_{N,M}$	0.093	0.83	0.063	0.008	0.005	0.001
	$s_{N,L}$	0.043	0.089	0.85	0.005	0.006	0.007
	$s_{T,H}$	0.027	0.004	0.00	0.87	0.106	0.007
	$s_{T,M}$	0.02	0.015	0.002	0.088	0.80	0.075
	$s_{T,L}$	0.009	0.011	0.016	0.035	0.099	0.83

Table 4.5: Transition probabilities for action $a_{nothing}$

then predicts the state of the system given the features defined above and the prediction is used as an observation for the POMDP.

For the observation probabilities, according to what we defined before, we will use the prediction performance in the form of a confusion matrix. This is shown in Table 4.6.

4.3.4. Costs

Since the most important performance metric for the DFRD is the response times, we focus on defining our POMDP cost structure based on the same metric. We are trying to reduce the response times as much as possible using the proposed policy.

The basic costs element would be the average response time of the emergency vehicles. Before that we need to mention that there seems to be a relationship between the response times and the remaining capacity of the system. This relation ship is best shown in Figure 4.5.

Observation Probabilities (Confusion Matrix)		Observation (Predicted)					
		$o_{N,H}$	$o_{N,M}$	$o_{N,L}$	$o_{T,H}$	$o_{T,M}$	$o_{T,L}$
	$s_{N,H}$	0.9	0.065	0.004	0.019	0.0015	0.0005
	$s_{N,M}$	0.1	0.82	0.063	0.008	0.006	0.003
State (Actual)	$s_{N,L}$	0.05	0.089	0.84	0.004	0.01	0.007
	$s_{T,H}$	0.03	0.004	0.00	0.86	0.096	0.009
	$s_{T,M}$	0.02	0.016	0.002	0.087	0.80	0.075
	$s_{T,L}$	0.01	0.02	0.016	0.035	0.099	0.82

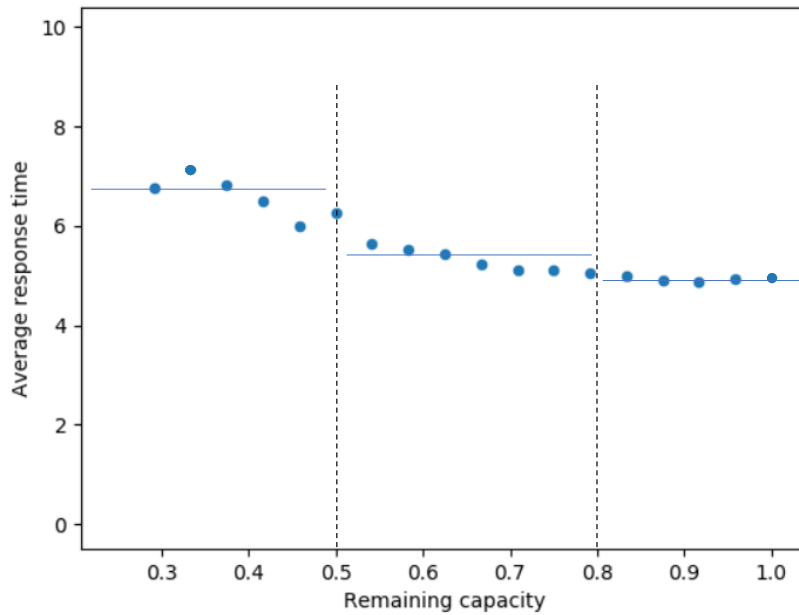Table 4.6: Observation probabilities (confusion matrix) from the predictive model

Figure 4.5: Average response time for each remaining capacity level of the system

Figure 4.5 shows that average response time increases as the remaining capacity of the system decreases. We descretize the remaining capacity of the system into 3 levels. Figure 4.5 shows that there are also 3 levels of response time. Table 4.7 shows the costs given the state of the system and the action taken.

Costs		State					
		$s_{N,H}$	$s_{N,M}$	$s_{N,L}$	$s_{T,H}$	$s_{T,M}$	$s_{T,L}$
Action	a_{ask}	4.2	6.8	8.7	9.7	10.2	12.4
	$a_{nothing}$	4.2	5.8	6.7	6.9	8.6	10.1

Table 4.7: Action and state dependent costs of the POMDP model (in minutes)

Since action $a_{nothing}$ indicates asking for help immediately when an incident happens and there are no resources available, the three elements of the matrix including $c(a_{nothing}, s_{N,H}), c(a_{nothing}, s_{N,M}), c(a_{nothing}, s_{N,L})$ are average response under different remaining capacity levels obtained from data. The other three elements of the matrix relating to action 2 including $c(a_{nothing}, s_{T,H}), c(a_{nothing}, s_{T,M}), c(a_{nothing}, s_{T,L})$ are obtained the same but they are for when the system is experiencing traffic.

For $c(a_{ask}, s_{N,H})$, there are no extra costs compared to the $c(a_{nothing}, s_{N,H})$ case, since asking for help not only does not affect the current remaining capacity of the zone 1, but also it does not take too much capacity from the zone asked from (literally just a few engines are asked and this does not change the capacity level of the other zone and also does not really change the current capacity of the zone 1 and thus really does not affect the response time). For $c(a_{ask}, s_{N,M})$, we have the cost of $c(a_{nothing}, s_{N,M})$ as well as the cost of the other zone changing from $s_{N,H}$ to $s_{N,M}$ (which is the difference between $c(a_{nothing}, s_{N,M})$ and $c(a_{nothing}, s_{N,H})$). For $c(a_{ask}, s_{N,L})$ we have the cost of $c(a_{nothing}, s_{N,L})$ as well as the cost of the other zone changing from $s_{N,H}$ to $s_{N,L}$ (which is the difference between $c(a_{nothing}, s_{N,L})$ and $c(a_{nothing}, s_{N,H})$). For the other three costs, i.e. $c(a_{ask}, s_{T,H})$, $c(a_{ask}, s_{T,M})$, and $c(a_{ask}, s_{T,L})$, the logic is the same but with one difference, and the difference is that, this time the other zone will have some costs, even for the $c(a_{ask}, s_{T,H})$ case and the reason is that we now have traffic and it is not a good idea to lose capacity. Note that the main logic behind these costs
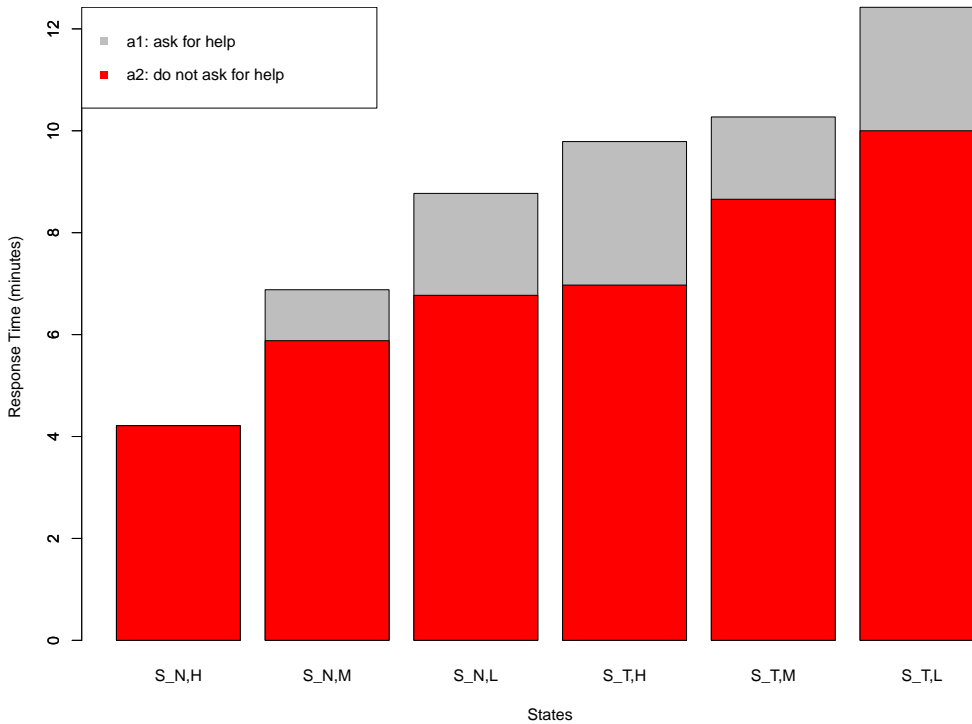
Figure 4.6: costs of POMDP model in minutes for each action and state

is that: by asking for help nothing happens immediately to zone 1 (since it takes time for vehicles to get there) but we have some costs inferred from the other zone immediately (since they dispatch the vehicles as soon as we ask and thus they immediately lose capacity).

Figure 4.6 also shows the costs in a single plot.

4.4. Optimal Reallocation Policy

Using the parameters estimated in previous section and the model developed before, we can now solve the formulated POMDP using the finite grid method over a horizon of 90 days. The optimal policy that POMDP returns is shown in Table 4.8 as follows.

4.5. Policy Evaluation using Simulation

In this section we try to evaluate the proposed policy using simulation. We simulate one

Policy		Action	Observations					
			$o_{N,H}$	$o_{N,M}$	$o_{N,L}$	$o_{T,H}$	$o_{T,M}$	$o_{T,L}$
	1	a_{ask}	10	4	1	1	1	1
	2	$a_{nothing}$	10	9	3	1	1	1
	3	$a_{nothing}$	10	8	9	1	1	1
Value function segments	4	$a_{nothing}$	10	9	7	1	1	2
	5	$a_{nothing}$	10	8	9	1	1	2
(Policy Nodes)	6	$a_{nothing}$	10	8	7	1	2	2
	7	$a_{nothing}$	10	8	9	1	3	5
	8	$a_{nothing}$	10	8	9	5	3	9
	9	$a_{nothing}$	10	8	9	3	3	3
	10	$a_{nothing}$	10	8	9	2	3	5

Table 4.8: Optimal reallocation policy produced by POMDP

zone of city during 90 days, with limited number of resources and compare the proposed policy to when the policy is to ask immediately for help if there are no resources available and an incident has happened.

To better simulate the system we estimate the parameters of the simulation including the distributions of response time and incidents' inter-arrival times.

4.5.1. Simulation parameter estimation

In this section we estimate the following set of parameters using the available data.

