Appl Intell (2010) 33: 3–20
DOI 10.1007/s10489-010-0228-1

Optimal sampling for estimation with constrained resources using a learning automaton-based solution for the nonlinear fractional knapsack problem

Ole-Christoffer Granmo · B. John Oommen

Published online: 24 April 2010

Abstract While training and estimation for Pattern Recognition (PR) have been extensively studied, the question of achieving these when the resources are both limited and constrained is relatively open. This is the focus of this paper. We consider the problem of allocating limited sampling resources in a "real-time" manner, with the explicit purpose of estimating multiple binomial proportions (the extension of these results to non-binomial proportions is, in our opinion, rather straightforward). More specifically, the user is presented with 'n' *training* sets of data points, $S_1, S_2, \ldots, S_n$, where the set S_i has N_i points drawn from two classes $\{\omega_1, \omega_2\}$. A random sample in set S_i belongs to ω_1 with probability u_i and to ω_2 with probability $1 - u_i$, with $\{u_i\}$, $i = 1, 2, \ldots n$, being the quantities to be learnt. The problem is both interesting and non-trivial because while both n and each N_i are large, the number of samples that can be drawn is bounded by a constant, c. A web-related problem which is based on this model (Snaprud et al., The Accessibility for All Conference, 2003) is intriguing because the sampling resources can only be allocated optimally if the binomial proportions are *already* known. Further, no non-automaton solution has ever been reported if these proportions are unknown and must be sampled.

Using the general LA philosophy as a paradigm to tackle this real-life problem, our scheme improves a current solution in an online manner, through a series of informed guesses which move towards the optimal solution. We solve the problem by first modelling it as a *Stochastic Nonlinear Fractional Knapsack Problem*. We then present a completely new on-line Learning Automata (LA) system, namely, the *Hierarchy of Twofold Resource Allocation Automata* (H-TRAA), whose primitive component is a *Twofold Resource Allocation Automaton* (TRAA), both of which are asymptotically optimal. Furthermore, we demonstrate empirically that the H-TRAA provides *orders of magnitude* faster convergence compared to the Learning Automata Knapsack Game (LAKG) which represents the state-of-the-art. Finally, in contrast to the LAKG, the H-TRAA scales sub-linearly. Based on these results, we believe that the H-TRAA has also tremendous potential to handle demanding real-world applications, particularly those dealing with the world wide web.

Keywords Sample size determination · Training with constrained resources · Constrained estimation for pattern recognition · Nonlinear knapsack problems · Hierarchical learning · Learning automata · Stochastic optimization

1 Introduction

In this paper we consider the fascinating problem of learning distributions when the number of "Classes" is large, and the

The second author was partially supported by NSERC, the Natural Sciences and Engineering Research Council of Canada. A preliminary version of this paper was presented at IEA/AIE'09, the 2009 International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, Taiwan, in June 2009. This paper won the *Best Paper Award* of the conference.

O.-C. Granmo (✉)
Department of Information and Communication Technology, University of Agder, Grooseveien 36, 4876 Grimstad, Norway
e-mail: ole.granmo@uia.no

B.J. Oommen
School of Computer Science, Carleton University, Ottawa, Canada, K1S 5B6
e-mail: oommen@scs.carleton.ca

B.J. Oommen
University of Agder, Grimstad, Norway

number of elements (i.e., data points) per class is also large, but when the resources available for obtaining data points are limited. In particular, we concentrate on methods which are forced to resort to sampling techniques because of the cardinalities of the set of classes and of the set of data points per class.

The phenomenon of estimation, although central to almost all aspects of science, engineering, research and the analysis of data, universally assumes that the observation of the events whose properties are to be estimated is feasible. Thus, for example, if we intend to learn the parameters of a Bernoulli distribution, we assume that we have access to the "Number of Heads and Tails" that have occurred in any experiment. At one extreme, such data points are abundant, making it possible to obtain reliable estimates for all involved parameters. However, in many cases (for example, in the PR of medical conditions), data points are not abundantly available. At the other extreme of the spectrum, we encounter situations when the underlying events associated with the estimation process may even be unobservable. In an earlier paper [32], which to the best of our knowledge was pioneering in this vein, we discussed how estimation could be achieved and how the corresponding actions (decisions) could be taken in such an "unobservable" setting. In less severe, and perhaps more typical situations, however, while observation of data points *is* feasible, the resources available for performing the observations may be limited. As a result, the data points must then be carefully obtained within the available resource bounds, in order to ensure as reliable estimates as possible. This is the primary focus of this paper—which, to the best of our knowledge, is a relatively little-studied problem.

We briefly explain the spectrum of problems alluded to above. The theory of estimation has been studied for hundreds of years [2, 5, 14, 22, 23, 25]. Besides, the learning (training) phase of a statistical pattern recognition system is, indeed, based on estimation theory [7, 9, 13, 29]. Estimation methods generally fall into various categories, including the Maximum Likelihood Estimates (MLE) and the Bayesian family of estimates [2, 5, 7, 9] which are well-known for having good computational and statistical properties. Consider the strategy used for developing the MLE of the parameter of a distribution, $f_X(\theta)$, whose parameter to be estimated is θ. The input to the estimation process is the set of points $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$, which are assumed to be generated independently and identically as per the distribution, $f_X(\theta)$. The process involves deriving the likelihood function, i.e., the likelihood of the distribution, $f_X(\theta)$, generating the sample points $\mathcal{X}$ given θ, which is then maximized (by traditional optimization or calculus methods) to yield the estimate, $\hat{\theta}$. The general characteristic sought for is that the estimate $\widehat{\theta}_{MLE}$ converges to the true (unknown) θ with probability one, or in a mean square sense. Bayesian and MLE estimates generally possess this desirable phenomenon. Typically, at one end of the spectrum, the assumption is that the data points are available so that the practitioner can obtain "meaningful" estimates.

We emphasize though that, *the whole process is meaningless* if we are not provided the sequence $\mathcal{X}$. Requiring the estimation of the parameters of the distribution generating $\mathcal{X}$ without being provided with any information[1] about $\mathcal{X}$ would be tantamount to "spinning gold out of straw". This constitutes the problem at the other end of the spectrum, namely, where one has to do training, learning and decision making when the events themselves are unobservable, which has been tackled in [32].

A problem in the middle of this spectrum is the one encountered when the points can be made available, but the resources to accomplish the estimation/training are both constrained and limited. While this may sound hypothetical, it truly is a relevant problem when one has to construct a good training set when the observations themselves are expensive, as in the medical domain. This situation also presents itself in cases when the domain is large, as in the problems (explained later) that deal with the world-wide web. In this paper we shall consider how we can extend learning (i.e., estimation and training) principles to the case when we are permitted to include "Constrained Sampling" in the process. However, to render the problem both interesting and non-trivial, we concentrate on the scenario when the number of classes and the size of the points per class is "large", while the number of samples that can be drawn is bounded by a constant. The intricacy of the problem becomes even more obvious if we note that a knowledge of the distributions for each class automatically leads to the optimal allocation of sampling resources—but these are the very quantities which are to be estimated. Hence the paradox!

The problem studied in this paper can be stated as follows: We consider the case when the user is presented with 'n' sets of data points, $S_1, S_2, \ldots, S_n$ (see Fig. 1). The set S_i has N_i points drawn from two classes $\{\omega_1, \omega_2\}$. A random sample in set S_i belongs to ω_1 with probability u_i and it belongs to ω_2 with probability $1 - u_i$. The problem we study involves estimating $\{u_i\}$ for $i = 1, 2, \ldots, n$. However, to render the problem both meaningful and non-trivial, we assume that we are allowed to sample *a maximum of only* c points from the collective group of sets, $\{S_i\}$, and thus we have to determine how samples must be drawn (with replacement) from each set so that we can obtain both accurate and efficient estimates of $\{u_i\}$. The purpose is to make the estimates as accurate as possible, and so, we will here

[1] It is possible to get very good estimates of θ if one is provided with random occurrences of a *known* function of X. Thus, instead of receiving $\{X_i\}$, if we are provided with $\{Y_i\}$, where, for example, if each $Y_i = X_i^2$, an MLE can be easily devised to estimate θ by observing the Y_i's.

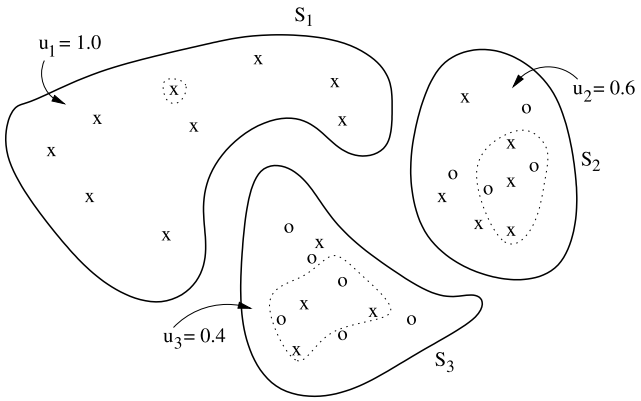

Fig. 1 The figure shows three sets S_1, S_2 and S_3 containing elements of two classes 'x' and 'o'. Since S_1 is homogenous and contains elements of only one class ('x'), a *single* sample is sufficient to determine this. Sets S_2 and S_3 are non-homogenous, although the proportions of the classes in each is different. The problem involves determining how the sampling is to be done when the proportions are unknown

pursue the goal of minimizing the variance of the proportion estimates when the total number of samples available for estimating the proportions is bounded by c. While brute force approaches to solve this problem can be easily devised, our aim is to arrive at a solution without exhaustively searching the search space, and by mapping it to a family of well-known problems, namely the family of knapsack problems.

1.1 Applications of these results

The world wide web is an extremely vast resource-thirsty field, which probably consumes a major portion of the computing resources available today. Searching, updating and examining web-pages is, undoubtedly, one of the primary tasks done by both individuals and companies today. This, in turn, leads to numerous extremely interesting real-life resource allocation and scheduling problems, and the problem that we emphasize in this paper is closely related to the so-called "Web Monitoring" problem.

Web page monitoring consists of repeatedly polling a selection of web pages (sites) so that the user can infer the contents of the documents at the respective sites. Clearly, as this task can be prohibitively expensive, in practical applications, the system imposes a constraint on the *maximum* number of documents at a web page that can be polled per unit of time. This bound is dictated by the governing communication bandwidth, and by the speed limitations associated with the processing. Since only a fraction of the web pages can be polled within a given unit of time, the problem which the system's analyst encounters is one of determining which web pages are to be examined, and the subset of documents to be examined at the site (once it is being examined). For example, consider the problem of investigating the number of documents in a site which are written in *English* or *Norwegian*. Clearly, it is not feasible to examine every single document in every single site. However, if one resorts to sampling, the question is one of knowing how documents from any site should be drawn. If any set S_i is homogenous (i.e., all the documents are in a given language), it is sufficient to examine only a single document in that class. The other extreme is the trivial case, when all the proportions are known. In such a case, the optimal sampling proportion is related to the latter proportions—which is the very quantity to be estimated. We refer the reader to Fig. 1 which concerns three classes. Clearly the number of samples to be drawn from each set depends on the homogeneity of the respective classes—which is what the learning problem entails.

