Acquiring an Optimal Amount of Information for Choosing from Alternatives*

Rina Azoulay-Schwartz[1] and Sarit Kraus[1,2]

[1] Department of Computer Science, Bar-Ilan University, Ramat-Gan, 52900 Israel
[2] Institute for Advanced Computer Studies
University of Maryland, College Park, MD 20742

Abstract. An agent operating in the real world must often choose from among alternatives in incomplete information environments, and frequently it can obtain additional information about them. Obtaining information can result in a better decision, but the agent may incur expenses for obtaining each unit of information. The problem of finding an optimal strategy for obtaining information appears in many domains. For example, in ecommerce when choosing a seller, and in solving programming problems when choosing heuristics. We focus on cases where the agent has to decide in advance on how much information to obtain about each alternative. In addition, each unit of information about an alternative gives the agent only partial information about the alternative, and the range of each information unit is continues. We first formalize the problem of deciding how many information units to obtain about each alternative, and we specify the expected utility function of the agent, given a combination of information units. This function should be maximized by choosing the optimal number of information units. We proceed by suggesting methods for finding the optimal allocation of information units between the different alternatives.

1 Introduction

An agent which has to choose from among alternatives in an incomplete information environment, would like to obtain information about them. The information about an alternative may enable the agent to compute its expected utility from choosing this alternative, and also the risk associated with this alternative. Given this knowledge, the agent will be able to make a better choice between the available alternatives.

Often, there is an expense associated with obtaining information. This expense may be for the time used to seek on-line information, the time spent talking to friends, the cost of buying relevant journals or the cost of searching a commercial database. We focus on situations where the information obtained is only partial. A unit of information about an alternative means one observation of the result due to the choice of this alternative. Thus, as the agent increases the number of information units about an alternative, it has better knowledge

* This work was supported in part by NSF under Grant No. IIS-0208608.

about the average value of this alternative. For example, an information unit in the ecommerce domain may be an impression of one customer from a product it bought. In choosing between heuristics, an information unit may be a result of one simulation. When using a remote information agent, one information unit may be a result of one query to an external database, etc.. The agent should evaluate its utility from additional information units about each alternative, in order to decide which information to obtain before making a decision.

This problem appears in ecommerce. A customer in an electronic market often has to choose between several suppliers of a product or a service. In ecommerce, the customer cannot view the product before buying it, and cannot form a personal impression of the supplier. Moreover, the price of the product cannot identify its quality, as shown in [4]. Thus, the customers have to collect information about the suppliers in order to learn about the quality of their products. The quality of the goods in ecommerce is measured by its properties, such as supply time, life time of the product, customer support, etc. The utility of the customer given a particular item depends on the item's quality and its price, and the customer should choose the supplier that will maximize its expected utility.

Another domain where information is crucial is in choosing between heuristics when developing a software or hardware product. Each heuristic can be tested a priory by simulations, but testing requires time and resources. Once a heuristic is chosen, it will be implemented in the product, and provided to the customers. The agent should determine the optimal number of trials to run on each heuristic in order to decide which heuristic will be integrated in the commercial product. Suppose the programmer has to decide how many simulations to run during the night for each heuristic. The problem can be formalized as a decision about allocation of information units, since each simulation result can be considered an information unit. Running a simulation has costs of computational power, and there is a limit on the total simulations that could be run in a given time period. In this case, there is a limit on the total number of information units, but not on the number of units for each alternative. Obviously, obtaining more information will improve the agent's knowledge about the different alternatives and will improve its decision making process. Since obtaining information incurs costs of time, communication and other resources, the agent should determine the optimal number of units of information to be obtained about each alternative, in order to optimize its utility from the final decision.

We assume that the agent should decide in advance on how many information units to obtain about each alternative, and it cannot change its decision during the information obtaining process. This assumption holds in different situations. For example, consider an agent that interacts with a search engine. The agent should specify how many answers it would like to obtain about each alternative. It cannot ask for additional answers from the search engine, since the answers will overlap with the previous ones. Thus, the agent should decide how many answers it wants. Another situation where the number of trials should be determined in advance is in sending queries or jobs to multiple machines. Consider, for example, the heuristics domain. Suppose the tester has M machines where he can test its heuristics, but he has to decide how many machines should be allocated to each heuristic. Again, the decision should be taken in advance, and it cannot be changed after observing part of the results, since the results are obtained simultaneously from all the machines.

A similar problem can be found in running a survey. Suppose an agent wants to ask clients of different companies about their satisfaction. The agent sends email to the clients of the different companies, and waits for answers. The agent should decide how many emails to send, assuming that there is a cost of sending emails. (time, communication, etc.) Also in this domain, the decision on the number of information units to be obtained on each alternative is made before answers are obtained.

Our problem is different from the k-armed bandit problem [5], since no costs or benefits are associated with the result of an observation, whereas in the bandit problem, the utility of the agent is composed of the results of the observations. In addition, in our problem, after the sampling process is over, the agent decides about the alternative to be used, while in the k-armed bandit problem, an item is chosen repeatedly, and the decision on the chosen item may be changed over time, according to the observed results. In ecommerce, a bandit problem arises when the agent has to choose repeatedly which item to buy, while our problem of optimal amount of information arises when the item is bought only once, but information can be collected from other entities prior to the buying event.