- Incidents inter-arrival time (distribution)

- Type of incident

- Number of resources to dispatch

- Response time (distribution)

- Clearance time (distribution)

Incidents inter-arrival time distribution estimation: For the incidents inter-arrival times, we fit multiple distributions including the exponential distribution as this is expected to be the one that fits the most using the maximum likelihood method. Figure 4.7 shows

fitted distributions alongside the Q-Q plot for the theoretical quantiles. As it is inferred from the plot, exponential distribution fits the data the best and also is not rejected by the tests. The results for the tests are provided in Table 4.9.
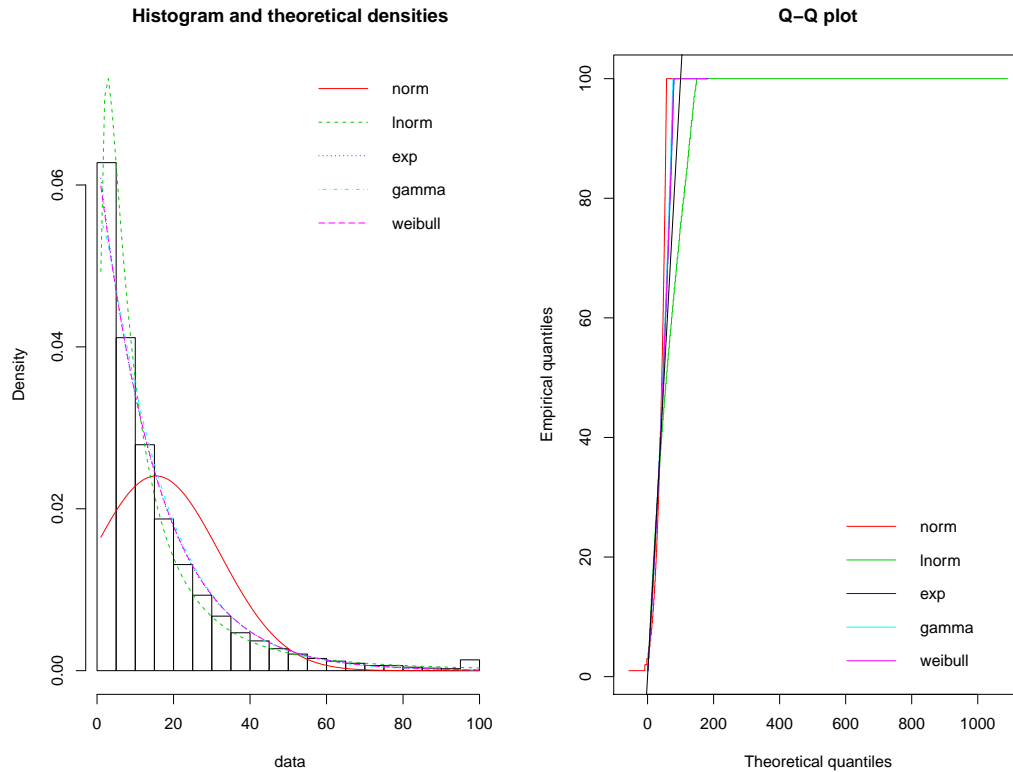
Figure 4.7: The fitted distributions of incidents inter-arrival times

Table 4.9: Distribution fitting analysis for incidents inter-arrival times

Inter-arrival times		Distributions				
		Log Normal	Normal	Exponential	Gamma	Weibull
Parameters		Meanlog= 2.2 (0.0041)	Mean= 15.41 (0.062)	Rate= 0.064 (0.0002)	Shape= 1.073 (0.005)	Shape= 1.007 (0.0028)
		Sdlog= 1.1 (0.0029)	Sd= 16.58 (0.043)		Rate= 0.069 (0.0004)	Scale= 15.47 (0.06)
Goodness of fit statistics	Kolmogorov-Smirnov statistic	0.057	0.192	0.062	0.053	0.061
	Cramer-von Mises statistic	44.56	793.3	32.27	48.46	34.97
	Anderson-Darling statistic	369.43	4521.8	295.8	347.79	303.0
Goodness-of-fit criteria	Akaike's Information Criterion	531352.1	603937.1	533648.4	533424.6	533642.4
	Bayesian Information Criterion	531370.4	603955.5	533657.6	533442.9	533660.8

Type if incidents: For the types of the incidents that happen, we simply use the number of resources required to handle the incident. We use the frequency of each type of incident as the probability of that type of incident. These frequencies and their respective probabilities are given in Table 4.11.

64

Table 4.10: Distribution fitting analysis for response times

Response times		Distributions				
		Log Normal	Normal	Exponential	Gamma	Weibull
Parameters		meanlog= 1.35 (0.0018) sdlog= 0.501 (0.0013)	Mean= 4.38 (0.009) SD= 2.42 (0.006)	Rate=0.228 (0.00024)	Shape= 4.21 (0.021) Rate= 0.96 (0.005)	Shape= 1.92 (0.0047) Scale= 4.95 (0.01)
Goodness of fit statistics	Kolmogorov-Smirnov statistic	0.144	0.172	3.354615e-01	0.125	0.157
	Cramer-von Mises statistic	213.535	430.68	1.987939e+03	195.86	317.37
	Anderson-Darling statistic	1246.5	Infinite	1.013887e+04	1111.83	1931.8
Goodness-of-fit criteria	Akaike's Information Criterion	296725.9	328277.4	353039.7	298307.1	309576.9
	Bayesian Information Criterion	296744.2	328295.8	353048.9	298325.4	309595.2

The number of resources required to handle each incident type depends the incident type as this is how we defined the incident type.

Response time distribution estimation: To find the distribution that best fits the distribution of response times, we fit multiple distributions including the normal, log normal, gamma, and Weibull distributions using the maximum likelihood method. Figure 4.7 shows fitted distributions alongside the Q-Q plot for the theoretical quantiles. As it is inferred from the plot, the log normal distribution fits the data the best and also is not rejected by the tests. The results for the tests are provided in Table 4.10.

Clearance time: The clearance time which is the time it takes for the resources to handle the situation, clear it and get back to the stations and eventually get ready for the next incident depends on the incident type. This is clearly shown in Figure 4.9 where clearance time is depicted with relation to the incident type. Thus we estimate this from the data for each incident type and provide the averages for each in Table 4.11.

Table 4.11: Incident types and their respective probabilities of happening

| Incident type	Frequency	Probability	Average clearance time (minutes)
1	45497	0.61	33.56
2	18473	0.24	18.58
3	7643	0.10	13.09
4	1468	0.019	12.36
5	358	0.004	12.12
6	153	0.002	16.9
7	173	0.002	14.14
8	78	0.001	15.16
9	49	0.0006	10.9
10	43	0.00057	34.09
11	36	0.00048	46.7
12	18	0.00024	27.57
13	14	0.00018	34.09
14	42	0.00056	18.12
15	39	0.00052	20.01
16	24	0.00032	19.7
17	11	0.00014	14.1721
18	5	0.00006	17.52
19	6	0.00008	17.93
20	2	0.000026	33.77
21	2	0.000026	61.9
22	1	0.000013	59.5
24	1	0.000013	65.3
26	1	0.000013	104.42
28	1	0.000013	68.07
29	1	0.000013	74.86
30	1	0.000013	76.56
31	1	0.000013	97.548
39	3	0.00004	108.48
40	1	0.000013	112.32
41	1	0.000013	94.9
44	1	0.000013	101.17
53	1	0.000013	163.45
56	1	0.000013	157.23
65	1	0.000013	232.32
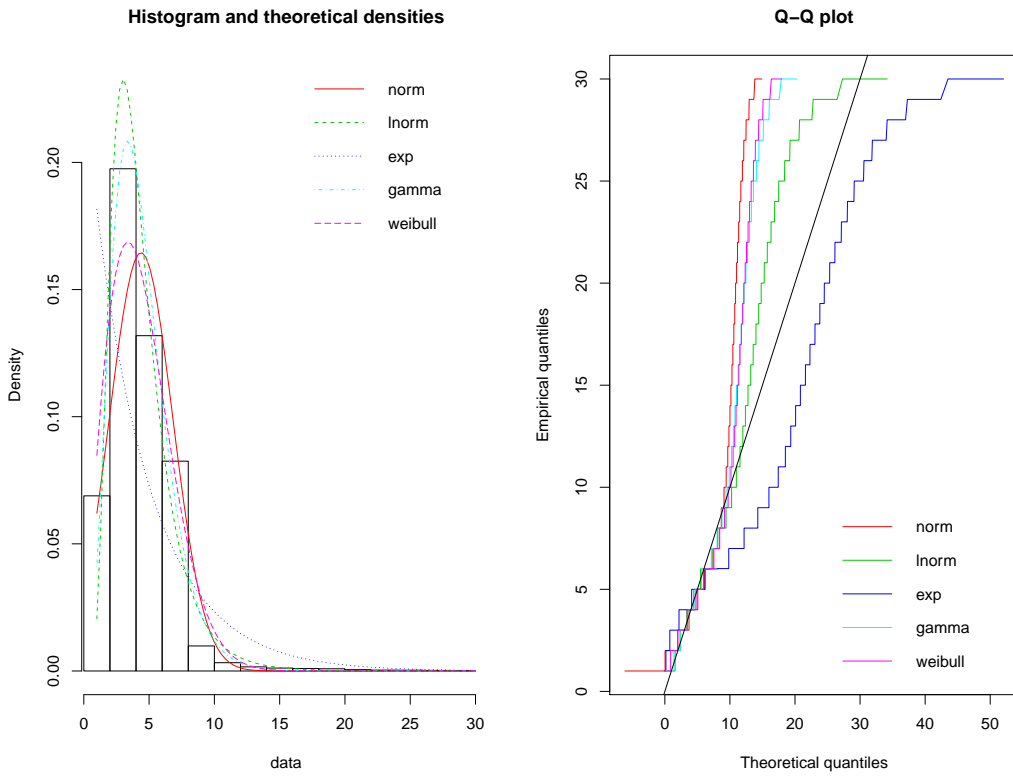
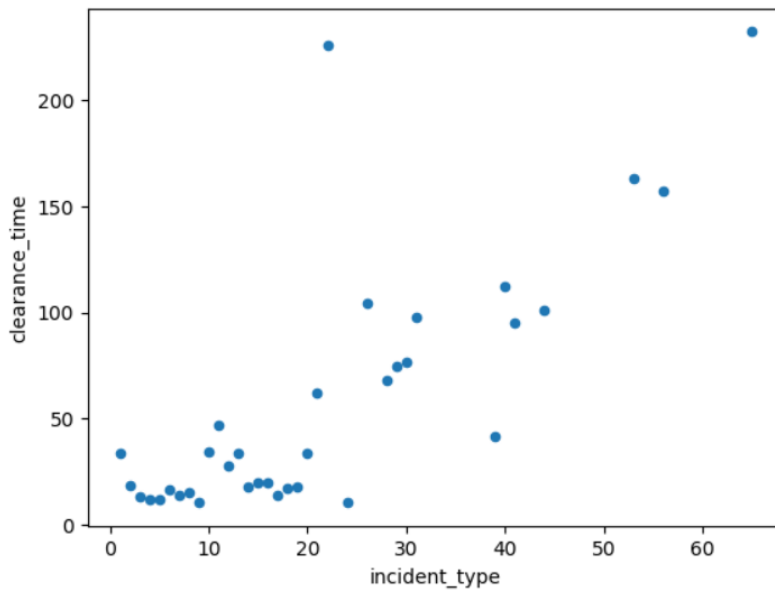Figure 4.8: The fitted distributions of response times

Figure 4.9: The average clearance time for each type of incident

Based on Figure 4.9 the average clearance time increases as the type of incident increases (more severe incidents).

4.5.2. Evaluation using simulation

Using the parameters estimated and calculated in the previous section, we simulate one zone of the city with a limited number of resources (similar to zone 1 of the city of Dallas, we only have 24 emergency vehicles available). We compare two different scenarios:

- Scenario 1: The DFRD asks for extra resources from other zones only when an incident happens and not enough resources are available to respond

- Scenario 2: The DFRD follows the proposed POMDP policy demonstrated in Table 4.8 and asks for extra resources whenever the policy implies.

We then run the simulation for 90 days and 30 replications and compare the two scenarios in various metric presented in Table 4.12.

According to Table 4.12 by implementing the proposed POMDP policy we can improve the average response time by almost 48 seconds (by 13%) which is a significant improvement in emergency situations. We can also decrease the number of times not enough resources are available to respond to an incident by almost 80 percent. This means that our system will have less chances of having not enough resources available to respond to an incoming incident. We also decrease the number of resources requested when not enough resources are available by around 66 percent. This means even if we are asking for extra resources from other zones, this request comes in very small amounts. POMDP policy tends to keep the system at full capacity while without such a policy the average remaining capacity of the system is around 67 percent. Implementing the proposed policy will result in spending more time in state 1 and 4 compared to the time where no such policy is implemented in which case the system is mostly in states 3 or 6. This is what we are trying to avoid.