The problem we study is also exactly the one encountered when, for instance, the task at hand is to determine the proportion of a web site that is successfully validated by an HTML validator [24] and/or a WCAG accessibility validator [31], and that n web sites are to be evaluated by only accessing c web pages.

The analogous parallel for the medical domain would be that of sampling individuals from different categories (for example, smokers *versus* non-smokers), where individuals of each category are found in different training sets. The issue here is that the testing of the individuals for a specific condition (say lung cancer) can be prohibitively expensive, and so the system is constrained to examine at most c individuals.

Although this problem is quite general, we shall proceed to solve it by suggesting that it falls within the model of knapsack-based problems.

1.2 Formal formulation as a knapsack problem

We first formulate, in a fairly general setting, a set of knapsack-based problems that are, in actuality, related to the web-polling and web monitoring problems. Indeed, we address one such model which can be translated into a family of problems:

> *Imagine that you have to allocate a limited amount of time among n different activities. The problem is such that spending a time instant on an activity randomly produces one of two possible outcomes—the time instant is either spent "fruitfully" or "unfruitfully". In this generic setting, your goal is to maximize the expected amount of fruitfully spent time. Unfortunately, you are only given the following information regarding the activities:*
>
> 1. *Each instant of time spent on an activity has a certain probability of being fruitful, and*

2. *This probability decreases with the amount of time spent on the activity.*

To render the problem even more realistic, you do not have access to the probabilities themselves. Instead, you must rely on solving the problem by means of trial-and-error, i.e., by attempting different allocations, and observing the resulting random outcomes.

Instances of the above problem can be formulated as instantiations of the *Stochastic Non-linear Fractional Equality Knapsack* (*NEFK*) *Problem* as clarified earlier [10–12], and as further explained below.

1.3 Related work and state-of-the-art

In order to appreciate the qualities of the Stochastic NEFK Problem, it is beneficial to view the problem in the light of the classical *linear* Fractional Knapsack (FK) Problem, which is what we shall do now.

1.3.1 Related foundational problems

The linear fractional knapsack (FK) problem The linear FK problem is a classical continuous optimization problem which also has applications within the field of resource allocation. The problem involves n materials of different value v_i per unit volume, $1 \leq i \leq n$, where each material is available in a certain amount $x_i \leq b_i$. Let $f_i(x_i)$ denote the value of the amount x_i of material i, i.e., $f_i(x_i) = v_i x_i$. The problem is to fill a knapsack of fixed volume c with the material mix $\vec{x} = [x_1, \ldots, x_n]$ to yield a maximal value $\sum_1^n f_i(x_i)$ [3].

The nonlinear equality FK (NEFK) problem One important extension of the above classical problem is the *Nonlinear Equality* FK problem with a separable and concave objective function. The problem can be stated as follows [15]:

$$\text{maximize} \quad f(\vec{x}) = \sum_1^n f_i(x_i)$$

$$\text{subject to} \quad \sum_1^n x_i = c \quad \text{and} \quad \forall i \in \{1, \ldots, n\}, \quad x_i \geq 0.$$

Since the objective function is considered to be concave, the value function $f_i(x_i)$ of each material is also concave. This means that the derivatives of the material value functions $f_i(x_i)$ with respect to x_i (hereafter denoted f_i'), are non-increasing. In other words, the material value *per unit volume* is no longer constant as in the linear case, but decreases with the material amount, and so the optimization problem becomes:

$$\text{maximize} \quad f(\vec{x}) = \sum_1^n f_i(x_i),$$

$$\text{where } f_i(x_i) = \int_0^{x_i} f_i'(x_i)dx_i$$

$$\text{subject to} \quad \sum_1^n x_i = c \quad \text{and} \quad \forall i \in \{1, \ldots, n\}, \quad x_i \geq 0.$$

Efficient solutions to the latter problem, based on the principle of Lagrange multipliers, have been devised. In short, the optimal value occurs when the derivatives f_i' of the material value functions are equal, subject to the knapsack constraints [4, 8]:

$$f_1'(x_1) = \cdots = f_n'(x_n),$$

$$\sum_1^n x_i = c \quad \text{and} \quad \forall i \in \{1, \ldots, n\}, \quad x_i \geq 0.$$

The stochastic NEFK problem In this paper we generalize the above nonlinear equality knapsack problem. First of all, we let the material value per unit volume for any x_i be a *probability* function $p_i(x_i)$. Furthermore, we consider the distribution of $p_i(x_i)$ to be *unknown*. That is, each time an amount x_i of material i is placed in the knapsack, we are only allowed to observe an instantiation of $p_i(x_i)$ at x_i, and not $p_i(x_i)$ itself. Given this stochastic environment, we intend to devise an on-line incremental scheme that learns the mix of materials of maximal *expected* value, through a series of informed guesses. Thus, to clarify issues, we are provided with a knapsack of fixed volume c, which is to be filled with a mix of n different materials. However, unlike the NEFK, in the Stochastic NEFK Problem the unit volume value of a material i, $1 \leq i \leq n$, is a random quantity—it takes the value 1 with probability $p_i(x_i)$ and the value 0 with probability $1 - p_i(x_i)$, respectively. As an additional complication, $p_i(x_i)$ is nonlinear in the sense that it decreases monotonically with x_i, i.e., $x_{i_1} \leq x_{i_2} \Leftrightarrow p_i(x_{i_1}) \geq p_i(x_{i_2})$.

Since unit volume values are random, we operate with the expected unit volume values rather than the actual unit volume values themselves. With this understanding, and the above perspective in mind, the expected value of the amount x_i of material i, $1 \leq i \leq n$, becomes $f_i(x_i) = \int_0^{x_i} p_i(u)du$. Accordingly, the expected value per unit volume[2] of material i becomes $f_i'(x_i) = p_i(x_i)$. In this stochastic and nonlinear version of the FK problem, the goal is to fill the knapsack so that the expected value $f(\vec{x}) = \sum_1^n f_i(x_i)$ of the

[2] We hereafter use $f_i'(x_i)$ to denote the derivative of the expected value function $f_i(x_i)$ with respect to x_i.

material mix contained in the knapsack is maximized as below

$$\text{maximize} \quad f(\vec{x}) = \sum_{1}^{n} f_i(x_i),$$

$$\text{where } f_i(x_i) = \int_0^{x_i} p_i(u)du, \text{ and}$$

$$p_i(x_i) = f_i'(x_i)$$

$$\text{subject to} \quad \sum_{1}^{n} x_i = c \quad \text{and} \quad \forall i \in \{1, \ldots, n\}, \quad x_i \geq 0.$$

A fascinating property of the above problem is that the amount of information available to the decision maker is limited—the decision maker is only allowed to observe the current unit value of each material (either 0 or 1). That is, each time a material mix is placed in the knapsack, the unit value of each material is provided to the decision maker. The actual outcome probabilities $p_i(x_i)$, $1 \leq i \leq n$, however, remain *unknown*. As a result of the latter, the expected value of the material mix must be maximized by trial-and-error, i.e., by experimenting with different material mixes and observing the resulting outcomes.

1.3.2 State-of-the-art

To the best of our knowledge, prior to our work reported in [12], the stochastic NEFK problem, which we address, was not covered in the literature before. However, several studies on related problems have been reported. For example, the works of [6, 26] consider solution policies for stochastic generalizations of the so-called NP-hard *linear* integer knapsack problem. In these papers, value distributions were considered known and constant, making dynamic programming a viable solution. Another variant of the knapsack problem is found in [21] where a deterministic knapsack is used, however, with objects arriving to and departing from the knapsack at random times. The optimization problem considered was to accept/block arriving objects so that the average value of the knapsack is maximized.

The first reported generic treatment of the stochastic NEFK problem itself can be found in [12]. Various instantiations of the problem have, however, appeared sporadically, particularly within the web monitoring domain. In these latter instantiations, the unknown parameters are *estimated* by means of a tracking phase where web pages are polled mainly for estimation purposes [20, 30]. One major disadvantage of such an approach is that the parameter estimation phase significantly delays the implementation of an optimal solution. This disadvantage is further aggravated in *dynamic* environments where the optimal solution changes over time, introducing the need for parameter re-estimation [10].

In contrast to the above approaches, we base our work on the principles of Learning Automata (LA) [16, 27]. LA have been used to model biological systems [28], and have attracted considerable interest in the last decade because they can learn the optimal actions when operating in (or interacting with) unknown stochastic environments. Furthermore, they combine rapid and accurate convergence with low computational complexity. The novel Learning Automata Knapsack Game (LAKG) scheme that we proposed in [12] does not rely on estimating parameters, and can be used to solve the stochastic NEFK problem in both static and dynamic settings. Accordingly, we believe that the LAKG can be considered to represent the state-of-the-art when it concerns research on the stochastic NEFK problem. This landmark is now extended to develop the TRAA (which, in itself is the first reported LA which is *artificially* ergodic),[3] and its hierarchical version, the H-TRAA.

With regard to the particular application domain, recent approaches to resource allocation in web monitoring attempt to *optimize* the performance of the system when the monitoring capacity is restricted [20, 30]. The principle cited in the literature essentially invokes Lagrange multipliers to solve a *nonlinear equality* knapsack problem with a separable and concave objective function [15]. Thus, for example, a basic web monitoring resource allocation problem may involve n web pages that are updated periodically, although with different periods. Clearly, each web page can be sampled maximally—which would result in a sluggish system. Thus, it would be fair to state that the problem which we study has a direct potential application in the domain of optimal web page sampling. Observe that this must be achieved without exceeding the available monitoring capacity—e.g., the maximum number of pages that can be accessed per unit of time as dictated by the governing bandwidth and processing speed limitations.

1.4 Contributions of this paper

The contributions of this paper are the following:

1. We report the first *analytical* results for schemes that solve the optimal sampling problem using a formal asymptotically optimal solution to the Stochastic NEFK Problem. The implications of the solution for estimation and training in PR are both direct and clearly validated.
2. We propose a novel scheme for the *two-set* sampling problem, namely, the *Twofold Resource Allocation Automaton* (*TRAA*). As mentioned, from the perspective of LA, the TRAA, in itself, is the first reported LA which is *artificially* rendered ergodic.

[3] LA which have been artificially made *absorbing* to yield specific properties, have been earlier reported [17]. However, we are not aware of any LA which, in essence are absorbing, but which have been made artificially *ergodic*.