Furthermore, our problem can be distinguished from classical *value of information* literature [17, 7]. In the latter, it is usually assumed that a piece of information means an exact evidence about the value of some random variable, while in our case, each observation gives only partial information about a given alternative. In addition, in our situations, there is an infinite number of possibilities of information units, since the value of the mean of each alternative, and the value of each observation has a continuous distribution. The assumption about a continuous set of possible answers seems to be more realistic in real world domains. For example, in ecommerce, the quality of an item can be a real number (such as weight, lifetime, etc.) and also an evidence of the quality of a particular item can be, again, a real number. Also in the heuristic domain, the time or the quality of the result can, again, obtain any real number in a given interval. Thus, different methods than those used in the classical literature should be used to solve the current problem.

In this paper, we first present relevant related work. We proceed by describing the case of deciding between two alternatives, and then we consider the general case of deciding between multiple alternatives. For both cases, we introduce a formal model of the problem and identify the utility from obtaining a given combination of information units for each alternative. We suggest how to decide on the optimal combination of information units to be obtained, in order to optimize the agent's expected utility. We also propose some heuristics for the case where there are multiple alternatives and a large possible number of information units for each of them. Finally, we conclude and propose directions for future work.

2 Related Work

In this section we describe research related to the information acquiring problem in AI and in statistics. Some work in AI has considered the decision about acquiring information, during the process of planning or decision making in incomplete information environments.

The problem of *value of information* was widely discussed in Artificial Intelligence, e.g., [8, 17, 7, 6]. However, as specified by Russel and Norvig, 'Usually, we assume that exact evidence is obtained about the value of some random variable.' In this paper, we consider the case where the acquired data gives only partial information, as described in Section 1. In particular, each unit of information about a particular alternative includes one evidence about the alternative.

Grass and Zilberstein [6] developed a decision theoretic approach that uses an explicit representation of the user's decision model in order to plan and execute information gathering actions. However, their system is based on information sources that return perfect information about the asked query. Similarly, Dean and Wellman [2] studied the problem of decision making on information acquisition to improve planning, but consider only cases where the answers consist of complete knowledge about the questions asked. In our paper, the information obtained about an alternative consists of a sample of this alternative, but does not provide a complete answer on the value of the given alternative.

Lesser et al. [11, 10] developed BIG, which is a sophisticated information gathering agent that retrieves and processes on-line information to support decision making. They supplied BIG with a design-to-criteria scheduler in order to control the following three factors: (a) the money spent on acquiring information from sites that charge a fee for accessing their information; (b) the balance between the coverage of information gathered and the precision of the results; and (c) the time of the overall process of information gathering. The design-to-criteria scheduler analyzes the agent's set of problem solving actions and chooses a course of actions for the agent. Lesser et al. consider the overall process of decisions about information gathering, but do not provide a formal model for optimal information acquisition, considering the precision of the results, and the time of the overall process.

Some ongoing research considers also the case of partial evidence about the missing information (the hypothesis) [7, 18]. Tseng and Gmytrasiewicz [18] considered sampled answers, but they assume that the number of possible answers for a query in the information gathering process is finite while in this paper we consider a continuous set of possible answers. Moreover, Tseng and Gmytrasiewicz consider a myopic sequential procedure for the information gathering process. Thus, their solution is not optimal: they only consider the nearest step of information gathering, assuming that in each step the agent can decide about the next information to be obtained. However, in our case the combination of information units to be obtained should be decided in advance and we are looking for optimal solutions.

Heckerman et al. [7] consider the case of partial information obtained about a hypothesis, but they consider a simplified model where there is one binary hypothesis, and the data that can be learned is based on binary evidences about several unknown variables. They suggest to consider all possible sets of results, in order to choose the next data item to obtain. In a case of a large number of binary variables, they suggest to use a normal approximation. The problem addressed in this paper is different, since each data item has a continuous set of possible values, so different computation techniques should be used. Moreover, the decision in this paper is which alternative to choose out of N possibilities, and not just to accept or reject a particular hypothesis. Thus, our problem is more complex than that of Heckerman et al.

Poh and Horvitz [16] consider a situation where the information obtained due to the model refinement is only partial, and it can have continuous probability distribution. They consider several types of model refinements, and they also consider the problem of which refinement to choose, when there are different alternative refinement steps. They consider the case where after each refinement, a new decision is made about the next refinement to do, and they suggest a greedy-myopic algorithm for the decision about the next refinement to be performed in each step. In this paper, we consider the problem of how to decide optimally about the sample allocation, in cases where the decision should be taken in advance.

A set of problems related to ours is the family of the bandit problems [1, 5]. In this kind of problems, the agent has to choose sequentially between alternatives. Each alternative when observed produces a result, and the agent's utility is based on a weighted sum of the results obtained over time. Our problem is different from the bandit problem since costs or benefits are associated with the result of each observation. However, in our problem the utility of the agent is composed only of the outcome of the final decision . In addition, in our problem, after the sampling process is over, the agent decides about the alternative to be used. In the k-armed bandit problem an item is chosen repeatedly, and the chosen item can be replaced over time, according to the observed results. For example, in the heuristics domain, if the final software includes a mechanism that can choose which heuristic to use in each step, and can change its decision using the results of the previous steps, then the problem will be similar to the k-armed bandit problem, since the result of each step is important. But, if there is a trial step, where all alternatives are tested, and only the winner is implemented in the final software, then there is a problem of the value of information: the agent does not care about the results of the sampling tests, and is only concerned with the final decision and the cost of performing the tests.