Table 4.12: Comparison of two scenarios in various metrics

		Average Response Time	Number of times extra resources requested	Number of times not enough resources were available	Number of extra resources requested	Average remaining capacity	Average time in state 1	Average time in state 2	Average time in state 3	Average time in state 4	Average time in state 5	Average time in state 6
Scenario 1	Mean	6.25	46.9	46.9	115.81	0.67	0.06	0.13	0.50	0.026	0.059	0.21
	STD	0.043	13.94	13.94	45.034	0.0033	0.0046	0.0047	0.0075	0.0024	0.0034	0.0051
Scenario 2	Mean	5.40	38907.75	9.7	38.98	1.032	0.56	0.06	0.071	0.24	0.026	0.032
	STD	0.029	113.72	7.68	35.72	0.006	0.005	0.0026	0.005	0.0038	0.0017	0.0029

Chapter 5

Conclusion

The framework proposed in this work focuses on using historical data to develop a predictive model that provides fewer but more accurate observations from a system modeled as a POMDP. The focus of the framework is first, to reduce the dimensionality of the observation space of the POMDP by selecting only a collection of signals that provides a single observations, and second, to provide more accurate observations by aggregating the selected signals using a predictive model that predicts the state of the system. Using the framework, we reduce the size of the problem by a large (exponential) factor while sacrificing only little in terms of the quality of the resulting POMDP policy. We also avoid the effort required to estimate all the observation probabilities which can be a difficult task when not enough data is available. This task is easily handled by the predictive model and from the data available and these probabilities are then provided more accurately.

The effectiveness of the proposed framework depends on the data available, but this is generally true for using POMDPs since parameters need to be estimated. The framework provides all the advantages of predictive modeling including methods for feature selection, dealing with missing data and data quality issues. When using this framework, the predictive model is trained supervised, i.e., the training data needs to be annotated with class information. This means that at some point in time, information on the actual states of the system must have been recorded along with the signal values. Otherwise, no predictive model can be developed based on the available data.

The framework can also be used by aggregating groups of signals into several types of observations. So far, we have only considered a single signal aggregated from a selected set of signals. This selection is only based on the accuracy of signals from the system. But in the future, the framework can expand in a way that provides various signals each a combination

of a different set of signals from the system. The sets can include signals that are more related to each other or are from the same part of the system or their nature is closer to each other. The relationship between signals can also be learned from data using unsupervised learning technique after the signal selection step to group signals into sets.

The proposed framework is then utilized in two different applications.

5.1. Optimal Individualized Diabetes Screening

Diabetes, a prevalent chronic disease affecting over 30 million American adults, is associated with multiple comorbidities and is the seventh-leading cause of death in the United States. The disease, associated with hundreds of billions of dollars in direct and indirect costs (ADA 2018), progresses with a lengthy asymptomatic period of 9 to 12 years, on average (Lu et al. 2010). Thus, it is critical to screen patients who have undiagnosed diabetes or those who are at an elevated risk of developing diabetes as this can result in substantial savings, since appropriate interventions can be put in place to prevent progression to diabetes and development of diabetes complications. Existing guidelines such as those from ADA (ADA 2019) are generic and cost-prohibitive if implemented on the entire population, since (i) only 9.4 percent of the population is at risk of developing diabetes (CDC 2017), and (ii) the gold standard test (using A1c) for screening is very expensive (Chatterjee et al. 2013). There does not exist, to the best of our knowledge, a personalized screening strategy for detecting patients with diabetes or prediabetes. This is exactly what this paper attempts to do.

In this study, we propose a targeted screening policy (equivalently, screening strategy) that uses all available information on individual patients to identify whom to screen (that is, which patients should receive the gold-standard test) and when to screen them; the policy is also age-specific. Our proposed policy relies on multiple methods and is based on actual patient data (available in the form of EHRs), making it practically implementable. In particular, POMDP is used for determining optimal decisions at each time period (a year) while HMM and PRM are used to generate the transition and observation probabilities, respectively, for POMDP. Thus, a key contribution of this work is the holistic integration of the three

71

methods to answer a practical healthcare decision-making problem of this magnitude.

We develop and validate our model on a detailed and proprietary dataset – of over 12,000 patients over an 18-month period – from a large safety-net hospital and demonstrate, using a detailed simulation analysis, that our proposed screening policy can improve patient outcomes by 106 percent - at only 65 percent of the cost – compared with existing guidelines. Our detailed sensitivity analyses show no significant or unjustifiable change in the results of the simulation due to changes in the model parameters.

The integration of data analytics with optimization methods has become increasingly critical to solve important problems in healthcare and beyond. In this study, we have demonstrated how existing methods can be combined with POMDP to produce an optimal screening policy that incorporates cohort-specific characteristics as well as individualized medical information.

Our study has several limitations. From a methodological standpoint, first, the use of POMDP approach relies on a potentially strong assumption that the Markov property holds, at least approximately. However, our assumption is consistent with previous work, where experts have argued that Markov models are useful approximations for disease progression models. Second, our state space is primarily based on a single measure (A1c), resulting in a simplistic single dimensional state space. However, our review of the literature and conversations with our clinical co-author reveal that A1c is the most commonly used test in clinical practice to screen and diagnose patients for diabetes. While inclusion of additional covariates may enrich the model, it also increases the model complexity and the computational effort required to solve such a model. Third, we only consider two actions – screen or do not screen (the latter being essentially an absence of action). However, expanding the action space will require the estimation of all associated transition parameters. Therefore, smaller state and action spaces are preferable from a practical standpoint. Fourth, estimation of the disease progression rates relies on the available screening results from the EHR data and is, thus, applicable only to screened patients with visits in the health system studied. We estimate the transition rates of unscreened patients by using a factor

representing treatment effectiveness. This is a simplistic approach that requires more research. Finally, our analysis is limited to patients at a single hospital who may not share the same characteristics with patients in other hospitals. However, we believe our methodological approach is generalizable and can be applied to patients in other settings, with modifications.

5.2. Resource Allocation Under Uncertainty for Emergency Vehicles

An important question for the DFRD is whether resources should be moved around in the city to cover areas where the resources are currently responding to an incidence. Every time an incident happens resources in a particular zone of the city will be dispatched and become unavailable for several hours. If another incident happens in that zone during that time, resources from other areas of the city will need to respond which will increase response time. To mitigate such situations, we can temporarily reallocate resources.

We apply framework proposed in chapter 2, and formulate the problem as a POMDP problem. We focus on one city zone in order to define our state space. We try to capture the availability of the resources in that zone in the near future; i.e. the states of our model are trying to describe the status of a particular zone in the near future. This near future depends on how long in advance we want to be aware of the future and also the risk associated with the problem. By defining what would be the status of a zone in the near future as the state of the model, we add unobservability to the model and since only in the future we will for sure know the answer to this question, the state space is always unobservable to us.

By implementing the proposed POMDP policy, and through simulation, we demonstrate that we can improve the average response time by almost 48 seconds which is a significant improvement in emergency situations. We can also decrease the number of times not enough resources are available to respond to an incident by almost 80 percent. This means that our system will have less chances of having not enough resources available to respond to an incoming incident. We also decrease the number of resources requested when not enough resources are available by around 66 percent. This means even if we are asking for extra

resources from other zones, this request comes in very small amounts. POMDP policy tends to keep the system at full capacity while without such a policy the average remaining capacity of the system is around 67 percent. Implementing the proposed policy will result in spending more time in state 1 and 4 compared to the time where no such policy is implemented in which case the system is mostly in states 3 or 6. This is what we are trying to avoid.

Appendix A

Diabetes Simulation Details

In this section we provide details on the simulation conducted for the fist application in this work. The simulation consists of a hypothetical cohort of 50,000 patients starting from age 30 to age 79, with characteristics described in Table 3.1. We simulate seven different scenarios representing different screening policies, including our proposed one. The only difference between these scenarios is the screening policy implemented. The six scenarios that use the existing or hypothetical guidelines are similar to each other thus only one of them will be explained here along with the proposed policy. All scenarios are compared to the base scenario. The base scenario is when there is not a specific screening policy (commonly called opportunistic screening). This means that we screen patients if they show up and have symptoms or if they request so, but we never prescribe screening for them or ask them to show up later for a screening test.

The simulation is an aging loop where patients enter with the age of 30 and leave either when they die or reach the age of 79 as demonstrated in Figure A.1. We explain here each part of the simulation loop in details.

A.1. Patient Instantiation

We instantiate each patient using the prevalence rate of each of the stages of the disease in the cohort described in Table 3.1. These prevalence rate are taken from Table 3.1. We also assign diabetes complications to the patients that already have Diabetes using the prevalence rates from Chen et al. (2001), Howard et al. (2010), and Kahn et al. (2010). These complications along with the diabetes progression Markov models are demonstrated in Figure 3.6. The prevalence rates for each of the states of the models in Figure 3.6 are provided in Table A.1.
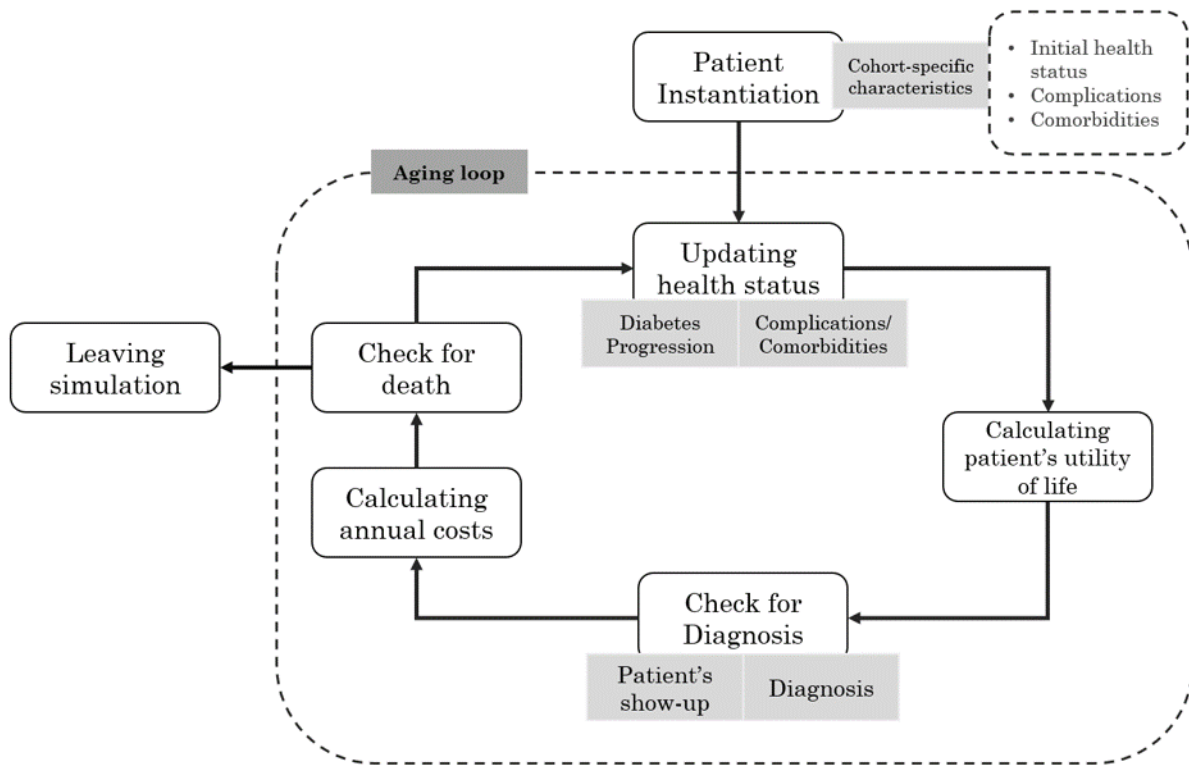
Figure A.1: Simulation loop for the base scenario

A.2. Updating Health Status

At the beginning of each iteration, patients' health status gets updated using the progression rates obtained via HMM and from Chen et al. (2001), Howard et al. (2010), and Kahn et al. (2010) according to the Markov models demonstrated in Figure 3.6. The progression rates for each transition is provided in Table A.2.