3. We prove that the TRAA is asymptotically optimal.
4. We report the first *hierarchical* solution to the Stochastic NEFK Problem, based on a hierarchy of TRAAs, namely, the H-TRAA.
5. We verify empirically that the H-TRAA provides orders of magnitude faster convergence when compared to the LAKG, and demonstrate its power to solve the underlying sampling problem.

As a result of the above contributions, we believe that the H-TRAA is the first reported viable and realistic strategy for solving the optimal sampling problem. Its applications in real PR problems is straightforward. Further, it can also be used to resolve other optimal resource allocation problems in large-scale web accessibility assessment [24].

1.5 Paper organization

The paper is organized as follows. In Sect. 2 we present the Twofold Resource Allocation Automaton (TRAA) for the *two-material* problems, and prove its asymptotic optimality. We then propose how TRAAs can be arranged in a hierarchy for solving *multi-material* Stochastic NEFK Problems. We proceed to empirically verify that the H-TRAA provides orders of magnitude faster convergence compared to the LAKG in Sect. 3 when applied to the optimal sampling problem. Indeed, we shall present results that clearly demonstrate that the H-TRAA allows us to tackle problems involving up to 16,384-sets of Bernoulli parameters in *real-time*, which we believe is no small achievement. Finally, we offer suggestions for further work before we conclude the paper in Sect. 4.

2 A hierarchy of twofold resource allocation automata (H-TRAA)

2.1 Overview of the non-LA solutions

In order to put our work in the right perspective, we start this section by providing a brief review of the concepts and the solution found in [12]—which are also relevant for more "primitive" variants of the knapsack problem.

As indicated in the introduction, solving the classical linear FK problem involves finding the most valuable mix $\vec{x}^* = [x_1^*, \ldots, x_n^*]$ of n materials that fits within a knapsack of fixed capacity c. The material value per unit volume for each material i is given as a constant v_i, and each material is available in a certain amount $x_i \leq b_i$, $1 \leq i \leq n$. Accordingly, the value of the amount x_i of material i, $f_i(x_i) = v_i x_i$, is linear with respect to x_i. In other words, the derivative of $f_i(x_i)$—i.e., the material value per unit volume—is fixed: $f_i'(x_i) = v_i$. Because a fraction of each material can be placed in the knapsack, the following greedy algorithm from [3] finds the most valuable mix: *Take as much as possible of the material that is most valuable per unit volume. If there is still room, take as much as possible of the next most valuable material. Continue until the knapsack is full.*

Let us now generalize this and assume that the material unit volume values are *random* variables with *constant* and *known* distributions. Furthermore, for the sake of conceptual clarity, let us only consider binary variables that *either* instantiate to the values of 0 or 1. Since the unit volume values are random, let p_i denote the probability of the unit volume value $v_i = 1$ for material i, $1 \leq i \leq n$, which means that the probability of the unit volume value $v_i = 0$ becomes $1 - p_i$. With some insight, it becomes evident that under such conditions, the above greedy strategy can again be used to maximize the *expected* value of the knapsack, simply by selecting material based on the *expected* unit volume values, $E[v_i] = 0 \times (1 - p_i) + 1 \times p_i$, rather than actual unit volume values.

The above indicated solution is, of course, inadequate when the p_i's are unknown. Furthermore, the problem becomes even more challenging when the p_i's are no longer constant, but rather depend on their respective material amounts x_i, $1 \leq i \leq n$. Let $p_i(x_i)$ denote the probability that the current unit volume value of material i is $v_i = 1$, given that the amount x_i has already been placed in the knapsack. Then, the expected value per unit volume of material i, $1 \leq i \leq n$, becomes $E[v_i] = 0 \times [1 - p_i(x_i)] + 1 \times p_i(x_i) = p_i(x_i)$, and accordingly, the expected value of the amount x_i becomes $f_i(x_i) = \int_0^{x_i} p_i(u)du$.

Our aim, then, is to find a scheme that moves towards optimizing the following NEFK problem on-line:

$$\begin{aligned} \text{maximize} \quad & f(\vec{x}) = \sum_{1}^{n} f_i(x_i), \\ & \text{where } f_i(x_i) = \int_0^{x_i} p_i(u)du, \text{ and} \\ & p_i(x_i) = f_i'(x_i) \\ \text{subject to} \quad & \sum_{1}^{n} x_i = c \quad \text{and} \quad \forall i \in \{1, \ldots, n\}, \quad x_i \geq 0. \end{aligned}$$

Note that we allow only instantiations of the material values per unit volume to be observed. That is, each time an amount x_i of material i is placed in the knapsack, an instantiation v_i at x_i is observed.

Because of the above intricacies, we approach the problem by relying on informed material mix *guesses*, i.e., by experimenting with different material mixes and learning from the resulting random unit volume value outcomes. We shall assume that x_i is any number in the interval (0, 1). The question of generalizing this will be considered later. The crucial issue that we have to address, then, is that of determining

how to change our current guesses on x_i, $1 \leq i \leq n$. We shall attempt to do this in a discretized manner by subdividing the unit interval into N points $\{\frac{1}{N+1}, \frac{2}{N+1}, \ldots, \frac{N}{N+1}\}$, where N is the resolution of the learning scheme. We will see that a larger value of N will ultimately imply a more accurate solution to the knapsack problem.

At this juncture, it is pertinent to mention that although the rationale for this updating is the stochastic point location solution proposed by Oommen in [18], the two schemes are quite distinct for the following reasons:

1. The method proposed in [18] assumes the existence of an Oracle which informs the LA whether to go "right" or "left". In our application domain, this now has to be *inferred* by the system.
2. The method proposed in [18] assumes that there is only a single LA in the picture. Here, we specifically understand that there are multiple LAs organized in a hierarchy—each of them being constrained to work together with the others.[4]
3. In [18] the problem of analyzing scenarios with space varying responses from the environment was left open. This problem is tackled in the present paper.
4. As opposed to the scheme in [18], our present approach is also applicable to dynamic (time varying) environments.
5. There is a "huge" fundamental difference between the LA which we devise here and the work of [18]. Unlike the latter, in which the system is truly ergodic, our present LA would be *absorbing* if the end-states of the probability space are also included. However, to forcefully render this present machine ergodic, we have artificially made the LA ergodic by *excluding* these states from the set of possible probability values. This makes the analysis both distinct and quite fascinating. As mentioned earlier, we are not aware of any LA which, in essence are absorbing, but which have been made artificially *ergodic*.

2.2 Details of the TRAA solution

2.2.1 *Design of the TRAA solution*

We first present our LA based solution to *two-material* Stochastic NEFK Problems. The two-material solution forms a critical part of the hierarchic scheme for multiple materials that is presented subsequently. As illustrated in Fig. 2, our solution to two-material problems constitutes of three modules:

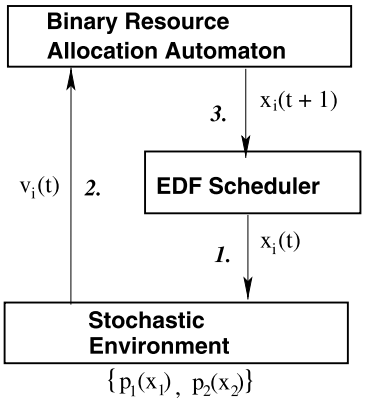

Fig. 2 The Twofold Resource Allocation Automaton (TRAA) interacting with a Scheduler and an unknown Stochastic Environment

1. A Stochastic Environment
2. The TRAA itself, and
3. An Earliest Deadline First (EDF) Scheduler.

We first detail each of the three modules, before we analyze the overall *feedback connection* between them. Finally, we prove that the TRAA that we have developed in this section is asymptotically optimal for two-material Stochastic NEFK Problems.

Stochastic environment The *Stochastic Environment* for the two-material case can be characterized by:

1. The capacity c of the knapsack;
2. Two material unit volume value probability functions $p_1(x_1)$ and $p_2(x_2)$.

In brief, if the amount x_i of material i is suggested to the Stochastic Environment, the Environment replies with a unit volume value $v_i = 1$ with probability $p_i(x_i)$ and a unit volume value $v_i = 0$ with probability $1 - p_i(x_i)$, $i \in \{1, 2\}$. It should be emphasized that to render the problem both interesting and non-trivial, we assume that $p_i(x_i)$ is unknown to the TRAA.

Twofold resource allocation automaton (TRAA) The scheme which attempts to learn the optimal allocation $\vec{x}^* = [x_1^*, x_2^*]$ can be described as follows. A finite fixed structure automaton with the states $s(t) \in \{1, 2, \ldots, N\}$ is used to decide the allocation of resources among the two materials. Let the current state of the automaton be $s(t)$. Furthermore, let $q_{s(t)}$ refer to the fraction $\frac{s(t)}{N+1}$, and let $r_{s(t)}$ refer to the fraction: $1 - q_{s(t)}$. Then the automaton's current guess is $\vec{x} = [q_{s(t)}, r_{s(t)}]$.

If the Stochastic Environment tells the automaton that the unit volume value of material i is $v_i(t)$ at time t, the automaton updates its state as follows:

[4] It is conceivable that this problem can be resolved with a single LA possessing an extended number of actions. But we do not recommend it for scalability reasons—the action space would grow exponentially.

$$s(t+1) := s(t)+1 \quad \textbf{If rand}() \le r_{s(t)} \textbf{ and } v_i(t) = 1 \textbf{ and } 1 \le s_i(t) < N \textbf{ and } i = 1, \tag{1}$$

$$s(t+1) := s(t)-1 \quad \textbf{If rand}() \le q_{s(t)} \textbf{ and } v_i(t) = 1 \textbf{ and } 1 < s_i(t) \le N \textbf{ and } i = 2, \tag{2}$$

$$s(t+1) := s(t) \quad \textbf{Otherwise}. \tag{3}$$

Figure 3 shows the resulting stochastic transition graphs for resolution $N = 5$. The upper graph shows the transitions for feedback from the Stochastic Environment on material 1, and the graph below shows the transitions for feedback on material 2. Notice how the stochastic state transitions are designed to offset the learning bias introduced by accessing the materials with frequencies proportional to $\vec{x} = [q_{s(t)}, r_{s(t)}]$. Also observe that the overall learning scheme does not produce any absorbing states, and is, accordingly, ergodic supporting dynamic environments. The effect of these properties is analysed in the next subsection.

Finally, after the automaton has had the opportunity to change its state, it provides output to the EDF Scheduler. That is, it outputs the material amounts $\vec{x} = [q_{s(t+1)}, r_{s(t+1)}]$ that have been changed.

Earliest deadline first (EDF) scheduler The Scheduler takes material amounts $\vec{x} = [x_1, \ldots, x_n]$ as its input (for the two-material case the input is $\vec{x} = [x_1, x_2]$). The purpose of the Scheduler is:

1. To provide accesses to the Stochastic Environment in a sequential manner, and
2. To make sure that the unit volume value functions are accessed with frequencies proportional to $\vec{x}$.