Most of the work done in statistical research in the context of determining the size of a sample, considers a criteria of reaching a required accuracy of the result of the sample units [13, 9]. Few researches deal with trying to maximize future benefits. Dunnett [3] considers the problem of determining the number of samples to be taken when deciding among multiple alternatives. However, he considers only the case of an equal number of samples to be taken about each alternative. Moreover, he assumes equal standard deviation of the different alternatives. Based on these assumptions, the optimal number of samples can be found analytically, at least for the case of two alternatives. However, in reality, the standard deviation of the different alternatives can be different, and there is also no reason for the number of samples to be the same for all the alternatives. Thus, the general case is much more complex and an analytical solution can not be found in most of the cases.

Lindley [12] describes a full Bayesian treatment for the problem of sample size determination, and compares it to other approaches. However, he considers a decision regarding whether to accept or reject a particular hypothesis, while in our research, the final decision is a choice between several alternatives.

Pezeshk and Gittins [15] consider the problem of sample size determining in the context of medical trials. They suggest how to choose the optimal number of trials to perform on a new medicine, in order to maximize the social benefits, or to maximize the benefits of the medical company. They also consider only a decision about accepting or rejecting one alternative, where the decision to be made is whether or not to accept the alternative.

In this paper, we consider the problem of sample size determination, where an agent that has to decide between several different alternatives by sampling them. Similar to the statistical research, we consider a continuous distribution of the values of each alternative. However, we suggest procedures for choosing from multiple alternatives, while current research in statistics mostly considers the decision of whether to accept or reject a particular hypothesis.

3 Environment Description

Consider a risk neutral agent that has to choose from among k alternatives. After choosing alternative i, the agent will obtain a value of x_i, which is unknown in advance. x_i has a particular distribution, but the distribution is unknown to the agent. In particular, we assume that for each alternative i, x_i is normally distributed, with an unknown mean μ_i, and a known standard deviation σ_i.[3] The agent does not know μ_i, but it has some prior beliefs about its distribution. The agent believes that the mean μ_i for each alternative i is normally distributed, with mean ζ_i and standard deviation τ_i. Formally, $x_i \sim N(\mu_i, \tau_i)$ and $\mu_i \sim N(\zeta_i, \tau_i)$. The prior beliefs are based on the knowledge of the agent about the world. For example, its knowledge about the average quality of an item, etc., but its prior beliefs may be inaccurate.

The agent has some available information about each alternative. For each alternative i, it has $n_i \geq 0$ units of information, with an average value of $\overline{x}_i$ per unit. The agent is able to obtain additional information units about the different alternatives, but this operation is costly. Collecting each unit of information takes one time period, and the agent has a discount factor of $0 < \delta \leq 1$ for each time delay. Suppose also that asking a query has a direct cost of $c \geq 0$. This may be

[3] In order to simplify the model, we assume a known standard deviation. If the standard deviation is unknown by the agent, then we can use the student (t) distribution, and analyze it accordingly.

the payment to the answering agent, the cost of a phone call, or other expenses associated with the query process.

Using the above parameters, the posterior distribution of x_i can be calculated, and used to calculate the value of information. We suggest that the final decision about the winner alternative be made only according to the collected information, and the prior distribution of x_i be considered, since this distribution is based on beliefs that are inaccurate. Dunnett [3] refers to this type of procedure as Procedure D_0. Therefore, the parameters x_i and n_i, which are used both for the calculation of posterior beliefs and for the final decision, and the parameters ζ_i and τ_i, which are used only for the calculation of posterior beliefs should be considered in a different manner.

Given the cost of time, δ, and the direct cost c, and given the list of alternatives, and the parameters $(\sigma_i, \zeta_i, \tau_i, x_i, n_i)$ for each of the alternatives, the agent should decide how much information to obtain about each alternative. We denote by m_i the number of information units to be obtained about alternative i. Thus, the agent should choose the combination $(m_1, m_2, .., m_n)$. We assume that $m_i \in \{0, .., M-1\}$, i.e., the maximum number of information units for each alternative is $M - 1$. Thus, there are M possibilities for the number of units to be obtained about each alternative. There are also situations where the total number of samples is limited, i.e., $\sum m_i < M$. For example, this is the case in the heuristics domain when there are M machines and the agent has to decide how to divide its trials among them. In this case, there is only a limit on the total number of samples, which is supposed to be at most $M - 1$.

In the following example, we present a particular ecommerce problem where there are two alternative suppliers and the agent has to choose between them. This example will be used in order to illustrate the evaluation process and the algorithm to determine the optimal amount of information to obtain.

Example 1. Suppose a customer has to buy a particular item, that can be sold by two different suppliers, A and B. Suppose the price of the item is the same for both suppliers, but the quality of the item cannot be observed in advance. The average quality of supplier A is μ_A and the average quality of supplier B is μ_B, but the customer does not know the values of μ_A and μ_B. However, the customer can collect information from friends or from the Web about these suppliers, in order to be able to decide between them. Since the prices of the suppliers are equal, the expected utility of the buyer depends on μ_A and μ_B. Obviously, the customer would like to choose the better supplier, but first it has to decide how much information to obtain about each of the suppliers.