A.3. Calculating Patient's Utility

At each iteration, patient's utility of life is calculated using the EQ-5D index from Ackermann et al. (2009), Bahia et al. (2017), and P. Zhang et al. (2012). These utilities are provided in Table A.3. The worst utility is assumed to be the utility of the patient with multiple chronic conditions and complications.

Table A.1: Prevalence rates of Diabetes and its complications stages

Disease	Stages				
Diabetes	Healthy	Pre-Diabetes	Diabetes		
	0.508	0.358	0.133		
Retinopathy	NDR	NPDR	PDR	ME	B
	0.5	0.2	0.05	0.25	0
Nephropathy	NNP	MA	PR	ESRD	CVD
	0.579	0.2	0.05	0.025	0.146
Neuropathy	NNR	SNR	LEA		
	0.7	0.3	0		

Table A.2: Progression rates for transitions in Diabetes and its complications

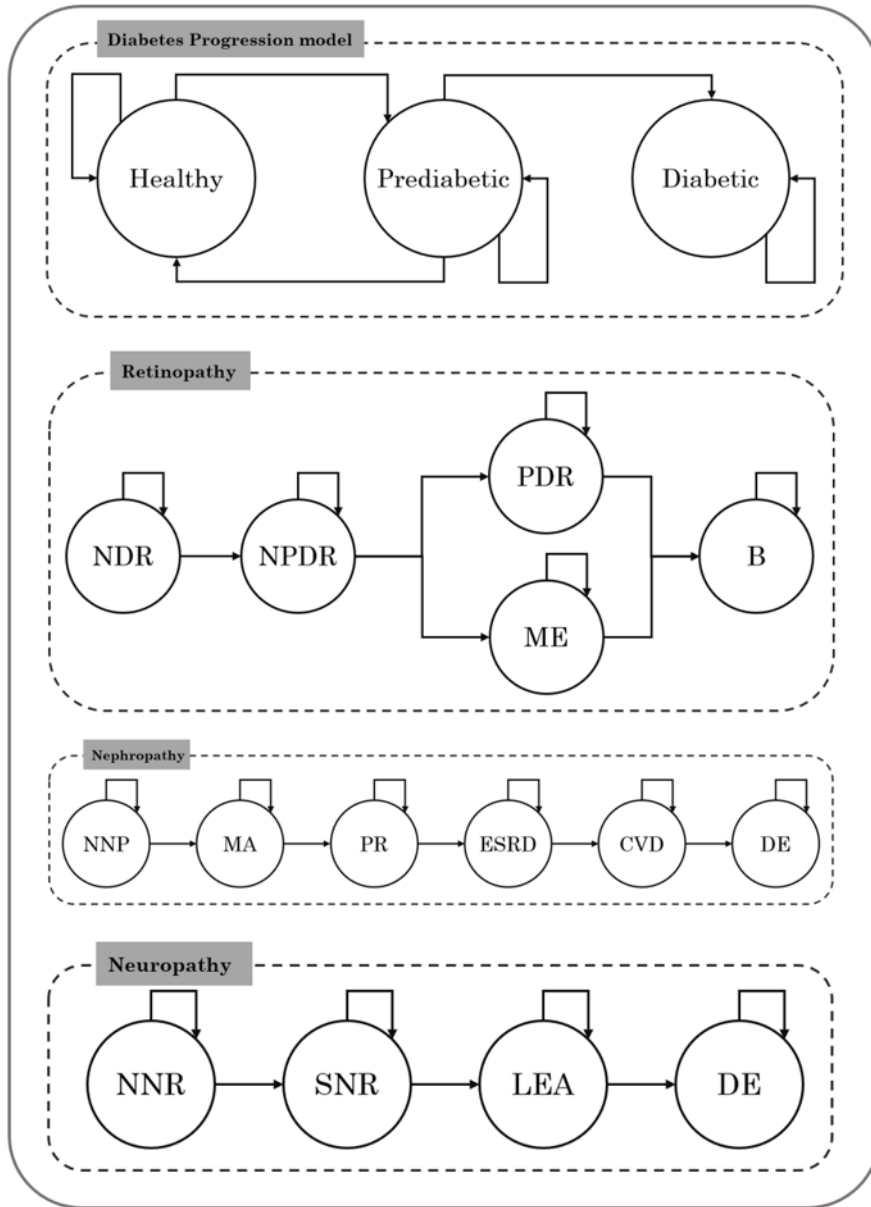| Disease	Transitions				
Diabetes	H to P	P to H	P to D		
	0.048	0.0328	0.0348		
Retinopathy	NDR to NPDR	NPDR to PDR	NPDR to ME	PDR to B	ME to B
	0.073	0.0103	0.1928	0.0148	0.033
Nephropathy	NNP to MA	MA to PR	PR to ESRD	ESRD to CVD	CVD to DE
	0.0267	0.1572	0.0042	0.5	0.2
Neuropathy	NNR to SNR	SNR to LEA	LEA to DE		
	0.0144	0.028	0.02		

Figure A.2: Diabetes and its complications Markov progression models

Table A.3: Life utilities for different health conditions

Condition	Utility of Life for living with the condition
Healthy	1
Pre-Diabetes	0.84
Diagnosed Diabetes	0.82
Undiagnosed Diabetes	0.8
Blindness	0.69
ESRD	0.61
CVD	0.63
LEA	0.59

A.4. Check for Diagnosis

For Diabetes and its complications there is usually a time to diagnosis from the onset of the disease. For Diabetes it is on average 10 years and for its complications it is on average 3 years (Lu et al. 2010). This means that given there is no screening policy, the patients will be diagnosed on average after the given amount of time due to major health problems they face. Thus at each iterations there is a certain chance based on the patient's conditions that the patient gets diagnosed with her condition. This is contingent upon the patient's showing up and visiting doctor. There is a certain chance associated with that called patient's tendency to visit doctor given their health conditions. These probabilities include on average 10%, 25% and 55% for when they have almost zero symptoms, medium risk symptoms and high-risk symptoms respectively. The high-risk symptoms are associated with later stages of the diseases presented in Figure A.2.

A.5. Patient's Annual Costs

At each year the total annual costs of the patient including Diabetes costs and Complications costs are calculated using the figures provided in Chen et al. (2001), Howard et al. (2010), and Kahn et al. (2010). All these costs are provided in US dollars in Table A.4. All costs are then discounted by 0.03 each year.

A.6. Leaving the Simulation

Table A.4: All costs associated with Diabetes and its complications

Type of Cost	Detail	Costs
Visit and Screening	Visit	$134
	Screening	$192
Intensive Glycemic control	Drugs	$862
	Outpatient	$910
Conventional Glycemic control	Total	$765
Complications Costs	Blindness	$1,997
	Photocoagulation	$2,682
	ESRD	$68,131
	LEA (per operation)	$31,139
	CVD	$2,757
Intensive Hypertension Control	Drugs	$686
	Outpatient	$217
Conventional Hypertension control	Drugs	$394
	Outpatient	$149

At the end of each year, the patient will leave the simulation if she dies by any of the complications or through the natural progression of Diabetes or by reaching the age of 79. This means that the patients will stay in simulation for at most 50 years and then leave it.

A.7. Implementing POMDP Policy/Other Guidelines

Almost everything is the same as the base scenario except for the screening policy here. In the POMDP scenario, we have the same chances for patients show-up at the doctors as the base scenario. Each time the patient shows up, the policy node at which the patient is, gets updated according to the observation made by the predictive model. At the show-up where the policy indicates screening, the screening would take place. For other guidelines it is the same except that guidelines have a fixed frequency of screening and if that matches with the time that the patient shows up, the screening will take place. Note that in all scenarios, no screening would happen if the patient does not show up. The simulation loop for POMDP scenario is depicted in Figure A.3.
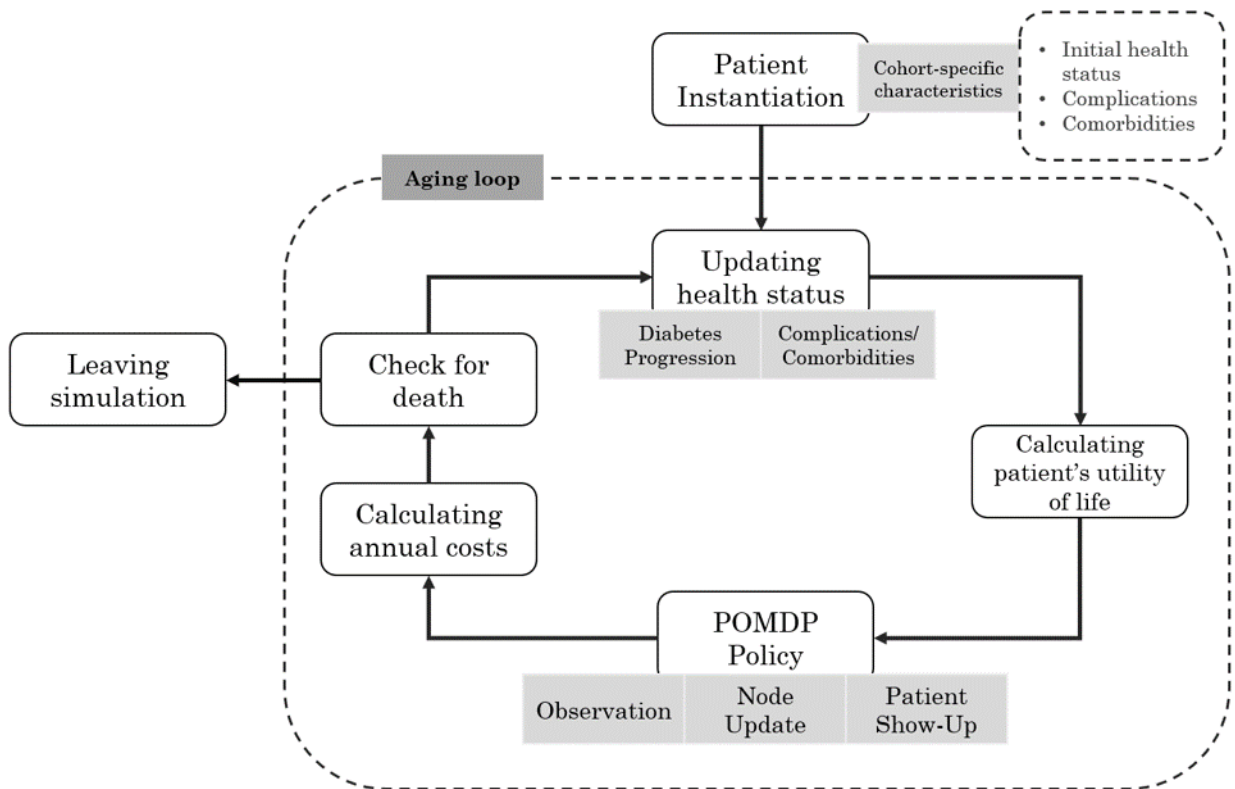
Figure A.3: Simulation loop for the POMDP policy

Appendix B

R Package 'pomdp'

Following is the manual to the R package 'pomdp' developed during this research. This package is a solver for Partially Observable Markov Decision Processes. The package enables the user to simply define all components of a POMDP model and solve the problem using several methods. The package also contains functions to analyze and visualize the POMDP solutions (e.g., the optimal policy)

The 'pomdp' package is available at CRAN at https://cran.r-project.org/package=pomdp

Package 'pomdp'

December 16, 2019

Title Solver for Partially Observable Markov Decision Processes
(POMDP)

Version 0.9.2

Date 2019-12-06

Description Provides an interface to pomdp-solve, a solver for Partially Observable Markov Deci-
sion Processes (POMDP). The package enables the user to simply define all compo-
nents of a POMDP model and solve the problem using several methods. The package also con-
tains functions to analyze and visualize the POMDP solutions (e.g., the optimal policy).

