Automatica 50 (2014) 378–388

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Attention allocation for decision making queues☆

Vaibhav Srivastava[a,1], Ruggero Carli[b], Cédric Langbort[c], Francesco Bullo[d]

[a] Department of Mechanical & Aerospace Engineering, Princeton University, NJ, USA

[b] Department of Information Engineering, University of Padova, Italy

[c] Coordinated Science Laboratory, University of Illinois at Urbana–Champaign, USA

[d] Center for Control, Dynamical Systems, and Computation, University of California, Santa Barbara, USA

ARTICLE INFO

Article history:
Received 21 February 2012
Received in revised form
13 May 2013
Accepted 16 October 2013
Available online 23 December 2013

Keywords:
Sigmoid utility
Human decision making
Control of queues
Human–robot interaction

ABSTRACT

We consider the optimal servicing of a queue with sigmoid server performance. There are various systems with sigmoid server performance, including systems involving human decision making, visual perception, human–machine communication and advertising response. Tasks arrive at the server according to a Poisson process. Each task has a deadline that is incorporated as a latency penalty. We investigate the trade-off between the reward obtained by processing the current task and the penalty incurred due to the tasks waiting in the queue. We study this optimization problem in a Markov decision process (MDP) framework. We characterize the properties of the optimal policy for the MDP and show that the optimal policy may drop some tasks; that is, may not process a task at all. We determine an approximate solution to the MDP using the certainty-equivalent receding horizon optimization framework and derive performance bounds on the proposed receding horizon policy. We also suggest guidelines for the design of such queues.

1. Introduction

The recent national robotic initiative (Guizzo, 2011) motivates research and applications emphasizing the interaction of humans with symbiotic co-robot partners. Such co-robots will facilitate better interaction between the human partner and the automaton. In complex and information rich environments, one of the key roles for these co-robots is to help the human partner focus their attention efficiently. A particular example of such a setting is a surveillance mission in which the human operator monitors the evidence collected by the autonomous agents (Bulkeley, 2009; Drew, 2010). The excessive amount of information available in such systems often results in poor decisions by the human operator (Shanker & Richtel, 2011). This emphasizes the need for the development of a support system that helps the human operator to focus their attention.

Recently, there has been a significant interest in understanding the mechanisms of human decision making (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006). Several mathematical models for human decision making have been proposed (Bogacz et al., 2006; Pew, 1969; Wickens & Hollands, 2000). These models suggest that the correctness of the decision of a human operator in a binary decision making scenario evolves as a sigmoid function of the time allocated for the decision. When a human operator has to serve a queue of decision making tasks *in real time*, the tasks (e.g., feeds from a camera network) waiting in the queue lose value continuously. This trade-off between the correctness of the decision and the loss in the value of the pending tasks is of critical importance for the performance of the human operator. In this paper, we address this trade-off, and determine the optimal duration allocation policies for the human operator serving such a decision making queue. The sigmoid function has also been used to model the quality of human–machine communication (Wickens & Hollands, 2000), human performance in multiple target search (Hong & Drury, 2002), advertising response function (Vakratsas, Feinberg, Bass, & Kalyanaram, 2004), and expected profit in simultaneous bidding (Rothkopf, 1977). Therefore, the analysis presented in this paper can also be used to determine optimal human–machine communication policies, optimal search strategies, the optimal advertisement duration allocation, and optimal bidding strategies. In this paper, we generically refer to the server with sigmoid performance as a human operator and the tasks as the decision making tasks.

There has been a significant interest in the study of the performance of a human operator serving a queue. In an early work, Schmidt (1978) models the human as a server and numerically studies a queueing model to determine the performance of a human air traffic controller. Recently, Savla, Temple, and Frazzoli

☆ This work has been supported in part by AFOSR MURI Award FA9550-07-1-0528. The material in this paper was partially presented at the 2011 American Control Conference (ACC 2011), June 29–July 1, 2011, San Francisco, California, USA (Srivastava, Carli, Bullo, & Langbort, 2011). This paper was recommended for publication in revised form by Associate Editor Antonis Papachristodoulou under the direction of Editor Frank Allgöwer.

E-mail addresses: vaibhavs@princeton.edu (V. Srivastava), carlirug@dei.unipd.it (R. Carli), langbort@illinois.edu (C. Langbort), bullo@engineering.ucsb.edu (F. Bullo).

[1] Tel.: +1 609 436 0292; fax: +1 6092586109.

0005-1098/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.automatica.2013.11.028

(2008) study human supervisory control for unmanned aerial vehicle operations: they model the system by a simple queueing network with two components in series, the first of which is a spatial queue with vehicles as servers and the second is a conventional queue with human operators as servers. They design joint motion coordination and operator scheduling policies that minimize the expected time needed to classify a target after its appearance. The performance of the human operator based on their utilization history has been incorporated to design maximally stabilizing task release policies for a human-in-the-loop queue in Savla and Frazzoli (2010, 2012). Bertuccelli, Pellegrino, and Cummings (2010) study the human supervisory control as a queue with re-look tasks. They study the policies in which the operator can put the tasks in an orbiting queue for a re-look later. An optimal scheduling problem in the human supervisory control is studied in Bertuccelli, Beckers, and Cummings (2010). Crandall, Cummings, Della Penna, and de Jong (2011) study optimal scheduling policy for the operator and discuss if the operator or the automaton should be ultimately responsible for selecting the task. Powel and Morgansen (2012) model mixed team of humans and robots as a multi-server queue and incorporate a human fatigue model to determine the performance of the team. They present a comparative study of the fixed and the rolling work-shifts of the operators.