The reader should note that our scheme does not rely on accessing the unit volume value functions sequentially with frequencies proportional to $\vec{x}$ for solving the knapsack problem. However, this restriction is obviously essential for solving the problem *incrementally* and *on-line* (or rather in a "real-time" manner). Note that since it, in some cases, may be essential to access each unit volume value function with a constant period and not randomly (for example, in the earlier-alluded-to problem which analyzes web page polling), we use the Earliest Deadline First (EDF) Scheduling to access the functions according to $\vec{x}$.

2.2.2 Analysis of the TRAA solution

In this section we characterize the optimal solution to a Stochastic NEFK Problem. Thereafter, we analyze the feedback connection of the TRAA and the Stochastic Environment—we prove that the TRAA is asymptotically optimal in the sense that it can find material allocations arbitrarily close to the solution of the Stochastic NEFK Problem.

Lemma 1 *The material mix $\vec{x} = [x_1, \ldots, x_n]$ is a solution to a given Stochastic NEFK Problem if* (1) *the derivatives of the expected material amount values are all equal at $\vec{x}$,* (2) *the mix fills the knapsack, and* (3) *every material amount is positive, i.e.*:

$$f_1'(x_1) = \cdots = f_n'(x_n),$$

$$\sum_{1}^{n} x_i = c \quad \textit{and} \quad \forall i \in \{1, \ldots, n\}, \quad x_i \ge 0.$$

The above lemma is based on the well-known principle of Lagrange Multipliers [4, 8], and its proof is therefore omitted here for the sake of brevity. Instead, we will start by analyzing the *two-material* problem and the TRAA. Multiple TRAAs will then be organized in a hierarchy with the aim of tackling *n-material* problems.

For the two-material problem, let $\vec{x}^* = [x_1^*, x_2^*]$ denote a solution, as defined above. Note that since x_2^* can be obtained from x_1^*, we will concentrate on finding x_1^*.

Theorem 1 *The TRAA solution scheme specified by* (1)–(3) *is asymptotically optimal.*

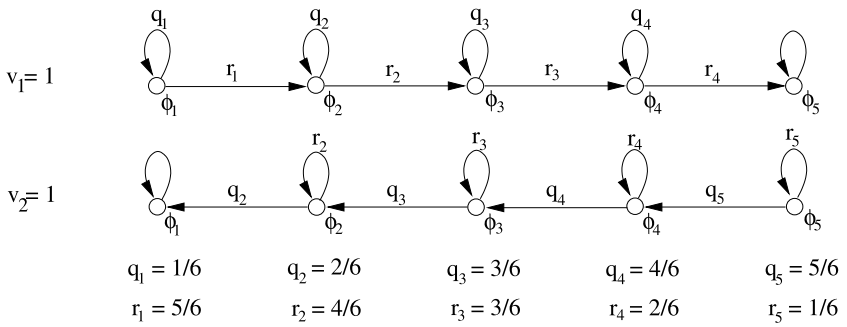

Fig. 3 The stochastic transition graphs of a TRAA with resolution $N = 5$

Proof Our aim is to prove that as the resolution, N, is increased indefinitely, the expected value of the TRAA output, $x_1(t)$, converges towards the solution of the problem, x_1^*, implying that:

$$\lim_{N\to\infty}\lim_{t\to\infty} E[x_1(t)] \to x_1^*.$$

We shall prove the above by analyzing the properties of the underlying Markov chain, which is specified by the EDF Scheduler, the rules (1)–(3) (the TRAA), and the Environment. As can be seen from (1)–(3), the states of the chain are the integers $j \in \{1, 2, \ldots, N\}$. In brief, rules (1)–(3), when interacting with the EDF Scheduler and the Environment, obey the Markov chain with transition matrix $H = [h_{ij}]$, where

$$h_{j,j-1} = r_j \cdot p_2(r_j) \cdot q_j, \quad 1 < j \le N, \tag{4}$$

$$h_{j,j+1} = q_j \cdot p_1(q_j) \cdot r_j, \quad 1 \le j < N, \tag{5}$$

$$h_{j,j} = 1 - h_{j,j-1} - h_{j,j+1}, \quad 1 < j < N, \tag{6}$$

and, accordingly,

$$h_{1,1} = 1 - h_{1,2}, \tag{7}$$

$$h_{N,N} = 1 - h_{N,N-1}. \tag{8}$$

Clearly, H represents a single closed communicating class whose periodicity is unity. The chain is ergodic, and the limiting probability vector is given by the eigenvector of H^T corresponding to eigenvalue unity. Let this vector be $\Pi = [\pi_1, \pi_2, \ldots, \pi_N]$. Then, Π satisfies:

$$\begin{bmatrix} h_{1,1} & h_{1,2} & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ h_{2,1} & h_{2,2} & h_{2,3} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & h_{3,2} & h_{3,3} & h_{3,4} & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & h_{N-2,N-3} & h_{N-2,N-2} & h_{N-2,N-1} & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & h_{N-1,N-2} & h_{N-1,N-1} & h_{N-1,N} \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & h_{N,N-1} & h_{N,N} \end{bmatrix}^T \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \cdot \\ \cdot \\ \pi_{N-2} \\ \pi_{N-1} \\ \pi_N \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \cdot \\ \cdot \\ \pi_{N-2} \\ \pi_{N-1} \\ \pi_N \end{bmatrix}. \tag{9}$$

The details of solving (9) are quite cumbersome, and we undertake it now. Observe that our aim is to prove that the probability mass of Π lies arbitrarily close to the solution of the knapsack problem, $\vec{x}^* = [x_1^*, x_2^*]$, as N goes to infinity. Before we go through the fine details, we outline the proof strategy as follows. We first explicitly solve for the quantities $\{\pi_i\}$ by solving the underlying difference equations. We then define a function U that forms an upper bound for Π. We proceed to show that the upper bound goes to zero outside an arbitrarily close vicinity of x_1^*, as the resolution, N, goes to infinity. Accordingly, since Π is a probability distribution, and since U is its upper bound, increasing the resolution towards infinity, moves the probability mass of Π arbitrarily close to x_1^*.

The details of the proof follow. Our first step is to reformulate the individual row-wise equations from the matrix (9) recursively. Expanding the first row of (9) yields:

$$\pi_1 \cdot h_{1,1} + \pi_2 \cdot h_{2,1} = \pi_1$$
$$\implies \quad \pi_2 = \frac{(1 - h_{1,1}) \cdot \pi_1}{h_{2,1}} = \frac{h_{1,2}}{h_{2,1}} \cdot \pi_1. \tag{10}$$

Expanding the second row of (9) and substituting (10) yields:

$$\pi_1 \cdot h_{1,2} + \pi_2 \cdot h_{2,2} + \pi_3 \cdot h_{3,2} = \pi_2$$
$$\implies \quad \pi_3 = \frac{h_{2,3}}{h_{3,2}} \cdot \pi_2. \tag{11}$$

Arguing in a similar way in a row-by-row manner, it can be seen[5] that

$$\pi_{k-1} = \frac{h_{k,k-1}}{h_{k-1,k}} \cdot \pi_k \tag{12}$$

for $0 < k \le N$, which, on reversing the recursion, yields for $0 \le k < N$,

$$\pi_{k+1} = \frac{h_{k,k+1}}{h_{k+1,k}} \cdot \pi_k. \tag{13}$$

Let $\alpha(x_1, N) = \lfloor \frac{x_1}{\frac{1}{N+1}} \rfloor$ and $\beta(x_1, N) = \lceil \frac{x_1}{\frac{1}{N+1}} \rceil$. Clearly, $[\frac{\alpha(x_1,N)}{N+1}, \frac{\beta(x_1,N)}{N+1}]$ is the interval that most accurately approximate x_1 given the resolution N. In particular, with

[5] We omit the laborious algebraic steps in the interest of readability.

$z = \alpha(x_1^*, N)$, the solution x_1^* is found in the interval $[\frac{z}{N+1}, \frac{z+1}{N+1}]$. The crucial part of our proof is to reformulate Π in terms of π_z and π_{z+1}, using (12)–(13). More specifically, for $j \in \{1, \ldots, z-1\}$ we have:

$$\pi_j = \pi_z \cdot \prod_{k=z}^{j+1} \frac{h_{k,k-1}}{h_{k-1,k}}. \tag{14}$$

Correspondingly, and arguing in an analogous manner, for $j \in \{z+2, \ldots, N\}$ we have:

$$\pi_j = \pi_{z+1} \cdot \prod_{k=z+1}^{j-1} \frac{h_{k,k+1}}{h_{k+1,k}}. \tag{15}$$

In other words, we represent Π in terms of two of its components: π_z and π_{z+1}.

We are now ready to define the upper bound U for Π:

$$U[i,z] = \begin{cases} \pi_z \cdot M^{z-i} & \text{if } i \leq z, \\ \pi_{z+1} \cdot M^{i-(z+1)} & \text{if } i \geq z+1, \end{cases} \tag{16}$$

where:

$$M = \max\left[\max_{k \leq z}\left\{\frac{h_{k,k-1}}{h_{k-1,k}}\right\}, \max_{k \geq z+1}\left\{\frac{h_{k,k+1}}{h_{k+1,k}}\right\}\right]. \tag{17}$$

As seen, the definition of M clearly makes U an upper bound for Π.

Our final goal is to show that as the resolution N goes to infinity, U goes to zero outside an arbitrarily close vicinity of x_1^*:

$$\lim_{N\to\infty} U[\alpha(x_1, N), \alpha(x_1^*, N)] \to 0 \quad \text{if } x_1 \neq x_1^*. \tag{18}$$

We shall argue that the latter is guaranteed to happen if we have $0 < \frac{h_{k,k-1}}{h_{k-1,k}} < 1$ for $k \in \{2, \ldots, z\}$ and $0 < \frac{h_{k,k+1}}{h_{k+1,k}} < 1$ for $k \in \{z+1, \ldots, N-1\}$, because then we get $0 < M < 1$. We argue this by considering the equilibrium (asymptotic) value of $E[\pi(t)]$ for any finite N. This argument can be separated into three different cases as in [18]:

1. The first case is when $\frac{z}{N+1}$ is close to zero. In this case the maximum is quickly reached and then geometrically falls away.
2. When $\frac{z}{N+1}$ is close to 1, the value of π_i geometrically increases but when the maximum is reached it quickly falls away. For both these cases when $N \to \infty$, most of the probability mass will be centered in a small interval around z.
3. The third case is slightly more complex because it involves $\frac{z}{N+1}$ being away from either end. This case must be broken down into two distinct geometric series, one representing the geometric series from π_1 to π_z and the other from π_{z+1} to π_N. The first series increases until it reaches the maximum at π_Z. The increase is geometric (or rather, exponential as $N \to \infty$), and the geometric ratio is bounded by the bound given by the quantity M above. The second series starts at the maximum at the value π_{z+1} and then decreases until π_N is reached. Again, the decrease is geometric (i.e., exponential as $N \to \infty$), and the geometric ratio is bounded by the quantity M above. In this case the probability mass will be centered within a small interval around $\frac{z}{N+1}$ and $\frac{z+1}{N}$ as $N \to \infty$ because of the law of the *sum* of the elements of a geometric series possessing a common ratio which is greater than unity.

