A Statistical Decision-making Model for Choosing among Multiple Alternatives

Shulamit Reches
reches@013.net

Shavit Talman
bimby@012.net.il

Sarit Kraus
sarit@cs.biu.ac.il

Department of Computer Science
Bar-Ilan University
Ramat-Gan 52900, Israel

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence —*Intelligent agents*

General Terms

Theory

Keywords

Decision theory

1. INTRODUCTION

Automated agents often have several alternatives to choose from in order to solve a problem. Usually the agent does not know in advance which alternative is the best one, so some exploration is required. However, in most cases there is a cost associated with exploring the domain, which must be minimized in order to be worthwhile. We concentrate on cases where the agent has some prior knowledge about each alternative, which is expressed in terms of *units of information*. A unit of information about an alternative is the result of choosing the alternative - for example, in the e-commerce domain one unit of information can be a customer's impression or feedback about a supplier; in the heuristic domain one unit of information can be the observed result of running one simulation with a given heuristic function. In our environments the agent has a-priori only a small number of units of information about each alternative, and it would like to use this knowledge in deciding between its alternatives. Nevertheless, since the agent has only a limited number of units of information, deciding between the alternatives solely based on these units may be risky. In extreme cases, they can even mislead the agent to choose the worst alternative rather than the best one.

This paper's primary contribution is a statistical model for making decisions about which problem solving approach to take with limited a-priori information. Towards this goal, we present the EURIKA model, or an innovating model for **e**stimating the **u**tility of **r**estricted **i**nformation among $k \geq 2$ **a**lternatives. The significance of this model is its ability to determine in a time complexity of $O(k)$ how many additional units of information the agent should obtain of each alternative in order to find the best alternative in minimal cost. We assume that in our environments the agent makes its choice in advance as occurs in many domains. For example, in the e-commerce domain, an agent is interested in choosing the best supplier for a certain product - once it chooses the supplier and buys the product it cannot change its mind. This paper is strongly connected to [1]. Azoulay-Schwartz and Kraus constructed a formal statistical model to find the optimal additional units of information in order to find the best alternative. They implemented the model on cases where the agent has to choose between two alternatives. However, when generalizing the model for choosing between $k > 2$ alternatives, their model involved solving a quadrate integral, which in most cases is analytically impossible to calculate. We succeeded to design a model that solves the problem in polynomial time. In [2] presented the *fixed number of experiments* model (FNE model), which decides in advance how many additional units of information the agent should obtain, denoted by N and divides it between the alternatives in proportion to their quality according to the agent's prior knowledge (initial units of information). Experimental results show that the gain of utilizing this approach was very small. In contrast to the FNE model, the EURICA model can be simplified and solved in polynomial time. Moreover, since it does not utilize any heuristic, and also takes into consideration the variance of prior knowledge, it outperforms the FNE model, which only refers to the average value. EURICA avoids an additional drawback of the FNE model in which surplus trials of *all* alternatives are carried out, including the worst ones. In contrast, in a large number of cases EURICA ignores the worst alternatives and does not waste time and resources on pursuing them. Moreover, in this paper we present a *polynomial*-time model, which finds the best alternative among k possible alternatives, at minimal cost within a time complexity of $O(k)$.

2. MODEL CONSTRUCTION

A risk neutral agent has to choose an alternative from a set of k independent alternatives. Each alternative i yields values denoted by x_i. We assume that x_i are normally distributed with an unknown parameter- μ_i. $x_i \sim N(\mu_i, \sigma_i^2)$. Although μ_i is unknown in advance, the agent has some prior beliefs about its distribution. These beliefs are based

AAMAS'07 May 14–18 2007, Honolulu, Hawai'i, USA.

on the knowledge of the agent about the world, but they may be inaccurate. We further assume that μ_i also follows the normal distribution - $\mu_i \sim N(\zeta_i, \tau_i^2)$. The distribution's parameters σ_i, ζ_i and τ_i are known and reflect the agent's beliefs.