We consider a case where the customer does not have specific beliefs about the suppliers A and B, but has prior beliefs about the distribution of μ_i for any unknown supplier i (suppose these are the average and standard deviations of the item given any arbitrary supplier). In particular, the customer believes that $\mu_A \sim N(\zeta_A = 50, \tau_A = 50)$ and $\mu_B \sim N(\zeta_B = 50, \tau_B = 50)$. Suppose also that σ_A and σ_B are equal and known to be 50. Finally, suppose that the discount factor of the agent is $\delta = 0.9$ and the cost is $c = 0$.

Suppose that the agent was able to obtain available information about both suppliers from some external sources (such as a free search engine), and it has collected 10 units of information about each alternative, with the results $\overline{x}_A = 49, \overline{x}_B = 47$. Given this available information, it is clear that without additional information, the agent will choose to buy the product from supplier A. However, the agent is able to collect additional information about each supplier, from some information source (for example, from another search engine, but reading each answer takes time, so a discount factor exists).

The agent should decide how much additional information to collect about each supplier. It should decide in advance how many units it wants for each alternative: it cannot ask for additional answers from the search engine, since the answers will overlap with the previous ones. Formally, we would like to find the optimal numbers M_A, M_B, where M_A is the number of information units to be obtained about supplier A, and M_B is the number of information units to be collected about supplier B.

In the following section we show how the agent should decide on the number of information units to obtain about each alternative i, given its prior beliefs about i, and given $\overline{x}_i$ and σ_i.

4 Choosing between Two Alternatives

Suppose the agent has to decide between alternatives A and B. Currently, $\overline{x}_A > \overline{x}_B$. Since the agent is risk neutral, and since the decision is made according to the collected information, alternative A will be chosen if no additional information is obtained. If additional information will be acquired the decision may be changed either because of additional negative information about A, or because of additional positive information about B. Thus, we should evaluate the expected utility resulting from the additional information.

Suppose that it is possible to obtain m_A additional units of information about alternative A and m_B additional units of information about alternative B. Denote by $\overrightarrow{x}_A$ and $\overrightarrow{x}_B$ the average value of these additional units, for alternative A and B, respectively. We start by calculating the value of the information from obtaining these additional units.

The agent has to evaluate its utility from obtaining additional information about the alternatives. In order to do so he has to evaluate the probability that the additional information will change his decision. We denote this probability by $Fchange(\mu_A, \mu_B, \sigma_A, \sigma_B, \overline{x}_A, n_A, m_A, \overline{x}_B, n_B, m_B)$. In the following lemma, we find the value of $Fchange$, as a function of its arguments.

Lemma 1. Calculating Fchange
The value of Fchange, is as follows:
$Fchange(\mu_A, \mu_B, \sigma_A, \sigma_B, \overline{x}_A, n_A, m_A, \overline{x}_B, n_B, m_B) = Pr(Z > Z_\alpha)$
where Z is a random variable, having the standard normal distribution (see [13]), $Pr(Z > Z_\alpha)$ is the probability that the random variable Z will take a value greater

than Z_α, and

$$Z_\alpha = \frac{\mu_A m_A(m_B+n_B) - \mu_B m_B(m_A+n_A)}{\sqrt{m_A^2(m_B+n_B)^2 \cdot \sigma_A^2 + m_B^2(m_A+n_A)^2 \cdot \sigma_B^2}} + \frac{n_A \overline{x}_A(m_B+n_B) - n_B \overline{x}_B(m_A+n_A)}{\sqrt{m_A^2(m_B+n_B)^2 \cdot \sigma_A^2 + m_B^2(m_A+n_A)^2 \cdot \sigma_B^2}}$$

Sketch of proof: Given the additional m_A, m_B units of information, the new average quality of A will be $(m_A \overrightarrow{x}_A + n_A \overline{x}_A)/(m_A+n_A)$, and the new average of B will be $(m_B \overrightarrow{x}_B + n_B \overline{x}_B)/(m_B+n_B)$. In order for the average of B to outperform the average of A, the following should hold: $(m_B \overrightarrow{x}_B + n_B \overline{x}_B)/(m_B+n_B) > (m_A \overrightarrow{x}_A + n_A \overline{x}_A)/(m_A+n_A)$. Manipulating the above formula, the above condition is $\overrightarrow{x}_B m_B(m_A+n_A) - \overrightarrow{x}_A m_A(m_B+n_B) > n_A \overline{x}_A(m_B+n_B) - n_B \overline{x}_B(m_A+n_A)$. Denote $e_A = m_A(m_B+n_B)$ and $e_B = m_B(m_A+n_A)$. Then we prove that $e_B \cdot \overrightarrow{x}_B - e_A \cdot \overrightarrow{x}_A \sim N(e_B \cdot \mu_B - e_A \cdot \mu_A, \sqrt{e_A^2 \cdot \sigma_A^2/m_A + e_B^2 \cdot \sigma_B^2/m_B})$. Based on this,

$$\frac{e_B \cdot \overrightarrow{x}_B - e_A \cdot \overrightarrow{x}_A - (e_B \cdot \mu_B - e_A \cdot \mu_A)}{\sqrt{e_A^2 \cdot \sigma_A^2/m_A + e_B^2 \cdot \sigma_B^2/m_B}} \sim N(0,1).$$

Using the above, we find the value of Z_α, such that

$$Z_\alpha = \frac{n_A \overline{x}_A(m_B+n_B) - n_B \overline{x}_B(m_A+n_A) - (e_B \cdot \mu_B - e_A \cdot \mu_A)}{\sqrt{e_A^2 \cdot \sigma_A^2/m_A + e_B^2 \cdot \sigma_B^2/m_B}}$$

and the probability for a change is the probability of a normal variable to be more then Z_α. □

In the following example we will calculate *Fchange* for different combinations of m_A and m_B, in the suppliers examples, and show how *Fchange* will be affected by changing m_A or m_B.