Depends R (>= 3.5.0)

License GPL (>= 3)

Suggests knitr, rmarkdown

VignetteBuilder knitr

LazyData true

Imports igraph

Copyright pomdp-solve is Copyright (C) Anthony R. Cassandra; LASPack
is Copyright (C) Tomas Skalicky; lp-solve is Copyright (C)
Michel Berkelaar, Kjell Eikland, Peter Notebaert; all other
code is Copyright (C) Hossein Kamalzadeh and Michael Hahsler.

NeedsCompilation yes

Author Hossein Kamalzadeh [aut, cph, cre],
Michael Hahsler [aut, cph],
Anthony R. Cassandra [ctb, cph]

Maintainer Hossein Kamalzadeh <hkamalzadeh@smu.edu>

Repository CRAN

Date/Publication 2019-12-16 09:30:08 UTC

R topics documented:

---

model   *Extract the User-defined Model Components from a Solved POMDP*

Description

The function returns the POMDP model components of a solved POMDP.

Usage

```
model(x)
```

Arguments

x   object of class POMDP returned by solve_POMDP.

Value

An object of class "POMDP_model", i.e., a list of all model components.

See Also

solve_POMDP

Examples

```
data("TigerProblem")
tiger_solved <- solve_POMDP(model = TigerProblem)
tiger_solved

model(tiger_solved)
```

---

plot   *Visualize a POMDP Policy Graph*

Description

The function plots the POMDP policy graph in an object of class POMDP. It uses plot in **igraph** with appropriate plotting options.

Usage

```
## S3 method for class 'POMDP'
plot(x, y = NULL, belief = TRUE, legend = TRUE, cols = NULL,...)
```

Arguments

x	object of class POMDP.
y	ignored.
belief	logical; display belief proportions as a pie chart in each node.
legend	logical; display a legend for colors used belief proportions?
cols	colors used for the states.
...	plotting options passed on to plot.igraph in **igraph** (see plot.common for available options).

Details

The policy graph nodes represent segments in the value function. Each segment represents one or more believe states. The pie chart in each node (if available) represent the average belief proportions of the belief states belonging to the node/segment.

See Also

solve_POMDP, plot.igraph, igraph_options, plot.common

Examples

```
data("TigerProblem")
tiger_solved <- solve_POMDP(model = TigerProblem)
tiger_solved

## policy graph
policy_graph(tiger_solved)

## visualization
plot(tiger_solved)

library(igraph)
```

```
## use a different graph layout (circle and manual)
plot(tiger_solved, layout = layout.circle)
plot(tiger_solved, layout = rbind(c(1,1), c(1,-1), c(0,0), c(-1,-1), c(-1,1)))

## hide edge labels
plot(tiger_solved, edge.label = NA)

## custom larger vertex labels (A, B, ...)
plot(tiger_solved,
  vertex.label = LETTERS[1:nrow(solution(tiger_solved)$pg)],
  vertex.label.cex = 2,
  vertex.label.color = "white")

## add a plot title
plot(tiger_solved, main = model(tiger_solved)$name)

## plotting using the graph object
## (e.g., using the graph in the layout and to change the edge curvature)
pg <- policy_graph(tiger_solved)
plot(pg,
  layout = layout_as_tree(pg, root = 3, mode = "out"),
  edge.curved = curve_multiple(pg, .2))
```

---

policy_graph	*Extract the Policy Graph (as an igraph Object)*

Description

Convert the policy graph in a POMDP solution object into an igraph object.

Usage

```
policy_graph(x, belief = TRUE, cols = NULL)
```

Arguments

x	A POMDP object.
belief	logical; add belief proportions as a pie chart in each node of the graph.
cols	colors used for the states in the belief proportions.

Value

An object of class igraph containing a directed graph.

Author(s)

Hossein Kamalzadeh, Michael Hahsler

86

See Also

solve_POMDP

Examples

```
data("TigerProblem")
tiger_solved <- solve_POMDP(model = TigerProblem)
tiger_solved

pg <- policy_graph(tiger_solved)

plot(pg)
```

POMDP	*Define a POMDP Problem*

Description

Defines all the elements of a POMDP problem including the discount rate, the set of states, the set of actions, the set of observations, the transition probabilities, the observation probabilities, and rewards.

Usage

```
POMDP(discount, states, actions, observations, transition_prob,
      observation_prob, reward, start = "uniform", max = TRUE, name = NA)

R_(action, start.state, end.state, observation, value)
O_(action, end.state, observation, probability)
T_(action, start.state, end.state, probability)
```

Arguments

discount   numeric; discount rate between 0 and 1.

states   a character vector specifying the names of the states.

actions   a character vector specifying the names of the available actions.

observations   a character vector specifying the names of the observations.

transition_prob

               Specifies the transition probabilities between states. Options are:

- a data frame with 4 columns, where the columns specify *action*, *start-state*, *end-state* and the *probability* respectively. The first 3 columns could be either character (the name of the action or state) or integer indices.
- a named list of m (number of actions) matrices. Each matrix is square of size $n \times n$, where n is the number of states. The name of each matrix the action it applies to. Instead of a matrix, also the strings "identity" or "uniform" can be specified.

observation_prob

> Specifies the probability that a state produces an observation. Options are:
>
> - a data frame with 4 columns, where the columns specify *action*, *end-state*, *observation* and the *probability*, respectively. The first 3 columns could be either character (the name of the action, state, or observation), integer indices, or they can be "*" to indicate that the observation probability applies to all actions or states. Use rbind() with helper function O_() to create this data frame.
> - a named list of m matrices, where m is the number of actions. Each matrix is of size $n \times o$, where n is the number of states and o is the number of observations. The name of each matrix is the action it applies to. Instead of a matrix, also the strings "identity" or "uniform" can be specified.

reward

> Specifies the rewards dependent on action, states and observations. Options are:
>
> - a data frame with 5 columns, where the columns specify *action*, *start.state*, *end.state*, *observation* and the *reward*, respectively. The first 4 columns could be either character (names of the action, states, or observation), integer indices, or they can be "*" to indicate that the reward applies to all transitions. Use rbind() with helper function R_() to create this data frame.
> - a named list of m lists, where m is the number of actions (names should be the actions). Each list contains n named matrices where each matrix is of size $n \times o$, in which n is the number of states and o is the number of observations. Names of these matrices should be the name of states.

start

> Specifies the initial probabilities for each state (i.e., the initial belief state) used to find the initial node in the policy graph and to calculate the total expected reward. The default initial belief state is a uniform distribution over all states. No initial belief state can be used by setting start = NULL. Options to specift start are:
>
> - a probability distribution over the n states. That is, a vector of n probabilities, that add up to 1.
> - the string "uniform" for a uniform distribution over all states.
> - an integer in the range 1 to n to specify a single starting state. or
> - a string specifying the name of a single starting state.
> - a vector of strings, specifying a subset of states with a uniform start distribution. If the first element of the vector is "-", then the following subset of states is excluded from the set of start states.

max

> logical; is this a maximization problem (maximize reward) or a minimization (minimize cost specified in reward)?

name

> a string to identify the POMDP problem.

action, start.state, end.state, observation, probability, value

> Values used in the helper functions O_(), R_(), and T_() to create an entry for observation_prob, reward, or transistion_prob above, respectively.

Details

POMDP problems can be solved using solve_POMDP. Details about the available specifications can be found in [1].

Value

The function returns an object of class POMDP which is list with an element called `model` containing a list with the model specification. `solve_POMDP` reads the object and adds a list element called `solution`.

Author(s)

Hossein Kamalzadeh, Michael Hahsler

References

[1] For further details on how the POMDP solver utilized in this R package works check the following website: <http://www.pomdp.org>

See Also

[solve_POMDP](#)

Examples

```
## The Tiger Problem

TigerProblem <- POMDP(
  name = "Tiger Problem",

  discount = 0.75,

  states = c("tiger-left" , "tiger-right"),
  actions = c("listen", "open-left", "open-right"),
  observations = c("tiger-left", "tiger-right"),

  start = "uniform",

  transition_prob = list(
    "listen" =     "identity",
    "open-left" =  "uniform",
    "open-right" = "uniform"),

  observation_prob = list(
    "listen" = rbind(c(0.85, 0.15),
                     c(0.15, 0.85)),
    "open-left" =  "uniform",
    "open-right" = "uniform"),

  # the rew helper expects: action, start.state, end.state, observation, value
  reward = rbind(
    R_("listen",     "*",           "*", "*", -1  ),
    R_("open-left",  "tiger-left",  "*", "*", -100),
    R_("open-left",  "tiger-right", "*", "*", 10  ),
    R_("open-right", "tiger-left",  "*", "*", 10  ),
    R_("open-right", "tiger-right", "*", "*", -100)
```

```
  )
)

TigerProblem

model(TigerProblem)
```

---

reward   *Calculate the Reward for a POMDP Solution*

Description

This function calculates the expexted total reward for a POMDP solution given a starting belief state.

Usage

```
reward(x, start = "uniform")
```

Arguments

x	a POMDP solution (object of class POMDP).
start	specification of the starting belief state (see argument start in POMDP for details).

Details

The value is calculated using the value function stored in the POMDP solution.

Value

A list with the components

total_expected_reward
> the total expected reward starting with the initial policy graph node representing the starting belief state.

initial_pg_node
> the policy graph node that represents the starting belief state.'

start_belief_state
> the starting belief state specified in start.

Author(s)

Michael Hahsler

See Also

POMDP, solve_POMDP

solution

Examples

```
data("TigerProblem")
tiger_solved <- solve_POMDP(model = TigerProblem)

# if no start is specified, a uniform belief is used.
reward(tiger_solved)

# we have additional information that makes us belief that the tiger
# is more likely to the left.
reward(tiger_solved, start = c(0.85, 0.15))

# we start with strong evidence that the tiger is to the left.
reward(tiger_solved, start = "tiger-left")

# Note that in this case, the total discounted expected reward is greater
# than 10 since the tiger problem resets and another game staring with
# a uniform belief is played which produces addional reward.
```

---

solution   *Extract the Solution of a POMDP*

Description

The function extracts the solution of a POMDP as an object of class POMDP_solution which is a list containing, e.g., the policy graph (pg) and the hyper-plane coefficients (alpha).

Usage

```
solution(x)
```

Arguments

x   object of class POMDP returned by solve_POMDP.

Value

returns an object is of class POMDP_solution, i.e., a list of all solution elements.

See Also

solve_POMDP

Examples

```
data("TigerProblem")
tiger_solved <- solve_POMDP(model = TigerProblem)
tiger_solved

solution(tiger_solved)
```

---

solver_output   *Display the Output of the POMDP Solver*

Description

Displays the output generated by the solver 'pomdp-solve'. This includes used parameters, and iterations (i.e., epochs). This produces the same output as running `solve_POMDP` with the argument `verbose = TRUE`.

Usage

```
solver_output(x)
```

Arguments

x   object of class POMDP returned by [solve_POMDP](#).

Value

returns invisibly a character string vector with the output of 'pomdp-solve'.

See Also

[solve_POMDP](#)

Examples

```
data("TigerProblem")
sol <- solve_POMDP(model = TigerProblem)

## solver output
solver_output(sol)
```

---

solve_POMDP   *Solve a POMDP Problem*

Description

This function utilizes the 'pomdp-solve' program (written in C) to use different solution methods [2] to solve problems that are formulated as partially observable Markov decision processes (POMDPs) [1]. The result is a (close to) optimal policy.