First of all, since the difference between $\frac{k}{N+1}$ and $\frac{k-1}{N+1}$ goes to zero as N goes to infinity, and since $p_1(x)$ is continuous, we have:

$$\lim_{N\to\infty} \frac{h_{k,k-1}}{h_{k-1,k}} = \lim_{N\to\infty} \frac{r_k \cdot p_2(r_k) \cdot q_k}{q_k \cdot p_1(q_k) \cdot r_k} \tag{19}$$

$$= \lim_{N\to\infty} \frac{p_2(r_k)}{p_1(q_k)}. \tag{20}$$

Secondly, from Lemma 1 we can conclude that $p_1(q_k) > p_2(r_k)$ for $k \in \{2, \ldots, z\}$. Therefore, $0 < \frac{h_{k,k-1}}{h_{k-1,k}} < 1$ for $k \in \{2, \ldots, z\}$ as N goes to infinity.

Showing that we have $0 < \frac{h_{k,k+1}}{h_{k+1,k}} < 1$ for $k \in \{z+1, \ldots, N-1\}$ follows analogously, and the proof is left out here for the sake of brevity.

Accordingly, Π must go to zero outside an arbitrarily close vicinity of x_1^* as the resolution N goes to infinity. This, in turn, means that the probability mass of Π will lie arbitrarily close to x_1^*. In other words, the TRAA is asymptotically optimal. □

2.3 Details of the H-TRAA solution

2.3.1 Design of the H-TRAA solution

In this section we propose a hierarchical scheme for solving n-material problems. The scheme takes advantage of the TRAA's ability to solve two-material problems asymptotically, by organizing them hierarchically.

Construction of hierarchy The hierarchy of TRAAs, which we hereafter will refer to as H-TRAA, is constructed as follows.[6] First of all, the hierarchy is organized as a balanced binary tree with depth $D = \log_2(n)$. Each node in the hierarchy can be related to three entities: (1) a set of materials, (2) a partitioning of the material set into two subsets of

[6] We assume that $n = 2^\gamma$, $\gamma \in \mathbb{N}^+$, for the sake of clarity. If the number of materials is less than this, we can assume the existence of additional materials whose values are "zero", and who thus are not able to contribute to the final optimal solution.

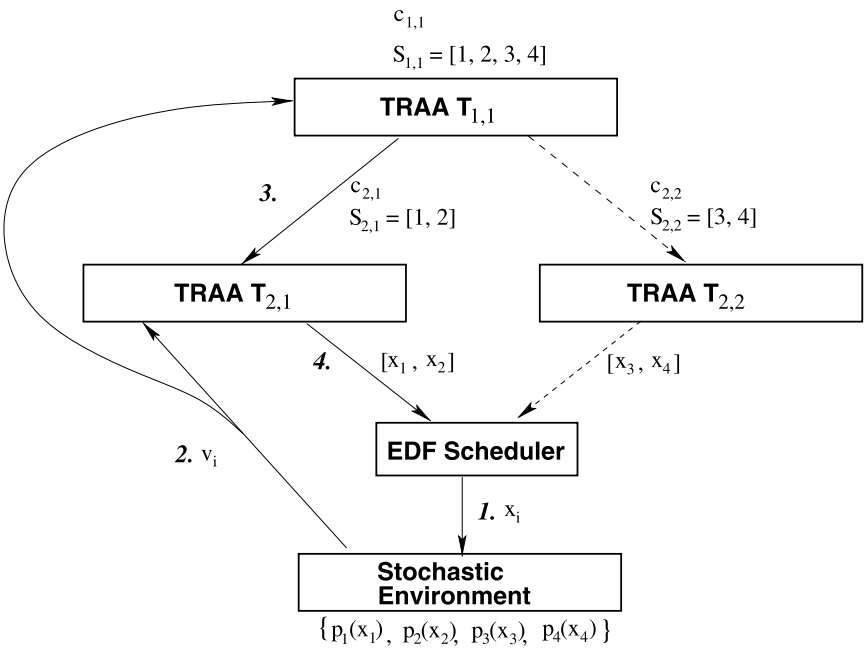

Fig. 4 A Hierarchy of Twofold Resource Allocation Automata (H-TRAA) interacting with a Scheduler and an unknown Stochastic Environment as explained in Example 1

equal size, and (3) a dedicated TRAA that allocates a given amount of resources among the two subsets.

Root node The hierarchy root (at depth 1) is assigned the complete set of materials $S_{1,1} = \{1, \ldots, n\}$. These n materials are partitioned into two disjoint and exhaustive subsets of equal size: $S_{2,1}$ and $S_{2,2}$. An associated TRAA, $T_{1,1}$, decides how to divide the full knapsack capacity c (which, for the sake of notational correctness will be referred to as $c_{1,1}$) among the two subsets. That is, subset $S_{2,1}$ receives the capacity $c_{2,1}$ and subset $S_{2,2}$ receives the capacity $c_{2,2}$, with $c_{2,1} + c_{2,2} = c_{1,1}$. Accordingly, *this* TRAA is given the power to prioritize one subset of the materials at the expense of the other.

Nodes at depth d Node $j \in \{1, \ldots, 2^{d-1}\}$ at depth d (where $1 < d \leq D$) refers to: (1) the material subset $S_{d,j}$, (2) a partitioning of $S_{d,j}$ into the subsets $S_{d+1,2j-1}$ and $S_{d+1,2j}$, and (3) a dedicated TRAA, $T_{d,j}$. Observe that since level $D+1$ of the H-TRAA is non-existent, we use the convention that $S_{D+1,2j-1}$ and $S_{D+1,2j}$ refer to the primitive materials being processed by the leaf TRAA, $T_{D,j}$. Assume that the materials in $S_{d,j}$ has, as a set, been assigned the capacity $c_{d,j}$. The dedicated TRAA, then, decides how to allocate the assigned capacity $c_{d,j}$ among the subsets $S_{d+1,2j-1}$ and $S_{d+1,2j}$. That is, subset $S_{d+1,2j-1}$ receives the capacity $c_{d+1,2j-1}$ and subset $S_{d+1,2j}$ receives the capacity $c_{d+1,2j}$, with $c_{d+,2j-1} + c_{d+1,2j} = c_{d,j}$.

At depth D, then, each individual material can be separately assigned a fraction of the overall capacity by way of recursion, using the above allocation scheme.

Interaction of H-TRAA with EDF scheduler and environment As in the single TRAA case, H-TRAA interacts with an EDF Scheduler, which suggests which unit volume value function $p_i(x_i)$ to access next. A response is then generated from the Stochastic Environment using $p_i(x_i)$. This response is given to all the TRAAs that were involved in determining the material amount x_i, that is, the TRAAs in the hierarchy that have allocated capacity to a material subset that contains material i. Finally, a new candidate material mix $\vec{x} = [x_1, \ldots, x_n]$ is suggested by the H-TRAA to the EDF Scheduler.

Example 1 Consider a 4-material problem. Figure 4 shows the associated hierarchy, constructed as described above. At the root level the TRAA $T_{1,1}$ divides the knapsack capacity among the two material subsets $\{1, 2\}$ and $\{3, 4\}$, respectively related to TRAA $T_{2,1}$ and $T_{2,2}$. At the level below, then, the TRAA $T_{2,1}$ allocates its share of the capacity among material 1 and material 2, while TRAA $T_{2,2}$ assigns *its* share of the capacity to material 3 and material 4. Based on the present assignment at time t, the EDF Scheduler selects material i, suggesting the amount $x_i(t)$ to the Stochastic Environment. The Stochastic Environment, in turn, responds with a randomly drawn material unit volume value, $v_i(t)$, using the probability value function $p_i(x_i)$. By way of example, if $i = 2$, the latter feedback is given to TRAAs $T_{1,1}$ and $T_{2,1}$, which update their states accordingly, and the feedback loop continues.

2.3.2 Analysis of the H-TRAA solution

In the previous section we proved that an individual TRAA is asymptotically optimal. We will now consider the H-TRAA and prove its optimality. More specifically, we shall show that if each individual TRAA in the hierarchy has

solved its own two-material problem, a solution to the complete n-material Knapsack Problem has also been produced.

Theorem 2 *Let $T_{d,j}$ be an arbitrary TRAA at level d of the H-TRAA associated with the node whose index is j. Then, if every single TRAA, $T_{d,j}$, in the H-TRAA has found a local solution with proportions $c_{d+1,2j-1}$ and $c_{d+1,2j}$ satisfying*

$$f'_{d+1,2j-1}(c_{d+1,2j-1}) = f'_{d+1,2j}(c_{d+1,2j}),$$

the overall Knapsack Problem involving n materials that are hierarchically placed in $\log_2 n$ levels of TRAAs, also attains the global optimum solution.

Proof We intend to prove the above theorem by means of induction, using the hierarchical H-TRAA structure defined in the paragraph titled *Construction of Hierarchy*.

Basis The *Basis* case concerns the nodes at the leaves, which, indeed, deal with the primitive materials themselves. Let a and b ($a, b \in \{1, \ldots, n\}$) be any two materials processed by a TRAA, $T_{D,u}$, at a leaf node (i.e., at depth $D = \log_2 n$) in the H-TRAA. The latter decides how to allocate an assigned capacity $c_{D,u}$ among the two materials a and b, with relative proportions x_a and x_b respectively. Observe that since a and b are the only two materials relevant to this TRAA, by virtue of the construction of the TRAA, $\frac{x_a}{x_a+x_b}$ and $\frac{x_b}{x_a+x_b}$ are the conditional probabilities of choosing a and b respectively, conditioned on the event that the knapsack had only to be filled with these primitive materials. Since, by virtue of Theorem 1, we know that the TRAA will find a local solution $[x_a, x_b]$, the foundation of the solution determined by the Lagrangian yields:

$$\begin{aligned} & f'_a(x_a) = f'_b(x_b) \\ \Longrightarrow \quad & f'_{D+1,2u-1}(c_{D+1,2u-1}) = f'_{D+1,2u}(c_{D+1,2u}), \\ & \text{with } c_{D+1,2u-1} + c_{D+1,2u} = c_{D,u}, \end{aligned}$$

thus proving the basis of the induction.