The agent has prior knowledge about each alternative in the form of n_i units of information about each alternative i, denoted by $x_{i_1}, \ldots, x_{i_{n_i}}$. These values are numerical and estimate the quality of each alternative. $\overline{x_i}$ is the average value of n_i units of information and it is an estimate of the mean quality of each alternative. Note therefore that $\overline{x_i}$ also follows the normal distribution - $\overline{x_i} \sim N(\mu_i, \sigma_i^2/n_i)$. The agent is interested in improving its decision-making by obtaining additional units of information about each alternative i. As the agent increases the number of information units about an alternative, it has better knowledge about the average value of this alternative. However, this operation is costly, either in time or in direct cost. Thus the agent should minimize the tradeoff between gaining additional information and the cost associated with it. Specifically, the agent is able to obtain m_i additional units of information for each alternative i, $comb = (m_1, \ldots, m_k)$. The agent's goal is to calculate the combination *comb* that will maximize its expected gain. Note that in some cases the utility function may determine that no additional units of information should be obtained for one or more alternatives, that is, in some cases $m_i = 0, 1 \le i \le k$. This may happen in one of two cases, when the agent - 1. Believes its prior knowledge is sufficient; 2. Estimates that obtaining additional information will cost more than the expected improvement in its decision-making.

Obtaining additional units of information takes time. Thus, the agent has a time discount factor of $0 < \delta \le 1$ for each time delay. Often obtaining an information unit involves a direct cost as a result of wasting the agent's resources. We will use $Cost(m_i)$ model to denote the cost associated with obtaining m_i additional units of information about alternative i. Given δ, a *Cost* model, the list of alternatives and the known parameters for each alternative, the agent has to evaluate its expected utility from obtaining *comb* additional information about the alternatives and than to find its maxima. The utility as a function of these parameters is very complex (as will be demonstrated in the next section). However, we were able to calculate a polynomial approximation for the function and, as a result, we were able to obtain the optimal solution in a time complexity of $O(k)$.

2.1 The k-alternative model

In this paper we suggest a statistical model that solves the difficult problem of finding the best alternative from k possible alternatives, in a minimal cost and in polynomial time. Without loss of generality, suppose alternative 1 is currently the best alternative among k alternatives (as determined by the n_i prior units of information). We denote the probability that after $m_1 \ldots m_k$ additional units of information, alternative i is superior to all other alternatives j, assuming they are independent of one another, by $Fchange_i$, $Fchange_i = \prod_{j \neq i} Fchange(m_j, m_i)$ where $Fchange(m_j, m_i)$ is the probability of changing the agent's decision from alternative j to alternative i. $Fchange(m_i, m_j) = Pr(Z > Z_\alpha)$ where Z is a random variable, having a standard normal distribution; $Pr(Z > Z_\alpha)$ is the probability that the random variable Z will take a value greater then Z_α; and Z_α represents the first value where j outperforms i (see [1]).

Since alternative 1 is currently the prior-best alternative, if the agent does not attain any additional information, it will choose alternative 1. In this case, the mean of the cost will be : μ_1. However, The agent will change its decision from alternative 1 to alternative i with the probability of $Fchange_i$. In this case the cost will be: $\sum_{i=2}^{k} Fchange_i \cdot \mu_i$. Furthermore, the agent will continue with alternative 1 with the probability of $Fchange_1 = 1 - \sum_{i=2}^{k} Fchange_i$. Here the total cost is: $(1 - \sum_{i=2}^{k} Fchange_i) \cdot \mu_1$. The expected benefit from obtaining *comb* additional units of information is the difference between the cost of the process without obtaining the additional information, and the cost after obtaining the additional information. Thus :

$$benefit(comb) = \sum_{i=2}^{k} Fchange_i \cdot \mu_i +$$
$$(1 - \sum_{i=2}^{k} Fchange_i) \cdot \mu_1 - \mu_1 = \sum_{i=2}^{k} Fchange_i \cdot (\mu_i - \mu_1) \quad (1)$$