Usage

```
solve_POMDP(model, horizon = NULL, method = "grid", parameter= NULL, verbose = FALSE)
solve_POMDP_parameter()
```

solve_POMDP

Arguments

model	a POMDP problem specification created with POMDP. Alternatively, a POMDP file or the URL for a POMDP file can be specified.
method	string; one of the following solution methods: "grid", "enum", "twopass", "witness", or "incprune". Details can be found in [1].
horizon	an integer with the number of iterations for finite horizon problems. If set to NULL, the algorithm continues running iterations till it converges to the infinite horizon solution.
parameter	a list with parameters passed on to the pomdp-solve program.
verbose	logical, if set to TRUE, the function provides the output of the pomdp solver in the R console.

Details

solve_POMDP_parameter() displays available solver parameter options.

Note: The parser for POMDP files is experimental. Please report problems here: https://github.com/farzad/pomdp/issues.

Value

The solver returns an object of class POMDP which is a list with the model specifications (model), the solution (solution), and the solver output (solver_output). The elements can be extracted with the functions model, solution, and solver_output.

Author(s)

Hossein Kamalzadeh, Michael Hahsler

References

[1] For further details on how the POMDP solver utilized in this R package works check the following website: http://www.pomdp.org

[2] Cassandra, A. Rocco, Exact and approximate algorithms for partially observable Markov decision processes, (1998). https://dl.acm.org/citation.cfm?id=926710

Examples

```
data("TigerProblem")
TigerProblem

tiger_solved <- solve_POMDP(model = TigerProblem, parameter = list(fg_points = 10))
tiger_solved

## look at the model
model(tiger_solved)

## look at the solution
solution(tiger_solved)
```

```
## look at solver output
solver_output(tiger_solved)

## plot the policy graph
plot(tiger_solved)

## display available solver options which can be passed on to the solver as parameter.
solve_POMDP_parameter()

## solve a POMDP from http://www.pomdp.org/examples
sol <- solve_POMDP("http://www.pomdp.org/examples/cheese.95.POMDP")
sol
plot(sol)
```

TigerProblem   *Tiger Problem POMDP Specification*

Description

The model for the Tiger Problem [1].

Usage

```
data("TigerProblem")
```

Format

A list with the elements: discount, states, actions, observations, start, transition_prob, observation_prob, reward, name.

Details

The Tiger Problem is defined as follows [1]. A tiger is put with equal probability behind one of two doors, while treasure is put behind the other one. You are standing in front of the two closed doors and need to decide which one to open. If you open the door with the tiger, you will get hurt by the tiger (negative reward), but if you open the door with the treasure, you receive a positive reward. Instead of opening a door right away, you also have the option to wait and listen for tiger noises. But listening is neither free nor entirely accurate. You might hear the tiger behind the left door while it is actually behind the right door and vice versa.

The states of the system are tiger behind the left door (tiger-left) and tiger behind the right door (tiger-right).

Available actions are: open the left door (open-left), open the right door (open-right) or to listen (listen).

Rewards associated with these actions depend on the resulting state: +10 for opening the correct door (the door with treasure), -100 for opening the door with the tiger. A reward of -1 is the cost of listening.

As a result of listening, there are two observations: either you hear the tiger on the right (tiger-right), or you hear it on the left (tiger-left).

The transition probability matrix for the action listening is identity, i.e., the position of the tiger does not change. Opening either door means that the game restarts by placing the tiger uniformly behind one of the doors.

References

[1] Anthony R. Cassandra, Leslie P Kaelbling, and Michael L. Littman (1994). Acting Optimally in Partially Observable Stochastic Domains. In Proceedings of the Twelfth National Conference on Artificial Intelligence, pp. 1023-1028.

Examples

```
data(TigerProblem)
TigerProblem

# solve the problem and look at the optimal policy graph (as a table and as a plot)
sol <- solve_POMDP(TigerProblem)
sol

solution(sol)$pg
plot(sol)
```

---

write_POMDP	*Write a POMDP Model to a File in POMDP Format*

Description

Writes a POMDP file suitable for the pomdp-solve program. This function is used internally.

Usage

```
write_POMDP(model, file)
```

Arguments

model   an object of class POMDP_model.

file   a file name.

Author(s)

Hossein Kamalzadeh, Michael Hahsler

References

POMDP solver website: http://www.pomdp.org

See Also

POMDP

Index

Bibliography

Ackermann, Ronald T., Sharon L. Edelstein, K. M. Venkat Narayan, Ping Zhang, Michael M. Engelgau, William H. Herman, David G. Marrero. 2009. Changes in health state utilities with changes in body mass in the diabetes prevention program. *Obesity* **17**(12) 2176–2181. doi:10.1038/oby.2009.114. URL `http://dx.doi.org/10.1038/oby.2009.114`.

Ahsen, Mehmet Eren, Elizabeth S Burnside. 2018. Preference-sensitive management of post-mammography decisions in breast cancer. *Production and Operations Management* **27**(12) 2313–2338. doi:10.1111/poms.12897.

Ahuja, Vishal, John R. Birge. 2018. An approximation approach for response adaptive clinical trial design. *SSRN Electronic Journal* doi:10.2139/ssrn.3212148. URL `https://www.ssrn.com/abstract=3212148`.

Alagoz, Oguzhan, Cindy L. Bryce, Steven Shechter, Andrew Schaefer, Chung-Chou H. Chang, Derek C. Angus, Mark S. Roberts. 2005. Incorporating biological natural history in simulation models: Empirical estimates of the progression of end-stage liver disease. *Medical Decision Making* **25**(6) 620–632. doi:10.1177/0272989X05282719. URL `http://www.ncbi.nlm.nih.gov/pubmed/16282213`.

Alagoz, Oguzhan, Heather Hsu, Andrew J. Schaefer, Mark S. Roberts. 2010. Markov decision processes: A tool for sequential decision making under uncertainty. *Medical Decision Making* **30**(4) 474–483. doi:10.1177/0272989X09353194. URL `http://www.ncbi.nlm.nih.gov/pubmed/20044582`.

Alagoz, Oguzhan, Lisa M. Maillart, Andrew J. Schaefer, Mark S. Roberts. 2004. The optimal timing of living-donor liver transplantation. *Management Science* **50**(10) 1420–1430. doi:10.1287/mnsc.1040.0287. URL `http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1040.0287`.

American Diabetes Association, American Diabetes. 2018. Economic costs of diabetes in the u.s. in 2017. *Diabetes care* **41**(5) 917–928. doi:10.2337/dci18-0007. URL `http://www.ncbi.nlm.nih.gov/pubmed/29567642`.

Aoki, Masanao. 1965. Optimal control of partially observable markovian systems. *Journal of the Franklin Institute* **280**(5) 367–386. doi:10.1016/0016-0032(65)90528-4. URL `https://www.sciencedirect.com/science/article/pii/0016003265905284`.

Astrom, K.J. 1965. Optimal control of markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications* **10**(1) 174–205. doi:10.1016/0022-247X(65)90154-X. URL `https://www.sciencedirect.com/science/article/pii/0022247X6590154X`.

Ayer, Turgay, Oguzhan Alagoz, Natasha K. Stout. 2012a. Or forum—a pomdp approach to personalize mammography screening decisions. *Operations Research* **60**(5) 1019–1034. doi:10.1287/opre.1110.1019. URL `http://pubsonline.informs.org/doi/abs/10.1287/opre.1110.1019`.

Ayer, Turgay, Oguzhan Alagoz, Natasha K. Stout. 2012b. Or forum—a pomdp approach to personalize mammography screening decisions. *Operations Research* **60**(5) 1019–1034. doi:10.1287/opre.1110.1019. URL `http://pubsonline.informs.org/doi/abs/10.1287/opre.1110.1019`.

Bahia, Luciana, Rosane Kupfer, Denise Momesso, Debora A P Cabral, Balduino Tschiedel, Marcia Puñales, Suzana Lavigne, Cristina F S Façanha, Adriana C Forti, Angela D N Mendes, et al. 2017. Health-related quality of life and utility values associated to hypoglycemia in patients with type 1 diabetes mellitus treated in the brazilian public health system: a multicenter study. *Diabetol Metab Syndr* **9**(9). doi:10.1186/s13098-017-0206-4. URL `https://link.springer.com/content/pdf/10.1186%2Fs13098-017-0206-4.pdf`.

Bertsimas, Dimitris, Nathan Kallus, Alexander M Weinstein, Ying Daisy Zhuo. 2017. Personalized diabetes management using electronic medical records. *Diabetes care* **40**(2) 210–217. doi:10.2337/dc16-0826. URL `http://www.ncbi.nlm.nih.gov/pubmed/27920019`.

Bertsimas, Dimitris, Allison O'Hair, Stephen Relyea, John Silberholz. 2016. An analytics approach to designing combination chemotherapy regimens for cancer. *Management*

Science **62**(5) 1511–1531. doi:10.1287/mnsc.2015.2363. URL `http://pubsonline.informs.org/doi/10.1287/mnsc.2015.2363`.

Bjork, S. 2001. The cost of diabetes and diabetes care. *Diabetes Res Clin Pract* **54 Suppl 1**(2001) S13–8. doi:10.1016/S0168-8227(01)00304-7. URL `http://ac.els-cdn.com/S0168822701003047/1-s2.0-S0168822701003047-main.pdf?_tid=6f3fa66807ee1d4a807e23ff3d342170&acdnat=1343026057_c31afa09e285d13d2a3c901f85b7f5ae`.

Brafman, Ronen I. 1997. A heuristic variable grid solution method for pomdps. *Aaai/Iaai* 727–733.

Calonge, N, DB Petitti. 2008. Screening for type 2 diabetes mellitus in adults: Us preventive services task force recommendation statement. *Annals of Internal Medicine* **148**(11) 846–54. URL `http://www.ncbi.nlm.nih.gov/pubmed/18519930`.

Care, Medical. 2013. Standards of medical care in diabetes - 2013. *Diabetes Care* **36**(SUPPL.1). doi:10.2337/dc13-S011.

Cassandra, Anthony Rocco. 1998. Exact and approximate algorithms for partially observable markov decision processes. *Doctoral Dissertation* URL `https://dl.acm.org/citation.cfm?id=926710`.

Cassandra, Anthony R. 1997. A survey of pomdp applications. *Uncertainty in Artificial Intelligence* **Fall Sympo** 472–480.

Cassandra, Anthony R., Michael L Littman, Nevin Lianwen Zhang. 2013. Incremental pruning: A simple, fast, exact method for partially observable markov decision processes URL `https://arxiv.org/ftp/arxiv/papers/1302/1302.1525.pdf`.

Cassandra, Anthony Rocco, Leslie Pack Kaelbling, Michael L Littman. 1994. Acting optimally in partially observable stochastic domains. *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI)*, vol. 2. 1023–1028. URL `http://www.scopus.com/inward/record.url?eid=2-s2.0-0028564629&partnerID=40`.

CDCP. 2017. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States. Atlanta, GA: Centers for Disease Control and Prevention; 2014*. Cdc, US Department of Health and Human Services. doi:10.1177/1527154408322560.

Chatterjee, Ranee, K M Venkat Narayan, Joseph Lipscomb, Sandra L Jackson, Qi Long, Ming Zhu, Lawrence S Phillips. 2013. Screening for diabetes and prediabetes should be cost-saving in patients at high risk. *Diabetes care* **36**(7) 1981–7. doi:10.2337/dc12-1752. URL `http://www.ncbi.nlm.nih.gov/pubmed/23393215`.

Chen, T H, M F Yen, T H Tung. 2001. A computer simulation model for cost-effectiveness analysis of mass screening for type 2 diabetes mellitus. *Diabetes research and clinical practice* **54 Suppl 1**(1) S37–42. doi:10.1016/S0168-8227(01)00307-2. URL `http://www.ncbi.nlm.nih.gov/pubmed/11580967`.

Chhatwal, Jagpreet, Oguzhan Alagoz, Elizabeth S Burnside. 2010. Optimal breast biopsy decision-making based on mammographic features and demographic factors. *Operations research* **58**(6) 1577–1591. doi:10.1287/opre.1100.0877. URL `http://www.ncbi.nlm.nih.gov/pubmed/21415931`.

