CCIT Report #577 March 2006

Universal Filtering Via Prediction*

Tsachy Weissman† Erik Ordentlich‡ Marcelo J. Weinberger‡
Anelia Somekh-Baruch§ Neri Merhav¶

March 6, 2006

Abstract

We consider the filtering problem, where a finite-alphabet individual sequence is corrupted by a discrete memoryless channel, and the goal is to causally estimate each sequence component based on the past and present noisy observations. We establish a correspondence between the filtering problem and the problem of prediction of individual sequences which leads to the following result: Given an arbitrary finite set of filters, there exists a filter which performs, with high probability, essentially as well as the best in the set, regardless of the underlying noiseless individual sequence. We use this relationship between the problems to derive a filter guaranteed of attaining the "finite-state filterability" of any individual sequence by leveraging results from the prediction problem.

1 Introduction

The study of prediction of individual sequences with respect to a set of predictors (also known as experts) was pioneered by Hannan [18] and Blackwell [6, 5], who considered competition with the set of constant predictors. Their work prompted further research on and refinements of the problem throughout the late fifties, sixties, and seventies, with notable examples including [9, 10], and references therein. More recently, the problem has seen a resurgence of interest by both the information and learning theory communities, generalizing the original framework to accommodate competition with more general, and in fact arbitrary, predictors, cf. [32, 7, 8, 23] and references therein.

On a parallel thread, study of the problem of estimating the components of a noise-corrupted individual sequence was initiated by Robbins in the seminal [25] and dubbed 'the compound decision problem'. The problem has been the focus of much attention during the fifties and sixties, notable references including [19, 28, 26, 27] (cf. [39] for a comprehensive account of this literature). Much in

*Elements in this paper were presented at the 37th Annual Allerton Conference on Communication, Control, and Computing, September 1999, and at the 2004 Data Compression Conference (DCC'04), Snowbird, Utah, March 2004.

†Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA. Part of this work was done while the author was visiting Hewlett-Packard Laboratories.

‡Hewlett-Packard Laboratories, Palo Alto, CA 94304, USA

§Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA.

¶Department of Electrical Engineering, Technion, Haifa 32000, Israel.