And when considering the various costs involved in obtaining $\sum_{i=1}^{k} m_i$ units of information the agent's utility function is given by:

$$benefit(comb) \cdot \delta^{m_1 + \ldots + m_k} - Cost(comb). \quad (2)$$

Finally, since μ_i are random variables, the agent must multiply the equation by their posterior functions, denoted by $f(\mu_i)$ which is the probability for alternative i to have a mean of μ_i given the prior belief $\mu_i \sim N(\zeta_i, \tau_i^2)$, and the average $\overline{x_i}$ of the n_i units of information about each alternative i. Considering all possible values of μ_i we attain the following proposition. The proof is immediate from the above explanation.

PROPOSITION 1. *The utility of obtaining* comb *additional units of information is*

$$utility(comb) = \int_{-\infty}^{+\infty} \ldots \int [benefit(comb) \cdot$$
$$\delta^{m_1 + \ldots + m_k} - Cost(comb)] \cdot \prod_{i=1}^{k} f(\mu_i) d\mu_1 \ldots d\mu_k \quad (3)$$

when $benefit(comb) = \sum_{i=2}^{k} Fchange_i \cdot (\mu_i - \mu_1)$ *and* $f(\mu_i) = f(u_i = \mu_i | u_i \sim N(\zeta_i, \tau_i^2))$

Now, the maxima of the utility function is actually the optimal allocation $m_1, ..., m_k$ that will maximize the gain. Computing the function's values for all possible $m_1, \ldots, m_k$ is exponential, and therefore not relevant. Another possibility is to find the maximum in the specific relevant radius when $0 \le \sum_{i+1}^{k} m_i \le M$ in an analytic way (such as division etc.). But, the function in its current format is very complicated, thus it is very difficult or even impossible to find its maxima in this manner. Consequently, in order to enable these analytic calculations, we used a polynomial approximation of the *Fchange* function, which enabled us to integrate it, to obtain a power series expansion and then to find the maxima of the utility analytically. This analytic way takes at most $O(k)$, and therefore we attain the expected optimal allocation in a polynomial time. Note that we are interested in $m_1, .., m_k$ which are integers and therefore some inexactitudes can be afforded.

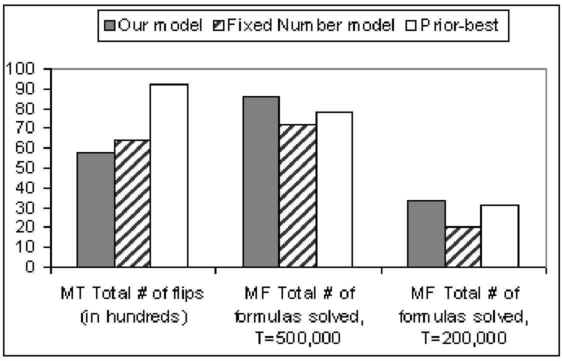

Figure 1: Average # of flips per formula in the MT scenario (the lower the better); Average # formulas solved in the MF scenario (left) T=500,000; (right) T=200,000; (the higher the better)

3. EXPERIMENTAL DESIGN AND ANALYSIS

We chose to evaluate EURIKA in the 3-SAT domain. In this domain the agent has k possible heuristic functions to utilize in order to solve 3-SAT formulas, and must select the best heuristic. We investigated two scenarios within this domain, namely - (1) Minimal Time (MT) scenario - the agent has to solve M formulas as quickly as possible; (2) Maximal Formulas (MF) scenario - the agent has T units of time to solve as many formulas as possible. We assumed each Truth-assignment flip takes one unit of time. The utility function of the MT scenario is as follows:

$$utility(comb) = \int \ldots \int [-\sum_{i=1}^{k} m_i \cdot (\mu_i - \mu_1) + (M - \sum_{i=1}^{k} m_i) \cdot \sum_{j=2}^{k} Fchange_j \cdot (\mu_1 - \mu_j)] d\mu_1 \ldots d\mu_k \quad (4)$$