The optimal control of queueing systems (Sennott, 1999) is a classical problem in queueing theory. There has been significant interest in the dynamic control of queues; e.g., see Stidham and Weber (1989) and references therein. In particular, Stidham and Weber (1989) study the optimal servicing policies for an $M/G/1$ queue of identical tasks. They formulate a semi-Markov decision process, and describe the qualitative features of the solution under certain technical assumptions. In the context of $M/M/1$ queues, Adusumilli and Hasenbein (2010) and George and Harrison (2001) relax some of technical assumptions in Stidham and Weber (1989). Hernández-Lerma and Marcus (1983) determine optimal servicing policies for queues with identical tasks and some unknown mean arrival rate. They adapt the optimal policy as the mean arrival rate is learned. In another related work, Zafer and Modiano (2008) study static queues with monomial and exponential utilities. They approximate the problem with a continuous time MDP. In the case of the dynamic queue, they propose a heuristic that solves the static problem at each stage.

In this paper, we study the problem of optimal duration allocation in a queue of binary decision making tasks with a human operator. We refer to such queues as *decision making queues*. In contrast to the aforementioned works in queues with human operators, we do not assume that the tasks require a fixed (potentially stochastic) processing time. We consider that each task may be processed for any amount of time, and the performance on the task is known as a function of the processing time. Moreover, we assume that tasks come with processing deadlines, and incorporate these deadlines as a soft constraint, namely, latency penalty (penalty due to delay in processing of a task). We consider two particular problems. First, we consider a static queue with latency penalty. Here, the human operator has to serve a given number of tasks. The operator incurs a penalty due to the delay in processing of each task. This penalty can be thought of as the loss in value of the task over time. Second, we consider a dynamic queue of decision making tasks. Tasks arrive according to a stochastic process and the operator incurs a penalty for the delay in processing each task. In both the problems, there is a trade-off between the reward obtained by processing a task and the penalty incurred due to the resulting delay in processing other tasks. We address this particular trade-off. The problem considered in this paper is similar to the problem considered in Adusumilli and Hasenbein (2010), George and Harrison (2001) and Stidham and Weber (1989). The main differences between these works and the problem considered in this paper are: (i) we consider a deterministic service process, and this yields an optimality equation significantly different from the optimality equation obtained for Markovian service process; (ii) we consider heterogeneous tasks, while the aforementioned works consider identical tasks. These works either propose approximate solution strategies customized to their setup, or rely on standard methods, e.g., the value iteration method in the case of a finite action space. In our problem, the heterogeneous nature of tasks significantly increases the dimension of the state space and makes the computation of optimal policies computationally intractable. We resolve this issue by utilizing the certainty-equivalent receding horizon framework (Bertsekas, 2005; Chang & Marcus, 2003; Mattingley, Wang, & Boyd, 2011) to approximately compute the solution.

The major contributions of this work are fourfold. First, we determine the optimal duration allocation policy for the static decision making queue with latency penalty. We show that the optimal policy may not process every task in the queue and may drop a few tasks, i.e., allocate zero duration to few tasks. Second, we pose a Markov decision process (MDP) to determine the optimal allocations for the dynamic decision making queue. We then establish some properties of this MDP. In particular, we show that an optimal policy exists and that it drops tasks if the queue length is greater than a critical value. Third, we employ the certainty-equivalent receding horizon optimization to approximately solve this MDP. We establish performance bounds on the certainty-equivalent receding horizon solution. Fourth and finally, we suggest guidelines for the design of decision making queues. These guidelines suggest the maximum mean arrival rate at which the operator expects a new task to arrive soon after optimally processing the current task.

The remainder of the paper is organized as follows. We present some preliminaries and the problem setup in Section 2. The static queue with latency penalty is considered in Section 3. We pose the optimization problem associated with the dynamic queue with latency penalty and study its properties in Section 4. We present and analyze a receding horizon algorithm for dynamic queue with latency penalty in Section 5. Our conclusions are presented in Section 6.

2. Preliminaries and problem setup

We consider the problem of attention allocation in queues with decision making tasks. We assume that the decision making tasks are independent of each other and arrive according to a Poisson process with a given mean rate. A human operator processes these tasks on a *first-come first-serve* (FCFS) basis (see Fig. 3.) The FCFS servicing discipline is a standard assumption in several queueing systems with a human operator (Koole & Mandelbaum, 2002; Savla & Frazzoli, 2010, 2012). The human operator receives a unit reward for the correct decision, while there is no penalty for a wrong decision. We assume that the tasks can be parametrized by some variable, which we will interpret here as the difficulty of the task, and the variable takes a value in a finite set $\mathcal{D} \subseteq \mathbb{R}$. Let the performance of the operator on a task with parameter $d \in \mathcal{D}$ be a function $f_d : \mathbb{R}_{\geq 0} \to [0, 1)$ of the duration the operator allocates to the task. A performance function relevant to the discussion in this paper is the probability of making the correct decision. The evolution of the probability of a correct decision by a human operator has been studied in cognitive psychology literature (Bogacz et al., 2006; Pew, 1969). We now briefly review some human decision making models:

Pew's model: For a two alternative forced choice task, the probability of the correct decision D_1 given that hypothesis H_1 is true and time t has been spent to make the decision is:

$$\mathbb{P}(D_1|H_1, t) = \frac{p_0}{1 + e^{-(at-b)}},$$

where $p_0 \in [0, 1]$, $a, b \in \mathbb{R}$ are some parameters specific to the human operator (Pew, 1969). The evolution of $\mathbb{P}(D_1|H_1, t)$ according to Pew's model is shown in Fig. 1(a).