Collins, Gary S, Susan Mallett, Omar Omar, Ly-Mee Yu. 2011. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Medicine* **9**(1) 103. doi:10.1186/1741-7015-9-103. URL `http://www.ncbi.nlm.nih.gov/pubmed/21902820`.

Davies, Melanie J. 1999. Is screening for type 2 diabetes justified? *Practitioner* **243**(1595) 93–100. URL `http://www.who.int/diabetes/publications/en/screening_mnc03.pdf`.

de Goeij, M. C. M., M. van Diepen, K. J. Jager, G. Tripepi, C. Zoccali, F. W. Dekker. 2013. Multiple imputation: dealing with missing data. *Nephrology Dialysis Transplantation* **28**(10) 2415–2420. doi:10.1093/ndt/gft221. URL `http://www.ncbi.nlm.nih.gov/pubmed/23729490`.

Denton, Brian T. 2018. *Optimization of Sequential Decision Making for Chronic Diseases: From Data to Decisions*. INFORMS, 316–348. doi:10.1287/educ.2018.0184. URL `http://pubsonline.informs.org/doi/10.1287/educ.2018.0184`.

Deo, Sarang, Kumar Rajaram, Sandeep Rath, Uday S. Karmarkar, Matthew B. Goetz. 2015. Planning for hiv screening, testing, and care at the veterans health administration. *Operations Research* **63**(2) 287–304. doi:10.1287/opre.2015.1353. URL `http://pubsonline.informs.org/doi/10.1287/opre.2015.1353`.

Dimitrakakis, Christos, Ronald Ortner. 2018. Decision making under uncertainty and reinforcement learning.

Drake, Alvin W. 1962. Observation of a markov process through a noisy channel. Ph.D. thesis, Massachusetts Institute of Technology. URL `https://dspace.mit.edu/handle/1721.1/11341`.

Dynkin, E. B. 1965. Controlled random sequences. *Theory of Probability Its Applications* **10**(1) 1–14. doi:10.1137/1110001. URL `http://epubs.siam.org/doi/10.1137/1110001`.

Eborall, Helen, Margaret Stone, Navneet Aujla, Nicholas Taub, Melanie Davies, Kamlesh Khunti. 2012. Influences on the uptake of diabetes screening: a qualitative study in primary care. *The British journal of general practice: the journal of the Royal College of General Practitioners* **62**(596) 4–11. doi:10.3399/bjgp12X630106. URL `http://www.ncbi.nlm.nih.gov/pubmed/22429438`.

Furukawa, Nagata. 1967. A bayes controlled process. *Memoirs of the Faculty of Science, Kyushu University. Series A, Mathematics* **21**(2) 249–258. doi:10.2206/kyushumfs.21.249. URL `https://www.jstage.jst.go.jp/article/kyushumfs/21/2/21_2_249/_article/-char/ja/`.

Gafa, Marija, Lalit Garg, Giovanni Masala, Sally I Mcclean, Marco Micocci, Giuseppina Cannas. 2012. Hiv disease progression modelling using phase type survival trees. *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*. IEEE,

2–5. doi:10.1109/CBMS.2012.6266408. URL http://ieeexplore.ieee.org/document/6266408/.

García-Mora, B., C. Santamaría, E. Navarro, G. Rubio. 2010. Modeling bladder cancer using a markov process with multiple absorbing states. *Mathematical and Computer Modelling* **52**(7–8) 977–982. doi:10.1016/j.mcm.2010.02.036. URL http://dx.doi.org/10.1016/j.mcm.2010.02.036.

Grosfeld-Nir, Abraham. 1996. A two-state partially observable markov decision process with uniformly distributed observations. *Operations Research* **44**(3) 458–463. doi:10.1287/opre.44.3.458. URL http://pubsonline.informs.org/doi/abs/10.1287/opre.44.3.458.

Group, Diabetes Prevention Program Research. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine* **346**(6) 393–403. doi:10.1056/NEJMoa012512. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa012512.

Hauskrecht, M. 2011. Value-function approximations for partially observable markov decision processes. *Journal Of Artificial Intelligence Research* doi:10.1613/jair.678. URL http://arxiv.org/abs/1106.0234.

Hauskrecht, Milos. 2000. Value-function approximations for partially observable markov decision processes. *J. Artif. Int. Res.* **13**(1) 33–94. doi:10.1613/jair.678. URL http://dl.acm.org.proxy.lib.uiowa.edu/citation.cfm?id=1622262.1622264.

Hinderer, K. 1970. *Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter*. Springer Berlin Heidelberg. doi:10.1007/978-3-642-46229-0. URL http://link.springer.com/10.1007/978-3-642-46229-0.

Hoerger, Thomas J, Katherine A Hicks, Albert D Bethke. 2004. *A Markov Model of Disease Progression and Cost-Effectiveness for Type 2 Diabetes*, vol. 3000. Centers for Disease Control and Prevention.

Hoey, Jesse, Pascal Poupart. 2005. Solving pomdps with continuous or large discrete observation spaces. *IJCAI International Joint Conference on Artificial Intelligence* 1332–1338.

Howard, Kirsten, Sarah White, Glenn Salkeld, Stephen McDonald, Jonathan C. Craig, Steven Chadban, Alan Cass. 2010. Cost-effectiveness of screening and optimal management for diabetes, hypertension, and chronic kidney disease: A modeled analysis. *Value in Health* **13**(2) 196–208. doi:10.1111/j.1524-4733.2009.00668.x. URL `http://dx.doi.org/10.1111/j.1524-4733.2009.00668.x`.

Huang, Xuedong., Alejandro. Acero, Hsiao-Wuen. Hon. 2001. *Spoken language processing:a guide to theory, algorithm, and system development*. Prentice Hall PTR, Upper Saddle River, NJ. URL `https://dl.acm.org/citation.cfm?id=560905`.

Jackson, Christopher H, Linda D Sharples, Simon G Thompson, Stephen W Duffy, Elisabeth Couto. 2003. Multistate markov models for disease progression with classification error. *The Statistician* **52**(2) 193–209. doi:10.1111/1467-9884.00351.

Jahani, Meysam, Mahdi Mahdavi. 2016. Comparison of predictive models for the early diagnosis of diabetes. *Healthcare Informatics Research* **22**(2) 95. doi:10.4258/hir.2016.22.2.95. URL `http://www.ncbi.nlm.nih.gov/pubmed/27200219`.

Kaelbling, Leslie Pack, Michael L. Littman, Anthony R. Cassandra. 1998a. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* **101**(1–2) 99–134. doi:10.1016/S0004-3702(98)00023-X. URL `http://linkinghub.elsevier.com/retrieve/pii/S000437029800023X`.

Kaelbling, Leslie Pack, Michael L. Littman, Anthony R. Cassandra. 1998b. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* **101**(1–2) 99–134. doi:10.1016/S0004-3702(98)00023-X. URL `http://linkinghub.elsevier.com/retrieve/pii/S000437029800023X`.

Kahn, Richard, Peter Alperin, David Eddy, Knut Borch-Johnsen, John Buse, Justin Feigelman, Edward Gregg, Rury R. Holman, M. Sue Kirkman, Michael Stern, et al. 2010.

Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis. *The Lancet* **375**(9723) 1365–1374. doi:10.1016/S0140-6736(09)62162-0. URL `http://dx.doi.org/10.1016/S0140-6736(09)62162-0`.

Kamalzadeh, Hossein, Michael Hahsler. 2019. pomdp: Solver for partially observable markov decision processes (pomdp). *r-project* URL `https://cran.r-project.org/package=pomdp`. R package version 0.9.0.

Khunti, Kamlesh, Hamidreza Mani, Felix Achana, Nicola Cooper, Laura J. Gray, Melanie J. Davies. 2015. Systematic review and meta-analysis of response rates and diagnostic yield of screening for type 2 diabetes and those at high risk of diabetes. *PLOS ONE* **10**(9) e0135702. doi:10.1371/journal.pone.0135702. URL `http://dx.plos.org/10.1371/journal.pone.0135702`.

Kochanek, Kenneth D, Sherry L Murphy, Jiaquan Xu, Elizabeth Arias. 2016. Mortality in the united states, 2016 key findings data from the national vital statistics system. *National Center for Health Statistics* URL `https://www.cdc.gov/nchs/data/databriefs/db293.pdf`.

Kuo, H S, H J Chang, P Chou, L Teng, T H Chen. 1999. A markov chain model to assess the efficacy of screening for non-insulin dependent diabetes mellitus (niddm). *Int J Epidemiol* **28** 233–240.

Lee, Chris P., Glenn M. Chertow, Stefanos A. Zenios. 2008. Optimal initiation and management of dialysis therapy. *Operations Research* **56**(6) 1428–1449. doi:10.1287/opre.1080.0613. URL `http://pubsonline.informs.org/doi/abs/10.1287/opre.1080.0613`.

Lee, S, J Ko, X Tan, I Patel, R Balkrishnan, J Chang. 2014. Markov chain modelling analysis of hiv/aids progression: A race-based forecast in the united states. *Indian Journal of Pharmaceutical Sciences* **76**(2) 107–115. doi:IJPhS-76-107 [pii].

Littman, Michael L. 2009. A tutorial on partially observable markov decision processes.

Journal of Mathematical Psychology **53**(3) 119–125. doi:10.1016/J.JMP.2009.01.005. URL
`https://www.sciencedirect.com/science/article/pii/S0022249609000042`.

Liu, Yu-Ying, Shuang Li, Fuxin Li, Le Song, James M Rehg. 2015. Efficient learning
of continuous-time hidden markov models for disease progression. *Advances in neural
information processing systems* **28** 3599–3607. URL `http://www.ncbi.nlm.nih.gov/`
`pubmed/27019571`.

Loukine, Lidia, Chris Waters, Bernard CK Choi, Joellyn Ellison. 2012. Impact of diabetes
mellitus on life expectancy and health-adjusted life expectancy in canada. *Population
Health Metrics* **10**(1) 7. doi:10.1186/1478-7954-10-7. URL `http://www.ncbi.nlm.nih.`
`gov/pubmed/22531113`.

Lovejoy, William S. 1991a. Computationally feasible bounds for partially observed markov
decision processes. *Operations Research* **39**(1) 162–175. doi:10.1287/opre.39.1.162. URL
`http://pubsonline.informs.org/doi/abs/10.1287/opre.39.1.162`.

Lovejoy, William S. 1991b. A survey of algorithmic methods for partially observed markov
decision processes. *Annals of Operations Research* **28**(1) 47–65. doi:10.1007/BF02055574.
URL `http://link.springer.com/10.1007/BF02055574`.

Lu, Zhong X, K Walker, O 'dea, Kerin, Ken A Sikaris, Jonathan Shaw. 2010. A1c for
screening and diagnosis of type 2 diabetes in routine clinical practice "a1c for screening
and diagnosis of type 2 diabetes in routine clinical practice" diabetes care. *American
Diabetes Association* **33**(4) 817–819. doi:10.2337/dc09. URL `https://doi.org/10.2337/`
`dc09-1763`.

Maillart, Lisa M., Julie Simmons Ivy, Scott Ransom, Kathleen Diehl. 2008. Assessing
dynamic breast cancer screening policies. *Operations Research* **28**(6) 1411–1427.
doi:10.1287/opre.1080.0614. URL `http://pubsonline.informs.org/doi/abs/10.1287/`
`opre.1080.0614`.

Monahan, George E. 1982. Survey of partially observable markov decision processes - theory,

models, and algorithms. *Management Science* **28**(1) 1–16. doi:10.1287/mnsc.28.1.1. URL `https://www.jstor.org/stable/2631070`.

Monahan, George E. 2008. A survey of partially observable markov decision processes: Theory, models, and algorithms. *Management Science* **28** 1–16. doi:10.2307/2631070. URL `https://www.jstor.org/stable/2631070`.