The utility function of the MF scenario is given by -

$$utility(comb) = \int \ldots \int [\sum_{i=2}^{k} m_i \cdot (1 - \frac{\mu_i}{\mu_1}) + \sum_{j=2}^{k} Fchange_j \cdot (\frac{1}{\mu_j} - \frac{1}{\mu_1}) \cdot (T - \sum_{i=1}^{k} m_i \mu_i)] d\mu_1 \ldots d\mu_k$$

4. EXPERIMENTAL RESULTS

We chose best-known search algorithms for solving 3-SAT problems: Greedy-SAT (GSAT) and GSAT with Random Walk (Random). The latter was implemented with three different walk probabilities: 40%, 60% and 80%. We chose the number of formulas to be solved in the MT scenario, M, to be equal to 300. The number of given flips to be used in the MF scenario, T, was set to 200,000 in one setting, and to 500,000 in an additional setting. This was done in order to compare our results to the results presented in [2].

We found the EURIKA model improved the average number of flips the agent used per formula in the MT scenario. Figure 1 presents the average number of flips per formula for this scenario (left hand-side). The average number of flips per formula after applying EURIKA was 5825, while the average number according to the prior-best alternative was 9273 (T-test PV < 0.002). Namely, the agent was able to save up to 47% flips per formula. Moreover, EURIKA used statistically significant less flips per formula than the FNE model (5825 vs. 6390, T-test PV$=0.054$). Furthermore, the average number of additional trials the model outputted was only 16.35 (similar to the FNE model's). Figure 1 also presents the average number of formulas solved in the MF scenario for the two settings (middle T=500,000 flips, right hand-side T=200,000 flips). The agent was able to solve more formulas by applying our model - in the first setting it solved 86 formulas on average vs. only 78 on average had it used the prior-best heuristic (T-test PV $= 0.05$). Furthermore, it was significantly superior to the FNE model, solving on average 14 more formulas (T-test PV < 0.001). The model outputted on average only 8.15 additional trials to perform - about 50% less than the FNE model. In the second setting the agent solved 34, 20 and 32 formulas for each approach, respectively.

According to ...	GSAT	Random 40%	Random 60%	Random 80%
Prior-best	7.5%	22.5%	50.0%	20.0%
FNE model	0.7%	25.1%	67.2%	7.0%
EURIKA	0.0%	19.3%	76.0%	4.7%

Table 1: 3-SAT % of choosing each heuristic in the MT scenario according to the different approaches

Table 1 summarizes the percentages of choosing each heuristic according to the prior-best alternative, according to the FNE model and, lastly, according to EURIKA model in the MT scenario. Recall that Random 60% proved to be the best alternative, Random 40% was second best, Random 80% took third place and GSAT was the worst. By applying the model, the agent succeeded in choosing the best heuristic 27.2% more frequently than by applying the prior-best alternative. It proved to be statistically significantly superior to using only the prior information the agent gathered (Chi-square test, PV < 0.001). Moreover, EURIKA outperformed the FNE model in each category (Chi-square test, PV $= 0.002$). Note that EURIKA did not spend any valuable flips on Random 80%, the worst alternative, as did the FNE model.

Acknowledgments: This work was supported in part by NSF under grant #IIS0222914.

5. REFERENCES

[1] R. Azoulay-Schwartz and S. Kraus. Acquiring an optimal amount of information for choosing from alternatives. In *Proceedings of the 6th International Workshop on Cooperative Information Agents VI*, pages 123–137, 2002.

[2] S. Talman, R. Toester, and S. Kraus. Choosing between heuristics and strategies: an enhanced model for decision-making. In *Proceedings of the International Joint Conference of Artificial Intelligence*, pages 324–330, 2005.