Induction step Consider any interior-node TRAA $T_{d,j}$ whose index at depth d is j in the H-TRAA hierarchy. The TRAA associated with this node decides how to allocate an assigned capacity $c_{d,j}$ among two disjoint subsets $S_{d+1,2j-1} = \{\alpha_1, \ldots, \alpha_m\}$ and $S_{d+1,2j} = \{\beta_1, \ldots, \beta_m\}$ of *composite* materials, where each α_i and β_i is, in itself, a primitive material. To simplify notation, let $\vec{\alpha} = \{\alpha_1, \ldots, \alpha_m\}$ and $\vec{\beta} = \{\beta_1, \ldots, \beta_m\}$. Observe that the union of the sets $\vec{\alpha}$ and $\vec{\beta}$ is the input to the present TRAA, and the task of *this* TRAA is to assign the current knapsack capacity, $c_{d,j}$, so as to satisfying the Lagrangian solution for these two mutually exclusive and exhaustive subsets. Let $T_{d,j}$ assign the relative proportions to $\vec{\alpha}$ and $\vec{\beta}$ by the quantities $x_{\vec{\alpha}}$ and $x_{\vec{\beta}}$. Observe that since $\vec{\alpha}$ and $\vec{\beta}$ are the only two materials[7] relevant to this TRAA, by virtue of the construction of the TRAA, $\frac{x_{\vec{\alpha}}}{x_{\vec{\alpha}}+x_{\vec{\beta}}}$ and $\frac{x_{\vec{\beta}}}{x_{\vec{\alpha}}+x_{\vec{\beta}}}$ are the conditional probabilities of choosing $\vec{\alpha}$ and $\vec{\beta}$ respectively, conditioned on the event that the knapsack had only to be filled with these composite materials $\vec{\alpha}$ and $\vec{\beta}$. The solution to this TRAA will thus satisfy:

$$f'_{\vec{\alpha}}(x_{\vec{\alpha}}) = f'_{\vec{\beta}}(x_{\vec{\beta}}) \quad \text{where,} \tag{21}$$

$$f'_{\vec{\alpha}}(x_{\vec{\alpha}}) = \sum_{\alpha_i \in \vec{\alpha}} \frac{x_{\alpha_i}}{\sum_{\alpha_j \in \vec{\alpha}} x_{\alpha_j}} f'_{\alpha_i}(x_{\alpha_i}) \quad \text{and} \tag{22}$$

$$f'_{\vec{\beta}}(x_{\vec{\beta}}) = \sum_{\beta_i \in \vec{\beta}} \frac{x_{\beta_i}}{\sum_{\beta_j \in \vec{\beta}} x_{\beta_j}} f'_{\beta_i}(x_{\beta_i}). \tag{23}$$

Since each α_i and β_i is a primitive material, and we are working our way *up* the H-TRAA hierarchy, we can invoke the inductive hypothesis to relate x_{α_i} and x_{β_i} for all i. By virtue of the inductive hypothesis and the Lagrangian solution at *every* level up the H-TRAA till level d, we know that for both of the material subsets $S_{d+1,2j-1}$ and $S_{d+1,2j}$ the following are true:

$$f'_{\alpha_1}(x_{\alpha_1}) = \cdots = f'_{\alpha_m}(x_{\alpha_m}), \tag{24}$$

$$f'_{\beta_1}(x_{\beta_1}) = \cdots = f'_{\beta_m}(x_{\beta_m}). \tag{25}$$

To simplify the notation, let each of the quantities in (24) equal $f'_{\alpha}(x_{\alpha})$, and each of the quantities in (25) equal $f'_{\beta}(x_{\beta})$.

Substituting (24) and (25) (which represent the induction hypothesis) into (22) and (23), the latter become:

$$f'_{\vec{\alpha}}(x_{\vec{\alpha}}) = f'_{\alpha}(x_{\alpha}) \sum_{\alpha_i \in \vec{\alpha}} \frac{x_{\alpha_i}}{\sum_{\alpha_j \in \vec{\alpha}} x_{\alpha_j}} \quad \text{and} \tag{26}$$

$$f'_{\vec{\beta}}(x_{\vec{\beta}}) = f'_{\beta}(x_{\beta}) \sum_{\beta_i \in \vec{\beta}} \frac{x_{\beta_i}}{\sum_{\beta_j \in \vec{\beta}} x_{\beta_j}}. \tag{27}$$

The summations on the RHSs of both of (26) and (27) can be trivially seen to sum to unity since they represent probabilities (in the conditioned spaces), implying that:

$$\forall i : f'_{\vec{\alpha}}(x_{\vec{\alpha}}) = f'_{\alpha_i}(x_{\alpha_i}) \quad \text{and} \tag{28}$$

$$\forall i : f'_{\vec{\beta}}(x_{\vec{\beta}}) = f'_{\beta_i}(x_{\beta_i}). \tag{29}$$

Combining the above with Equation (21) yields:

$$\begin{aligned} f'_{\alpha_1}(x_{\alpha_1}) = \cdots = f'_{\alpha_m}(x_{\alpha_m}) = f'_{\beta_1}(x_{\beta_1}) = \cdots \\ = f'_{\beta_m}(x_{\beta_m}) \end{aligned} \tag{30}$$

implying that the global optimum required by the Lagrangian has been found. Hence the theorem! □

[7] The fact that these are composite materials is irrelevant to the present TRAA. It merely treats $\vec{\alpha}$ and $\vec{\beta}$ as individual materials.

Remarks Theorem 2 has some very interesting consequences listed below:

1. The proof of Theorem 2 has tacitly assumed that all the automata have converged before the global convergence can be asserted. This implies that the TRAA $T_{d,j}$ is aware of its capacity, and that this is a known quantity to the TRAAs $T_{d+1,2j-1}$ and $T_{d+1,2j}$. In other words, if all the individual TRAAs converge to their local optimum, Theorem 2 states that the global optimum is attained. Conceptually, this can pose a small implementation-related problem. The fact is that the TRAAs of the lower level are converging even while the TRAA at the higher level is attempting to find *its* capacity. Therefore, essentially, the lower level TRAAs are working in a non-stationary environment. The strategy by which we can resolve this is to ensure that the higher level automata converge at a slower rate than the lower ones (thus guaranteeing a certain level of stationarity). In practice, however, we have observed that if the resolution parameter N is large enough (in the order of hundreds) the time varying phenomenon is marginal, and the TRAAs at all the levels tend to converge simultaneously.
2. Theorem 2 claims that the solution obtained by the convergence of the individual TRAAs leads to the global convergence of the overall optimization problem. But this claim means that the ordering of the materials at the leaf nodes does not carry any significance. This is, indeed, true! It turns out that if the nodes at the leaves are ordered in such a way that "more precious materials" lie in the same sub-tree, the weight associated with the sub-tree of the composite material containing these "more precious materials" will have a much larger weight, and the weight of the other sub-trees will be much smaller. As opposed to this, if the "more precious materials" lie in distinct sub-trees, the weights associated with the respective sub-trees will be correspondingly compensated for.

3 Application: sample size determination for constrained estimation

In this section we consider the problem of estimating the proportion of a population having some specific characteristic. Specifically, we assume that n populations are to be evaluated, and that each population i is characterized by an independent unknown binomial proportion u_i. We will here pursue the goal of minimizing the variance of the proportion estimates when the total number of samples available for estimating the proportions is restricted to c. The purpose is to make the estimates as accurate as possible. The PR implications of this were explained earlier, and will not be visited again to avoid repetition. However, in the interest of clarification, the task at hand could, for instance, be to determine the proportion of a web site that is successfully validated by an HTML validator [24] and/or a WCAG accessibility validator [31], assuming that n web sites are to be evaluated by only accessing c web pages.

3.1 Problem specification

Let x_i be the number of elements sampled randomly from population i and let the count Y_i be the number of the sampled elements that possess a chosen characteristic. For large x_i and when u_i is not too near 0 or 1, the estimator $\hat{u_i} = \frac{Y_i}{x_i}$ is approximately normal with mean u_i and standard deviation $s_i = \sqrt{\frac{u_i(1-u_i)}{x_i}}$ [1]. This standard deviation can be reduced (and the estimate accuracy increased) by increasing the number of samples x_i. In the problem targeted in this section, n different populations can be sampled c times and the goal is to distribute the samples among the populations to minimize the aggregated variance of the estimates. The problem can be reformulated as follows:

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{n} -\frac{u_i(1-u_i)}{x_i} \\ \text{subject to} \quad & \sum x_i = c, \\ & 0 \le x_i, \quad i = 1, \ldots, n. \end{aligned}$$

The above optimization problem is an NEFK problem with concave and separable objective function. Since the u_i's are assumed unknown, we apply our H-TRAA to find a near-optimal solution incrementally and online.

3.2 The H-TRAA solution

We must first define the Stochastic Environment that the H-TRAA is to interact with. That is, we must define the stochastic functions $\mathcal{F}' = \{F'_1(x_1), F'_2(x_2), \ldots, F'_n(x_n)\}$. By applying the principles of Lagrange multipliers we find the following conditions that characterize the optimal solution:

$$\begin{aligned} & \frac{u_1(1-u_1)}{{x_1}^2} = \cdots = \frac{u_n(1-u_n)}{{x_n}^2}, \\ & \sum x_i = c, \\ & 0 \le x_i, \quad i = 1, \ldots, n. \end{aligned}$$

Accordingly, we define $F'_i(x_i)$ as follows. First of all, each time $F'_i(x_i)$ is accessed by the H-TRAA, population i is sampled once and the proportion estimate $\hat{u}_i$ is updated accordingly.[8] After $\hat{u}_i$ has been updated, we instantiate $F'_i(x_i)$

[8] For a dynamic environment we would utilize a "window-based" strategy and only use the last c samples to estimate the u_i's. However, we are currently studying how recently proposed weak estimators can be used in this setting [19].

by a random draw—$F_i'(x_i)$ is instantiated to the value 0 with probability $1 - \frac{\hat{u}_i(1-\hat{u}_i)}{{x_i}^2}$ and to the value 1 with probability $\frac{\hat{u}_i(1-\hat{u}_i)}{{x_i}^2}$. In other words, we keep running estimates of the u_i's in order to calculate the outcome probabilities of the $F_i'(x_i)$'s.[9]

The H-TRAA can be configured by various means. First of all, the material amount space $(0, 1)$ need not be discretized uniformly. Instead, a nonlinear material amount space can be formed, as done for the LAKG in [12]. Furthermore, the discretization resolution N must also be set for each TRAA, possibly varying from TRAA to TRAA in the hierarchy. In short, the performance achieved for a particular problem can be optimized using these different means of configuring the H-TRAA. In this section, however, our goal is to evaluate the overall performance of the H-TRAA, without fine tuning. Therefore, we will only use a linear material amount space, as specified in Sect. 2. Furthermore, we will use the same resolution $N = 500$ for all the TRAAs in the hierarchy, independent of the specific knapsack problem at hand. Thus, our aim is to ensure a fair comparison with the present state of the art, namely, the LAKG scheme.