Murphy, Sherry L, Jiaquan Xu, Kenneth D Kochanek, Sally C Curtin, Elizabeth Arias. 2015. National vital statistics reports volume 66, number 6, november 27, 2017 URL `https://www.cdc.gov/nchs/data/nvsr/nvsr66/nvsr66_06.pdf`.

Neumann, Anne, Olaf Schoffer, Fredrik Norström, Margareta Norberg, Stefanie J Klug, Lars Lindholm. 2014a. Health-related quality of life for pre-diabetic states and type 2 diabetes mellitus: a cross-sectional study in västerbotten sweden. *Health and Quality of Life Outcomes* **12**(1) 150. doi:10.1186/s12955-014-0150-z. URL `http://hqlo.biomedcentral.com/articles/10.1186/s12955-014-0150-z`.

Neumann, Peter J., Joshua T. Cohen, Milton C. Weinstein. 2014b. Updating cost-effectiveness — the curious resilience of the \$50,000-per-qaly threshold. *New England Journal of Medicine* **371**(9) 796–797.

Orton, E., A. Forbes-Haley, L. Tunbridge, S. Cohen. 2013. Equity of uptake of a diabetic retinopathy screening programme in a geographically and socio-economically diverse population. *Public Health* **127**(9) 814–821. doi:10.1016/J.PUHE.2013.04.015. URL `https://www.sciencedirect.com/science/article/pii/S0033350613001303`.

O'connor, Patrick J, William A Rush, Linda M Cherney, Nicolaas P Pronk. 2001. Screening for diabetes mellitus in high-risk patients: Cost, yield, and acceptability patterns of practice. *Eff Clin Pract* **4**. URL `https://pdfs.semanticscholar.org/1fee/88305cd5f826045b451137b308789cae6630.pdf`.

Park, Paul, Rebecca K Simmons, A Toby Prevost, Simon J Griffin. 2008. Screening for type 2 diabetes is feasible, acceptable, but associated with increased short-term anxiety: A randomised controlled trial in british general practice. *BMC Public Health* **8**(1)

350. doi:10.1186/1471-2458-8-350. URL `http://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-8-350`.

Patrick, Jonathan. 2011. Access to long-tem care: The true cause of hospital congestion. *Production and Operations Management* **20**(3) 347–358.

Petersen, Matt. 2016. Economic costs of diabetes in the u.s. in 2012. *Diabetes Care* **39**(7) 1033–1046. doi:10.2337/dc12-2625.

Platzman, Loren K. 1980. Optimal infinite-horizon undiscounted control of finite probabilistic systems. *SIAM Journal on Control and Optimization* **18**(4) 362–380. doi:10.1137/0318028. URL `http://epubs.siam.org/doi/10.1137/0318028`.

Poupart, Pascal, Craig Boutilier. 2004. Bounded finite state controllers. *Neural Information Processing Systems 16 (NIPS 2003)*. 823–830. URL `https://papers.nips.cc/paper/2372-bounded-finite-state-controllers`.

prevention Program, Diabetes. 2008. Diabetes prevention program (dpp) study design and goals. *New England Journal of Medicine.* 1–6URL `https://www.niddk.nih.gov/about-niddk/research-areas/diabetes/diabetes-prevention-program-dpp/Documents/DPP_508.pdf`.

Puterman, Martin L. 2005. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, vol. 37. 3rd ed. Wiley. doi:10.2307/1269932. URL `http://www.jstor.org/stable/1269932?origin=crossref`.

Rhenius, Detlef. 1974. Incomplete information in markovian decision models. *The Annals of Statistics* **2**(6) 1327–1334. doi:10.1214/aos/1176342886. URL `http://projecteuclid.org/euclid.aos/1176342886`.

Rieder, Ulrich. 1975. Bayesian dynamic programming. *Advances in Applied Probability* **7**(2) 330–348. doi:10.2307/1426080. URL `https://www.cambridge.org/core/product/identifier/S0001867800046012/type/journal_article`.

Sandikci, B., L. M. Maillart, a. J. Schaefer, Oguzhan Alagoz, Mark S. Roberts. 2008. Estimating the patient's price of privacy in liver transplantation. *Operations Research* **56**(6) 1393–1410. doi:10.1287/opre.1080.0648.

Sandikci, Burhaneddin, James J Cochran, Louis A Cox, Pinar Keskinocak, Jeffrey P Kharoufeh, J Cole Smith. 2010. Reduction of a POMDP to an MDP. *Wiley Encyclopedia of Operations Research and Management Science* 1–9URL `http://onlinelibrary.wiley.com/doi/10.1002/9780470400531.eorms0710/abstract`.

Sandikci, Burhaneddin, Lisa M Maillart, Andrew J. Schaefer, Mark S. Roberts. 2013. Alleviating the patient's price of privacy through a partially observable waiting list. *Management Science* **59**(8) 1836–1854. doi:10.1287/mnsc.1120.1671. URL `http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1120.1671`.

Santoso, L., I.M.Y. Mareels. 2001. Markovian framework for diabetes control. *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228)* **2** 1705–1710. doi:10.1109/.2001.981147.

Satia, Jay K., Roy E. Lave. 1973. Markovian decision processes with uncertain transition probabilities. *Operations Research* **21**(3) 728–740. doi:10.1287/opre.21.3.728. URL `http://pubsonline.informs.org/doi/abs/10.1287/opre.21.3.728`.

Sawaragi, Yoshikazu, Tsuneo Yoshikawa. 1970. Discrete-time markovian decision processes with incomplete state observation. *The Annals of Mathematical Statistics* **41**(1) 78–86. doi:10.1214/aoms/1177697190. URL `http://projecteuclid.org/euclid.aoms/1177697190`.

Schaefer, Andrew J, Matthew D Bailey, Steven M Shechter, Mark S Roberts. 2004. Modeling medical treatment using markov decision processes. *Operations research and health care* 593–612doi:10.1007/1-4020-8066-2_23.

Shams, Issac, Saeede Ajorlou, Kai Yang. 2015. A predictive analytics approach to reducing 30-day avoidable readmissions among patients with heart failure, acute myocardial infarction, pneumonia, or copd. *Health Care Management Science* **18**(1)

19–34. doi:10.1007/s10729-014-9278-y. URL `http://link.springer.com/10.1007/s10729-014-9278-y`.

Shechter, Steven M., Matthew D. Bailey, Andrew J. Schaefer, Mark S. Roberts. 2008. The optimal time to initiate hiv therapy under ordered health states. *Operations Research* **56**(1) 20–33. doi:10.1287/opre.1070.0480. URL `http://pubsonline.informs.org/doi/abs/10.1287/opre.1070.0480`.

Shih, Hui-Chuan, Pesus Chou, Chi-Ming Liu, Tao-Hsin Tung. 2007. Estimation of progression of multi-state chronic disease using the markov model and prevalence pool concept. *BMC medical informatics and decision making* **7** 34. doi:10.1186/1472-6947-7-34. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2241590&tool=pmcentrez&rendertype=abstract`.

Siebert, Uwe, Oguzhan Alagoz, Ahmed M. Bayoumi, Beate Jahn, Douglas K. Owens, David J. Cohen, Karen M. Kuntz. 2012. State-transition modeling. *Medical Decision Making* **32**(5) 690–700. doi:10.1177/0272989X12455463. URL `http://www.ncbi.nlm.nih.gov/pubmed/22990084`.

Sipser, Michael. 1997. *Introduction to the theory of computation*. PWS Pub. Co.

Sukkar, Rafid, Elyse Katz, Yanwei Zhang, David Raunig, Bradley T. Wyman. 2012. Disease progression modeling using hidden markov models. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* **2012** 2845–8. doi:10.1109/EMBC.2012.6346556. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6346556%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/23366517`.

Sukkar, Rafid, Bradley Wyman, Elyse Katz, Yanwei Zhang, David Raunig. 2011. Modeling alzheimer's disease progression using hidden markov models. *Alzheimer's & Dementia* **7**(4) S147. doi:10.1016/j.jalz.2011.05.397. URL `https://linkinghub.elsevier.com/retrieve/pii/S155252601100536X`.

Team, R Core. 2018. R: A language and environment for statistical computing URL `https://www.r-project.org/`.

UK, Diabetes. 2010. *UK Diabetes* **692** 1–21. URL `https://www.diabetes.org.uk/resources-s3/2017-11/diabetes_in_the_uk_2010.pdf`.

Wang, Xiang, David Sontag, Fei Wang. 2014. Unsupervised learning of disease progression models doi:10.1145/2623330.2623754. URL `http://dx.doi.org/10.1145/2623330.2623754`.

Wessels, Jacobus. 1968. Decision rules in markovian decision processes with incompletely known transition probabilities URL `https://www.semanticscholar.org/paper/Decision-rules-in-Markovian-decision-processes-with-Wessels/ccba6474a981841c95405d2a15ef20f7913ab833`.

White, Chelsea C. 1975. Cost equality and inequality results for a partially observed stochastic optimization problem. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-5**(6) 576–582. doi:10.1109/TSMC.1975.4309398. URL `http://ieeexplore.ieee.org/document/4309398/`.

White, Chelsea C. 1976. Procedures for the solution of a finite-horizon, partially observed, semi-markov optimization problem. *Operations Research* **24**(2) 348–358. doi:10.1287/opre.24.2.348. URL `http://pubsonline.informs.org/doi/abs/10.1287/opre.24.2.348`.

Will M. Bertrand, J, Jan C Fransoo. 2002. Operations management research methodologies using quantitative modeling. *International Journal of Operations & Production Management* **22**(2) 241–264.

Yuen, M.-F., C.-L. Lai. 2003. Screening for hepatocellular carcinoma: survival benefit andcost-effectiveness. *Annals of Oncology* **14**(10) 1463–1467. doi:10.1093/annonc/mdg400. URL `https://academic.oup.com/annonc/article-lookup/doi/10.1093/annonc/mdg400`.

Zargoush, Manaf, Stella S Daskalopoulou. 2018. Designing risk-adjusted therapy for patients with hypertension. *Production and Operations Management* **27**(12) 2291–2312. doi:10.1111/poms.12872.

Zhang, Nevin L, Wenju Liu. 1996. *Planning in Stochastic Domains: Problem Characteristics and Approximation*. Version II. URL `https://www.researchgate.net/profile/Nevin_Zhang/publication/2825882_Planning_in_Stochastic_Domains_Problem_Characteristics_and_Approximation/links/546b04f50cf20dedafd3a417/Planning-in-Stochastic-Domains-Problem-Characteristics-and-Approximation.pdf`.

Zhang, Ping, Morton B Brown, Dori Bilik, Ronald T Ackermann, Rui Li, William H Herman. 2012a. Health utility scores for people with type 2 diabetes in u.s. managed care health plans: results from translating research into action for diabetes (triad). *Diabetes care* **35**(11) 2250–6. doi:10.2337/dc11-2478. URL `http://www.ncbi.nlm.nih.gov/pubmed/22837369`.

Zhang, Ping, Morton B Brown, Dori Bilik, Ronald T Ackermann, Rui Li, William H Herman. 2012b. Health utility scores for people with type 2 diabetes in u.s. managed care health plans: results from translating research into action for diabetes (triad). *Diabetes care* **35**(11) 2250–6. doi:10.2337/dc11-2478. URL `http://www.ncbi.nlm.nih.gov/pubmed/22837369`.

Zhang, Ping, Michael M Engelgau, Rodolfo Valdez, Stephanie M Benjamin, Betsy Cadwell, K M Venkat Narayan. 2003. Costs of screening for pre-diabetes among us adults: a comparison of different screening strategies. *Diabetes care* **26**(9) 2536–42. doi:10.2337/DIACARE.26.9.2536. URL `http://www.ncbi.nlm.nih.gov/pubmed/12941715`.