Example 2. We return to the case of Example 1. Recall that the data collected before the decision should be made includes 10 units of information about each alternative, where $\overline{x}_A = 49$ and $\overline{x}_B = 47$. In order to calculate the values of *Fchange* for different combinations of m_A and m_B, we should consider all possible values of mean value μ_A and mean value μ_B. Intuitively, the influence of m_A and m_B on *Fchange* depends on the values of μ_A and μ_B.

For example, suppose that $\overline{x}_A = 49$, and additional units are obtained only about alternative B. In this case, the additional units may cause a change if the combined average, which includes $\overline{x}_B$ and $\overrightarrow{x}_B$, will be above 49 (since the average quality of alternative B is 49). Whenever $\mu_B > 49$, the probability of a new example to have a value higher than 49 is greater than the probability of it to have a lower value. Thus, as we collect more data about alternative B, the probability for a change increases, since more often the collected data will be higher than 49, and will push the combined average above 49. Whenever this value will be over 49, alternative B will win (since the average of alternative A is 49). But, if $\mu_B < 49$, then obtaining additional information units about B will more often cause the average of alternative B to remain less than the average of alternative A, so a change will occur more rarely. The global influence on *Fchange* depends on the probability for each value of μ_A and μ_B.

In the following table, we display the different values of $Fchange$. The values of $Fchange$ for $m_A = 0$, and $m_B = 0,..,5$ are presented in the second column. We see that the positive influence of enlarging m_B is higher than the negative one, i.e., as m_B increases, the probability of a change increases too. The third column presents the influence of enlarging m_A when $m_B = 0$. Again, the total positive effect is higher than the negative effect: as m_A increases, the probability of a change increases.

m	$Fchange$ $(m_A = 0, m_B = m)$	$Fchange$ $(m_A = m, m_B = 0)$
1	.41543	.41055
2	.43648	.42998
3	.44536	.43784
4	.45041	.44217
5	.45371	.44492

The next evaluation step is the calculation of the gains due to the additional information. We should consider all possible pairs μ_A, μ_B, and for each pair, calculate the expected benefits from m_A and m_B additional information units, and multiply this by the probability for μ_A, μ_B being the actual mean values.

The expected benefits from obtaining additional m_A units of information on A and m_B units on B, is $Fchange \cdot (\mu_B - \mu_A)$, since obtaining these additional units can cause the final decision to be changed from choosing A to choosing B with a probability of $Fchange$, and in this case, the expected utility of choosing B instead of A is $\mu_B - \mu_A$. In order to compute the probability for each particular assignment of μ_A and μ_B, we evaluate their posterior distribution, using their prior distributions, and the collected information $n_A, n_B, \overline{x}_A$ and $\overline{x}_A$.

The following lemma provides the expected benefits from obtaining additional information about each alternative, by considering all possible distributions of both alternatives, i.e., by considering each possible pair μ_A, μ_B. In this lemma, we use the following notations. Denote by $f(\mu_i|\zeta_i, \tau_i, \sigma_i, n_i, \sigma_i, \overline{x}_i)$ the probability for alternative i to have a mean of μ_i, given the prior belief $\mu_i \sim N(\zeta_i, \tau_i)$, given σ_i, and given a sample with n_i units, and average $\overline{x}_i$. Denote by $f(u_i = \mu_i | u_i \sim N(\mu(i), \sigma(i))$ the probability of the variable u_i to have the value μ_i, given that u_i is normally distributed, with mean $\mu(i)$ and standard deviation $\sigma(i)$.

Lemma 2. Calculating the benefits *The benefits due to obtaining m_A units of information about A and m_B units of information about B are as follows*

$$
\begin{aligned}
&benefits(m_A, m_B) = \\
&\int\int Fchange(\mu_A, \mu_B, \sigma_A, \sigma_B, \overline{x}_A, n_A, m_A, \overline{x}_B, n_B, m_B)\cdot \\
&(\mu_B - \mu_A) \cdot f(\mu_A|\zeta_A, \tau_A, \sigma_A, n_A, \overline{x}_A) \cdot f(\mu_B|\zeta_B, \tau_B, \sigma_B, n_B, \overline{x}_B) d\mu_A d\mu_B = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Fchange(\mu_A, \mu_B, \sigma_A, \sigma_B, \overline{x}_A, n_A, m_A, \overline{x}_B, \\
&n_B, m_B) \cdot (\mu_B - \mu_A) \cdot f(u_A = \mu_A | u_A \sim N(\mu(A), \sigma(A))) \\
&\quad \cdot f(u_B = \mu_B | u_B \sim N(\mu(B), \sigma(B))) d\mu_A d\mu_B
\end{aligned}
$$

where

$$\mu(i) = \frac{\sigma_i^2\zeta_i + n_i\tau_i^2\overline{x_i}}{\sigma_i^2 + n_i\tau_i^2}, \sigma(i)^2 = \frac{\sigma_i^2\tau_i^2}{\sigma_i^2 + n_i\tau_i^2}. \quad (1)$$