3.3 Empirical results

3.3.1 Experimental set-up

In this sub-section we evaluate our learning scheme by comparing it with the optimal and uniform policies using simulated data. These policies are described as follows:

1. *Uniform*: The uniform policy allocates the sampling resources uniformly across all 'n' training sets. This is the only classical policy that can be applied directly in an unknown environment.
2. *Optimal*: The optimal policy requires that the binomial proportions are known *a priori*. The optimal solution is then determined by using the principle of Lagrange multipliers [20, 30].

The reader should appreciate that, in practice, we can only apply the uniform policy, because the optimal policy requires that the u_i's are known.

The data used in the experiment is summarized in Table 1. The table shows the true population proportions used, and the number of populations associated with each proportion. The first set of experiments encompasses 512 populations, and the corresponding proportions are to be estimated by allocating 50,000 samples (window based). We have then proceeded to consider a whole suite of experimental settings

Table 1 The true population proportions used in the experiment, and the number of populations associated with each proportion

True Proportion	Populations
0.5	6
0.750/0.250	5
0.900/0.100	41
0.990/0.010	51
0.999/0.001	409

(for example, using up to 16,384 populations), the results of which are conclusive.

As we will see in the following, it turns that one of the strengths of the H-TRAA is its ability to take advantage of so-called spatial dependencies among materials. To be more specific, as seen in Table 1, the materials are spatially ordered in the sense that the proportion of each material decreases/increases with *its* row-index in the table. In order to starve the H-TRAA from this information, we have rather opted to perturb this spatial structure, and present the *perturbed* information to the learning algorithm. We emphasize though that the algorithm is unaware of the ordering, or the fact that such a perturbation has taken place in the background! Each perturbation swapped the proportions of a randomly selected material and the corresponding proportion of a material succeeding it in the ordering. To enable us to conduct experiments with increasing orders of complexity, we have done our experiments with 10^3, 10^4, 10^5 and 10^6 perturbations. For each of these values, an ensemble of several independent replications with different random number streams was performed so as to minimize the variance of the reported results.

The results of our experiments are truly conclusive and confirm the power of the H-TRAA. Although several experiments were conducted using various setting for various numbers of automata, we report, in the interest of brevity, a brief overview of the results obtained.

Figure 5 plots the variance of the current solution (as a function of time) each time a unit volume value function $f_i'(x_i)$ has been sampled. The graphs show the results of applying the H-TRAA with 500 states and the LAKG with 12,500 states (where the amount of the material added on a transition in the latter is not fixed but varying in a nonlinear manner).[10]

As seen in the figure, the H-TRAA steadily reduces the variance of the initial solution in which the populations are sampled uniformly. Indeed, even by the first 50,000 samples, one can observe a very *significant* reduction. The reader should notice that the H-TRAA converges to a near optimal allocation more expediently and far quicker than the

[9] Because the outcome probabilities are always available for the populations, we can normalize the outcome probabilities to speed up convergence.

[10] The details of this are omitted. They can be found in [12].

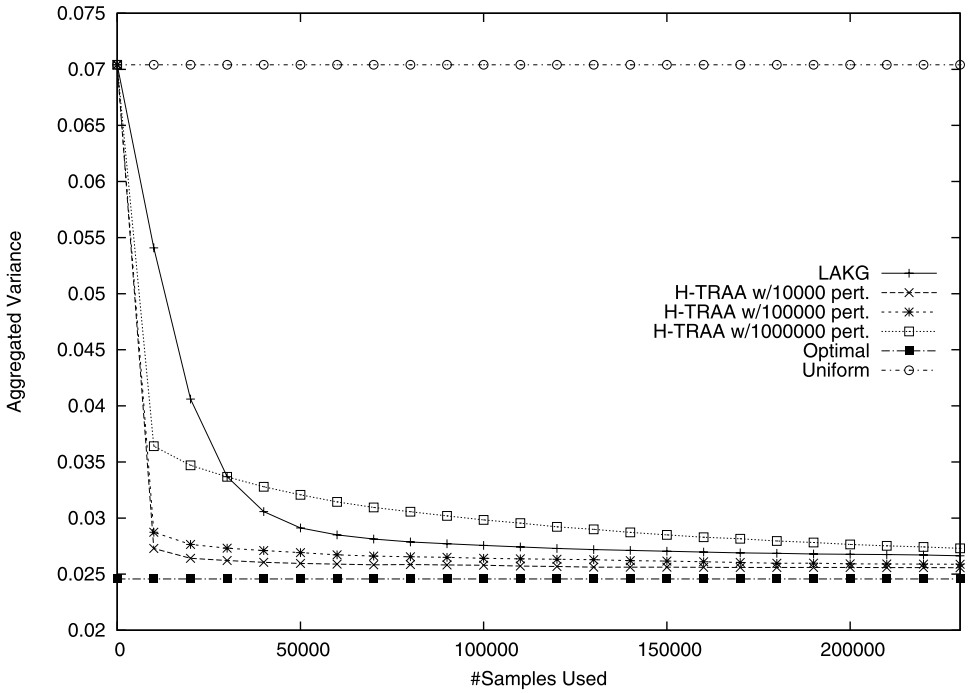

Fig. 5 The H-TRAA steadily reduces the total variance of the initial solution (the uniform policy) as it moves towards near-optimal solutions

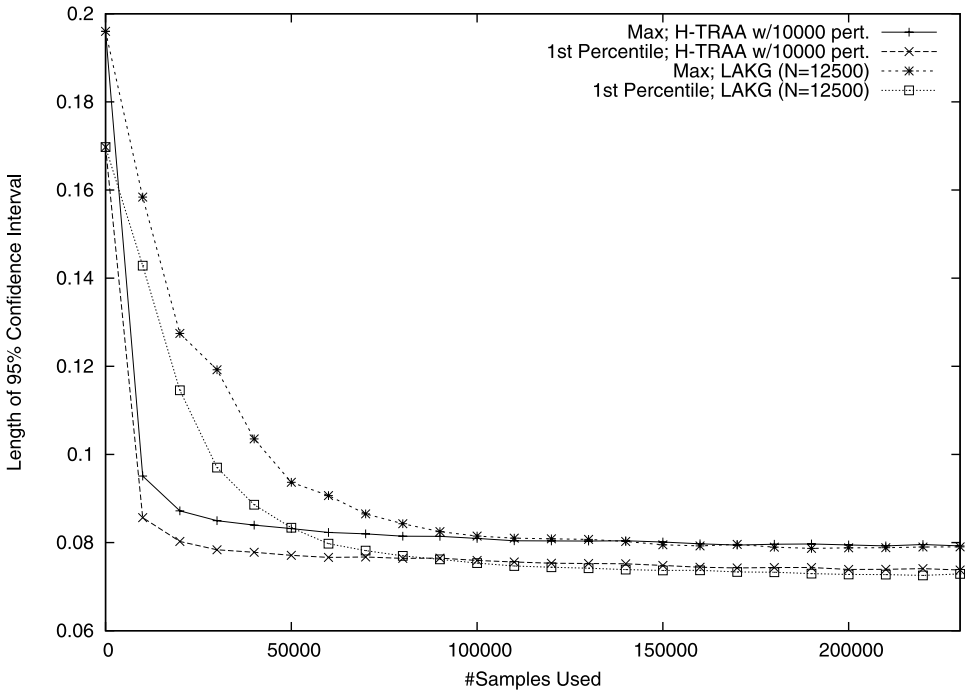

Fig. 6 The confidence interval of each estimated proportion is reduced as the total variance is minimized

LAKG-scheme, expect for the case with 1,000,000 perturbations where the H-TRAA initially converges faster but subsequently in a more conservative manner.

Figure 6 plots the length of the widest 95% confidence interval among the n estimates after each sampling. We also plot the length of the 5th widest interval (1st percentile), whence we see that the confidence interval of each estimated proportion is reduced by minimizing the total variance.

To conclude, our experimental results demonstrate that the H-TRAA is superior to LAKG in spatially structured environments.

3.3.2 Scalability

One of the motivations for designing the H-TRAA was the improved scalability by means of hierarchical learning. As

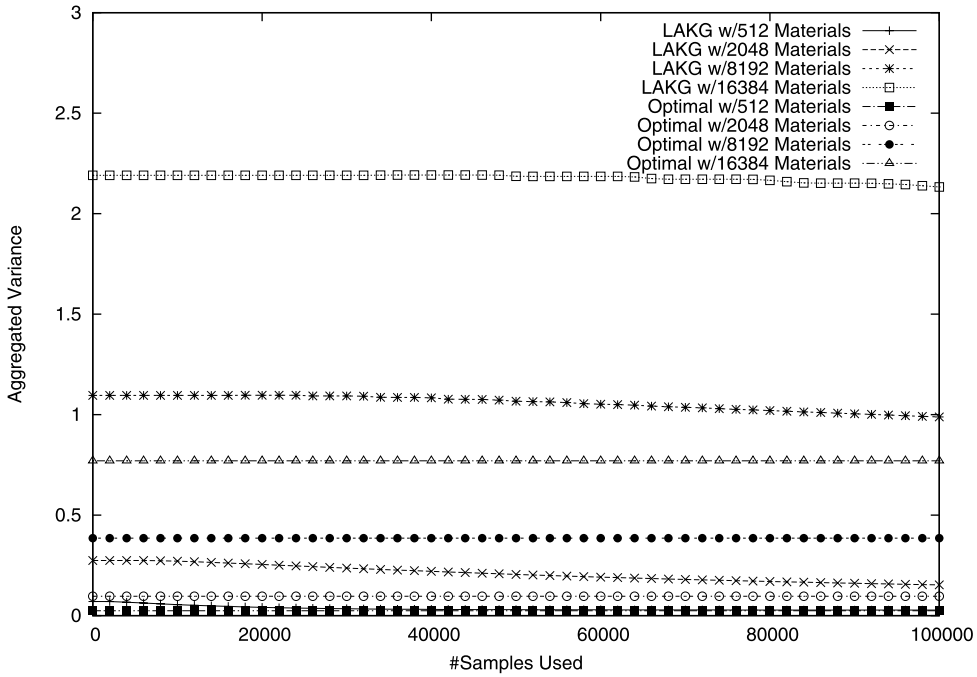

Fig. 7 From the figure we see that extending the number of sets significantly increases the convergence and adaption time of the LAKG

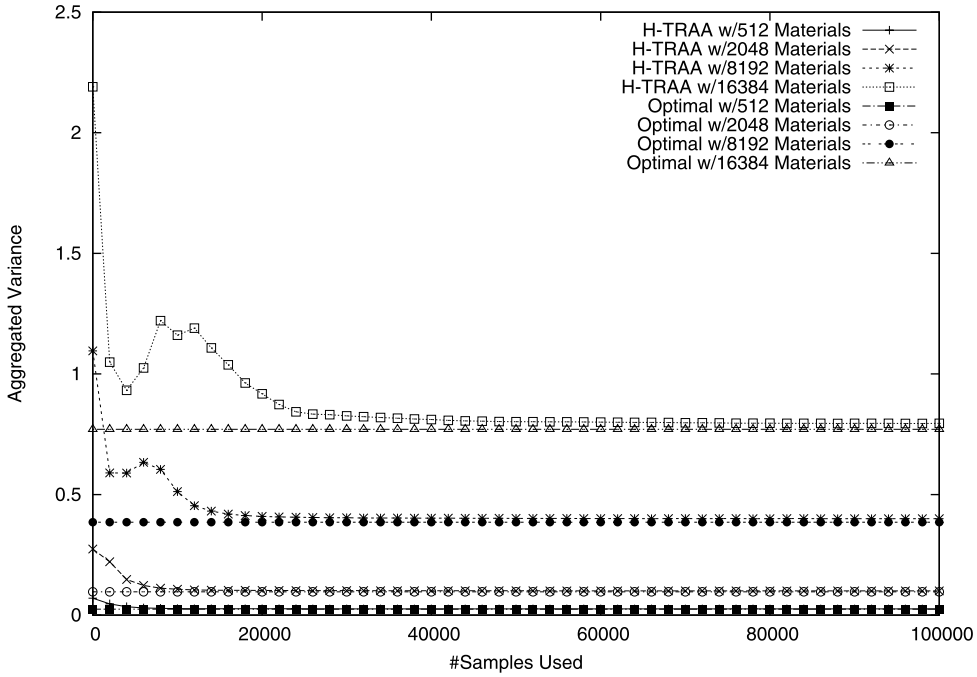

Fig. 8 From the figure we see that the H-TRAA scales sub-linearly with the number of materials. Observe the ability of the H-TRAA to emerge out of local optima

seen in Fig. 7, extending the number of materials significantly increases the convergence time of the LAKG. By comparing Figs. 7 and 8, we further observe that while the LAKG does not even converge, the H-TRAA scales sub-linearly *in every case* with the number of materials. However, the most interesting phenomenon that we observe from Fig. 8 is the ability of the H-TRAA to emerge out of local optima. The H-TRAA first decreases to a minimum, but when it "discovers" that there is a better solution (which in this cases implies a superior partitioning of the nodes in the tree to their left and right subtrees), it is capable of unlearning the inferior configuration and converging to a superior solution. This, we believe, is quite remarkable, especially because the size of the underlying tree is very large, implying that the number of possible binary trees (which grows exponentially with the size), is even larger. However, by arriving at the global optimum, we see that the H-TRAA has

succeeded in learning the best tree structure to resolve the sampling proportions!

4 Conclusions

In this paper, we have considered a problem with real-life PR implications, namely that of allocating limited sampling resources in a "real-time" manner with the purpose of estimating multiple binomial proportions. More specifically, the user is presented with 'n' sets of data points, and with each set containing points drawn from two classes $\{\omega_1, \omega_2\}$, with probabilities u_i and $1 - u_i$ respectively. The problem we have considered involves the interesting and non-trivial cases when both n and each N_i are large, but the number of samples that can be drawn is bounded by a constant, c. The applications of the problem in PR have been alluded to, and (by way of a specific example) shown to be closely related to a world-wide-web application. The problem is further shown to be particularly intriguing because the sampling resources can only be allocated optimally if the binomial proportions are known *a priori*.

Using the general LA philosophy as a paradigm to tackle this real-life problem, our scheme improves a current solution in an online manner, through a series of informed guesses which move towards the optimal solution. To solve the problem, we first modelled it as a *Stochastic Nonlinear Fractional Knapsack Problem*. We then presented a completely new on-line Learning Automata (LA) system, namely, the *Hierarchy of Twofold Resource Allocation Automata* (H-TRAA), whose primitive component is a *Twofold Resource Allocation Automaton* (TRAA). Both the TRAA and the H-TRAA have been proven to be asymptotically optimal.

Comprehensive experimental results demonstrated that performance of the H-TRAA is superior to the previous state-of-the-art scheme, namely, the LAKG. Finally, we have also provided empirical evidence to show that the H-TRAAs possess a sub-linear scaling property, capable of emerging out of local optima.

References

1. Bhattacharyya GK, Johnson RA (1977) Statistical concepts and methods. Wiley, New York
2. Bickel P, Doksum K (2000) Mathematical statistics: basic ideas and selected topics, vol 1, 2nd edn. Prentice Hall, New York
3. Black PE (2004) Fractional knapsack problem. In: Dictionary of algorithms and data structures
4. Bretthauer KM, Shetty B (2002) The nonlinear knapsack problem—algorithms and applications. Eur J Oper Res 138:459–472
5. Casella G, Berger R (2001) Statistical inference, 2nd edn. Brooks/Cole, Pacific Grove
6. Dean BC, Goemans MX, Vondrdk J (2004) Approximating the stochastic knapsack problem: the benefit of adaptivity. In: 45th annual IEEE symposium on foundations of computer science. IEEE, Los Alamitos, pp 208–217
7. Duda R, Hart P, Stork D (2000) Pattern classification, 2nd edn. Wiley, New York
8. Fox B (1966) Discrete optimization via marginal analysis. Manag Sci 13(3):211–216
9. Fukunaga K (1990) Introduction to statistical pattern recognition. Academic Press, San Diego
10. Granmo O-C, Oommen BJ (2006) On allocating limited sampling resources using a learning automata-based solution to the fractional knapsack problem. In: Proceedings of the 2006 international intelligent information processing and web mining conference (IIS:IIPW'06). Advances in soft computing. Springer, Berlin, pp 263–272
11. Granmo O-C, Oommen BJ, Myrer SA, Olsen MG (2006) Determining optimal polling frequency using a learning automata-based solution to the fractional knapsack problem. In: Proceedings of the 2006 IEEE international conferences on cybernetics & intelligent systems (CIS) and robotics, automation & mechatronics (RAM). IEEE, Los Alamitos, pp 73–79
12. Granmo O-C, Oommen BJ, Myrer SA, Olsen MG (2007) Learning automata-based solutions to the nonlinear fractional knapsack problem with applications to optimal resource allocation. IEEE Trans Syst Man Cybern, Part B 37(1):166–175
13. Herbrich R (2001) Learning kernel classifiers: theory and algorithms. MIT Press, Cambridge
14. Jones B, Garthwaite P, Jolliffe I (2002) Statistical inference, 2nd edn. Oxford University Press, London
15. Kellerer H, Pferschy U, Pisinger D (2004) Knapsack problems. Springer, Berlin
16. Narendra KS, Thathachar MAL (1989) Learning automata: an introduction. Prentice Hall, New York
17. Oommen BJ (1986) Absorbing and ergodic discretized two action learning automata. IEEE Trans Syst Man Cybern 16:282–293
18. Oommen BJ (1997) Stochastic searching on the line and its applications to parameter learning in nonlinear optimization. IEEE Trans Syst Man Cybern, Part B 27(4):733–739
19. Oommen BJ, Rueda L (2006) Stochastic learning-based weak estimation of multinomial random variables and its applications to pattern recognition in non-stationary environments. Pattern Recogn 39:328–341
20. Pandey S, Ramamritham K, Chakrabarti S (2003) Monitoring the dynamic web to respond to continuous queries. In: 12th international world wide web conference. ACM, New York, pp 659–668
21. Ross KW, Tsang DHK (1989) The stochastic knapsack problem. IEEE Trans Commun 37(7):740–747
22. Ross S (2002) Introduction to probability models, 2nd edn. Academic Press, San Diego
23. Shao J (2003) Mathematical statistics, 2nd edn. Springer, Berlin
24. Snaprud M, Ulltveit-Moe N, Granmo O-C, Rafoshei-Klev M, Wiklund A, Sawicka A (2003) Quantitative assessment of public web sites accessibility—some early results. In: The accessibility for all conference
25. Sprinthall J (2002) Basic statistical analysis, 2nd edn. Allyn and Bacon, Needhom Heights
26. Steinberg E, Parks MS (1979) A preference order dynamic program for a knapsack problem with stochastic rewards. J Oper Res Soc 30(2):141–147
27. Thathachar MAL, Sastry PS (2004) Networks of learning automata: techniques for online stochastic optimization. Kluwer Academic, Dordrecht
28. Tsetlin ML (1973) Automaton theory and modeling of biological systems. Academic Press, San Diego
29. Webb A (2000) Statistical pattern recognition, 2nd edn. Wiley, New York

30. Wolf JL, Squillante MS, Sethuraman J, Ozsen K (2002) Optimal crawling strategies for web search engines. In: 11th international world wide web conference. ACM, New York, pp 136–147
31. Xiaoming Z (2004) Evaluation and enhancement of web content accessibility for persons with disabilities. PhD thesis, University of Pittsburgh
32. Zhu Q, Oommen BJ (2009) Estimation of distributions involving unobservable events: The case of optimal search with unknown target distributions. Pattern Anal Appl J 12(1):37–53

Ole-Christoffer Granmo was born in Porsgrunn, Norway. He obtained his M.Sc. in 1999 and the Ph.D. degree in 2004, both from the University of Oslo, Norway. He is currently an Associate Professor in the Department of ICT, University of Agder, Norway. His research interests include Intelligent Systems, Stochastic Modelling and Inference, Machine Learning, Pattern Recognition, Learning Automata, Distributed Computing, and Surveillance and Monitoring. He is the author of more than 40 refereed journal and conference publications.

B. John Oommen was born in Coonoor, India on September 9, 1953. He obtained his B.Tech. degree from the Indian Institute of Technology, Madras, India in 1975. He obtained his M.E. from the Indian Institute of Science in Bangalore, India in 1977. He then went on for his M.S. and Ph.D. which he obtained from Purdue University, in West Lafayettte, Indiana in 1979 and 1982 respectively. He joined the School of Computer Science at Carleton University in Ottawa, Canada, in the 1981–82 academic year. He is still at Carleton and holds the rank of a Full Professor. Since July 2006, he has been awarded the honorary rank of Chancellor's Professor, which is a lifetime award from Carleton University. His research interests include Automata Learning, Adaptive Data Structures, Statistical and Syntactic Pattern Recognition, Stochastic Algorithms and Partitioning Algorithms. He is the author of more than 330 refereed book chapters, journal and conference publications, and is a Fellow of the IEEE and a Fellow of the IAPR. Dr. Oommen has also served on the Editorial Board of the IEEE Transactions on Systems, Man and Cybernetics, and Pattern Recognition.