Sketch of proof: We should consider all the possible pairs of μ_A, μ_B, and for each pair, to evaluate the expected value of additional m_A and m_B units of information, given the pair μ_A, μ_B, multiplied by the probability for this pair. The expected value of the additional units is equal to the probability of a change, multiplied by the gains $\mu_B - \mu_A$ due to a change. The probability for each value μ_i is evaluated, by calculating its posterior distribution, by the Bayesian estimation [13]. □

Finally, we will also consider the various costs involved in obtaining $m_A + m_B$ units of information and present the agent's utility. Recall that there is a discount factor of $0 < \delta \leq 1$ on the utility of the agent, and a direct cost of c for each unit. Given δ and c, the expected utility of the agent from obtaining a utility of x_i from the alternative chosen, after collecting m units of information, is $(x_i)\delta^m - \sum_{i=0}^{m-1} c \cdot \delta^i$. The explicit value is $x_i\delta^m - c(\frac{\delta^m - 1}{\delta - 1})$. Considering all possible values of x_i, we obtain the following theorem. The proof is immediate from lemma 2 and from the cost of each data unit.

Theorem 3. *The expected utility from obtaining m_A and m_B additional information units about alternative A and B, respectively, is* $utility(m_A, m_B) = benefits(m_A, m_B) \cdot \delta^{m_A + m_B} - c\frac{\delta^{m_A + m_B} - 1}{\delta - 1}$.

If there is no discount factor, then $\delta = 1$. In this case, the utility of the agent is $benefits(m_A, m_B) - cm$. If there are no direct costs, then $c = 0$, and in this case, the utility of the agent is $benefits(m_A, m_B)\delta^m$. Based on the above theorem, we proceed by presenting an algorithm that computes the optimal number of units to obtain for each alternative. The agent's goal is to find the pair m_A and m_B, that yields the highest value of $utility(m_A, m_B)$, i.e., we should find the values M_A and M_B such that, $(M_A, M_B) = argmax_{m_A, m_B} utility(m_A, m_B)$.

We suggest to consider all possible combinations of information units about A and B, and choose the optimal combination. Since there are only two alternatives, this process is polynomial. We proceed by implementing the above process on a particular problem, and then we generalize the problem and consider $k > 2$ alternatives to choose from. In the following suppliers example, we show how the values of m_A and m_B should be determined.

Example 3. We return to the situation of Examples 1-2. One may hypothesis that as m_A and m_B increase, the benefits due to the additional information increase, since the additional information results in more knowledge about the alternative, and a better decision. However, an increase of m_A and m_B also has a negative effect, due to the cost of time. The following table, that describes $Utility(m_A, m_B)$ for each pair m_A, m_B, shows that there is a peak level for the optimal values of m_A and m_B, and an additional increment of them can only reduce the expected utility of the agents.

m_B / m_A	0	1	2	3	4	5
0	0	2.0241	2.5998	2.7867	2.795	2.711
1	1.7951	2.6685	2.9528	2.9959	2.920	2.784
2	2.3551	2.8884	**3.0344**	3.0055	2.889	2.729
3	2.5547	2.9113	2.9794	2.9118	2.776	2.609
4	2.5831	2.8319	2.8529	2.7647	2.622	2.455
5	2.5205	2.6982	2.6899	2.5915	2.449	2.287

In the above example, the optimal combination of m_A and m_B is $m_A = 2, m_B = 2$, which gives a utility of 3.0344. As the number of units about alternative A or B decreases or increases, the utility decreases. Note that if there are at most 5 available information units for both alternatives, i.e., when $M = 6$, then we should consider only the part of the table from the main diagonal and up, while $m_A + m_B < 6$, but the same result will be obtained in our case, since the pair $m_A = 2, m_B = 2$ is valid also in the case of a total of $M - 1$ units.

There are cases where the optimal M_A and M_B can be found by analytically finding the optimal value of $utility(m_A, m_B)$, as done by [3] for a simpler case. This calculation can be performed in simple situations. For example, if μ_A and μ_B have a uniform distribution. But, in more complex situations, analytical derivation of the optimal values of M_A and M_B becomes hard or even impossible.

In the above example, we could simply use a greedy algorithm in order to find the optimal allocation of samples. The algorithm should choose in each step the most beneficial sample to add, and it should stop when there is no additional sample that increases the expected utility. However, in the following lemma, we show that the greedy algorithm is not optimal. The reason being that there may be situations where obtaining one unit of information about a particular alternative is not worthwhile, but obtaining two and more may be worthwhile to the agent, as stated in the following lemma.

Lemma 4. *(1) There are situations where obtaining one unit of information about a given alternative is not beneficial, but obtaining more than one is beneficial. (2) There are situations where obtaining information about one alternative is not beneficial, while obtaining information about two or more is beneficial.*

Proof. Both claims will be proven by examples. In order to show claim (1), consider a case where prior beliefs of the user are: $\mu_A \sim N(\zeta_A = 10, \tau_A = 27), \mu_B \sim N(\zeta_B = 10, \tau_B = 76)$, and $\sigma_A = 34$, $\sigma_B = 50$. Suppose that the discount factor is $\delta := 0.968$, and the cost is $c = 0$. Suppose that there are ten collected information units about alternative A and only two units about B, with the results $\overline{x}_A = 9.5, \overline{x}_B = 12.9$. In this case, $utility(m_A = 1, m_B = 0) = -.02551$, while $utility(m_A = 2, m_B = 0) = .63431$. Thus, although it is not worthwhile to obtain one unit of information about alternative A, it may be worthwhile to obtain 2 units of information about this alternative.

We proceed by an example to show claim (2). Suppose the prior beliefs of the agent are $\mu_A \sim N(\zeta_A = 27, \tau_A = 72), \mu_B \sim N(\zeta_B = 27, \tau_B = 72)$. Suppose also that $x_A \sim N(\mu_A, 89)$ and $x_B \sim N(mu_B, 67)$. The discount factor is $\delta := 0.991$, and the cost of time is $c = 0.03$. Eight units of information were collected about each alternative, A and B, with the results $\overline{x}_A = 31.05$ and $\overline{x}_B = 34.83$. In this case, $utility(m_A = 1, m_B = 0) = -0.85096$, $utility(m_A = 0, m_B = 1) = -0.62808$, but $utility(m_A = 1, m_B = 1) = 0.86613$. This means that although it is not worthwhile to obtain one unit of information about alternative A or B, it becomes beneficial to obtain one unit about each of these two alternatives. □

5 Choosing from Multiple Alternatives

We proceed by considering a general case, where there are $k > 2$ alternatives, and the agent can obtain up to $M-1$ units of information about each of them. Thus, there are M^k possible combinations of information units that can be obtained by the agent, and if there are a total of $M-1$ units to be used, then there are $(M+k)!/((M-1)! \cdot (k+1)!)$ possible combinations (Each sample can be allocated to a particular alternative or not be allocated at all). In the following, we describe the agent utility function given a combination of information units.

In order to calculate the value of each combination of information units, we should know the prior beliefs ζ_i, τ_i about each alternative i, and also the value of $\overline{x}_i$ for the k_i information units that were already obtained. In the case of two alternatives, we presented the expected utility to be the additional utility due to replacing alternative A with alternative B. In this section, for simplicity reasons, we calculate the absolute utility from each alternative, without comparing it to the previous winner. Of course, given the utility from each alternative, the new winner could be identified.

Let $comb = (m_1, m_2, ..., m_k)$ denote a combination of m_i units of information about alternative i, and $totalM$ the total number of information units obtained about all the alternatives, i.e., $totalM = \sum_{i=1}^{k} m_i$. Also $\overrightarrow{x}_i$ will denote the average of the m_i information units about alternative i. Finally, let $R_{mi} = m_i/(m_i + n_i)$ and $R_{ni} = n_i/(m_i + n_i)$. The following theorem expresses the utility of the agent, given the information units combination $comb$, the prior beliefs about each alternative and the prior average value of each of them.

Theorem 5. *Suppose that given $\overline{x_i}$ and n_i, alternative 1 will be chosen. The utility from obtaining m_i additional information units about each alternative i is*

$$Util(comb) = \sum_{i=2}^{k} \int_{-\infty}^{+\infty} f(i,\mu_i) \cdot \int_{-\infty}^{+\infty} f(1,\mu_1) \cdot (\mu_i - \mu_1) \cdot$$
$$f(wins_i|\mu_i, \mu_1, comb) d\mu_i d\mu 1 \cdot \delta^{totalM} - c\frac{\delta^{totalM-1}}{\delta-1},$$

where

$$f(wins_i|\mu_i, \mu_1, comb) = \int_{\overrightarrow{x}_i=-\infty}^{\infty} f(\overrightarrow{x}_i|\mu_i) \cdot$$
$$\left(\int_{-\infty}^{(R_{mi}\cdot\overrightarrow{x_i^{?}} + R_{ni}\cdot\overline{x_i} - R_{n1}\cdot\overline{x_1})/R_{m1}} f(\overrightarrow{x}_1|\mu_1) d\overrightarrow{x}_1 \cdot\right.$$
$$\prod_{j>1, j\neq i} F\left(\frac{(R_{mi}\cdot\overrightarrow{x_i^{?}} + R_{ni}\cdot\overline{x_i} - R_{nj}\cdot\overline{x_j})/R_{mj} - \mu(j)}{\sqrt{\sigma_j^2 + \sigma(j)^2}}\right) d\overrightarrow{x}_i$$

where $f(i,\mu_i) = f(u_i = \mu_i | u_i \sim N(\mu(i), \sigma(i))$ as defined in equation 1, $f(\overrightarrow{x}_i|\mu_i) = f(x = \overrightarrow{x}_i | x \sim N(\mu_i, \sigma_i/\sqrt{m_i}))$ and finally, $F(z)$ is the standard normal density of the random variable z.

Sketch of proof: Given *comb*, for each alternative i we consider the probability for it to win, and the expected utility in this case, and we also consider the cost of obtaining the combination *comb* of information units. Thus, we run all the values of μ_i and μ_1, which is the default alternative to be chosen. For each pair, we calculate the probability for this pair, and the utility from the additional information, which is $\mu_i - \mu_1$ (since we changed the decision from alternative 1 to alternative i). We multiply the above by $f(wins_i|\mu_i, \mu_1, comb)$, the probability for alternative i to win.

In order to evaluate $f(wins_i|\mu_i, \mu_1, comb)$, we run each possible value of $\overrightarrow{x}_i$, and calculate its probability to win, i.e., we require that for each $j \neq i$, $(n_j \cdot \overline{x} + m_j \cdot \overrightarrow{x}_j)/(m_j + n_j) < (n_i \cdot \overline{x} + m_i \cdot \overrightarrow{x}_i)/(m_i + n_i)$. So, the maximum value of $\overrightarrow{x_j}$ should be $(R_{ni} \cdot \overline{x} + R_{mi} \cdot \overrightarrow{x}_i - R_{nj} \cdot \overline{x})/R_{mj}$.

The distribution of the average of alternative $j > 1, j \neq i$ depends on μ_j. Denote $M' = (R_{ni} \cdot \overline{x} + R_{mi} \cdot \overrightarrow{x}_i - R_{nj} \cdot \overline{x})/R_{mj}$. Then, calculate $\int_{-\infty}^{\infty} f(j, \mu_j) \cdot \int_{-\infty}^{M'} f(\overrightarrow{x}|\mu_j)$. Manipulating the above formula, and using integration rules of the normal distribution from [14], we reveal that the above formula is equal to $F(\frac{M' - \mu(j)}{\sqrt{\sigma_j^2 + \sigma(j)^2}})$. Finally, we consider the cost of obtaining the additional information: we multiply the expected utility by δ^{totalM}, and subtract $c\frac{\delta^{totalM-1}}{\delta - 1}$. $\square$

Given the ability to evaluate the expected utility from a particular combination of additional information units, we are also able to calculate a beneficial combination, in order to maximize its expected utility. There are two methods for finding the optimal combination. First, this can be done analytically, by finding the derivation of the expected utility according to the $m_1, .., m_k$, and finding the values of $m_1, .., m_k$ for which the derivation is equal to 0. This method was performed by Dunnett [3] for the case of equal standard deviation σ for the different populations, and equal sample sizes k. In cases of multiple $m_1, .., m_k$ and $\sigma_1, .., \sigma_k$, the calculation becomes very complex, or even analytically impossible. In such cases, we consider different combinations of $m_1, .., m_k$, and choose the combination that maximizes $Util$.

We can suggest an optimal algorithm that considers all possible pairs of $m_1..m_k$, but this algorithm is clearly exponential in the number of alternatives. Different heuristics for the choosing problem can be used such as greedy-myopic heuristic [16], or a local search technique. In future work we intend to compare the different heuristics by simulations, and to compare their results with the optimal solution.

6 Conclusion

In this paper, we consider the problem faced by an agent that has to choose from alternatives, and is able to acquire additional information about them. We consider situations where the agent should decide in advance how much information it would like to obtain about each alternative, before it obtains any answer. We describe the expected utility of the agent due to the acquired information units, and we provide an optimal and polynomial decision procedure for the case of choosing between two alternatives, and demonstrate it with a particular example. We proceed by considering the case of choosing from multiple alternatives, and we describe the expected utility of the agent given this case. We suggest an optimal algorithm for this case, which has an exponential complexity, and suggest how sub-optimal solutions can be found.

References

1. D. A. Berry and B. Fristedt. *Bandit Problems: Sequential Allocation of Experiments.* Chapman and Hall, London, UK, 1985.
2. T. L. Dean and M. P. Wellman. *Planning and Control.* Morgan Kaufman, Publishers, California, 1991.
3. C. W. Dunnett. On selecting the largest of k normal population means. *Journal of the Royal Statistical Society. Series B (Methodological)*, 22 (1):1–40, 1960.
4. C. Eric, K. Hann, and I. Hitt. The nature of competition in electronic markets: An empirical investigation of online travel agent offerings. WP, The Wharton School of the Univ. of Pennsylvania, 1998.
5. J. C. Gittins. *Multi-armed Bandit Allocation Indices.* John Wiley & Sons, 1989.
6. J. Grass and S. Zilberstein. A value-driven system for autonomous information gathering. *Journal of Intelligent Information Systems*, 14:5–27, 2000.
7. D. E. Heckerman, E. J. Horvitz, and B. Middleton. An approximate nonmyopic computation for value of information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:292–298, 1993.
8. R. A. Howard. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2:22–26, 1966.
9. L. Joseph and D. B. Wolfson. Interval-based versus decision theoretic criteria for the choice of sample size. *Statistician*, 46 (2):145–149, 1997.
10. V. Lesser, B. Horling, F. Klassner, A. Raja, T. Wagner, and S. Zhang. Big: An agent for resource-bounded information gathering and decision making. *Artificial Intelligence*, 118:197–244, 2000.
11. V. Lesser, B. Horling, A. Raja, T. Wagner, and S. X. Zhang. Sophisticated information gathering in a marketplace of information providers. In *Proceedings of IEEE Internet Computing, Agents on the Net*, volume 4 (2), pages 49–58, 2000.
12. D. V. Lindley. The choice of sample size. *The Statistician*, 46:129–138, 1997.
13. I. Miller and J. E. Freund. *Probability and Statistics for Engineers.* Premtoce-Hall, Inc., 1985.
14. J. K. Patel and C. B. Read. *Handbook of the normal distribution.* Marcel Dekker, Inc., 1996.
15. H. Pezeshk and J. Gittins. Sample Size Determination in Clinical Trials. *Student*, 3 (1):19–26, 1999.
16. K. L. Poh and E. Horvitz. Reasoning about the value of decision model refinement: Methods and application. In *UAI-99*, pages 174–182, Washington DC, 1993.
17. S. Russel and P. Norvig. *Artificial Intelligence: a Modern Approach.* Prentice-Hall Inc., 1996.
18. C. Tseng and P. J. Gmytrasiewicz. Time sensitive sequential myopic information gathering. In *Proc. of HICSS-32*, Maui, Hawaii, 1999.