From: AAAI Technical Report SS-94-06. Compilation copyright © 1994, AAAI (www.aaai.org). All rights reserved.

A Statistical Approach to Adaptive Problem–Solving for Large–Scale Scheduling and Resource Allocation Problems

Steve Chien*, Jonathan Gratch+, and Michael Burl*

*Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109-8099
{chien,burl}@aig.jpl.nasa.gov

+Beckman Institute
University of Illinois
405 N. Mathews Av., Urbana, IL 61801
gratch@cs.uiuc.edu

Abstract

Our work focuses upon development of techniques for choosing among a set of alternatives in the presence of incomplete information and varying costs of acquiring information. In our approach, we model the cost and utility of various alternatives using parameterized statistical models. By applying techniques from an area of statistics called parameter estimation, statistical models can be inferred from data regarding the utility and information cost of each of the various alternatives. These statistical models can then be used to estimate the utility and cost of acquiring additional information and the utility of selecting specific alternatives from the possible choices at hand. We apply these techniques to adaptive problem–solving, a technique in which a system automatically tunes various control parameters on a performance element to improve performance in a given domain. We present empirical results comparing the effectiveness of these techniques on speedup learning from a real–world NASA scheduling domain and schedule quality data from the same real–world NASA scheduling domain.

1 INTRODUCTION

In machine learning and basic decision making in AI, a system must often reason about alternative courses of action in the absence of perfect information. In many cases, acquiring additional information may be possible, but has associated with it some cost. A central problem of interest to AI researchers, and statisticians is that of designing strategies to balance the cost of acquiring additional information against the expected utility of the information to be acquired. For example, when choosing among a set of actions, one could gather further information which might help make your decision (at some cost), or one might simply make a decision. If one decides to gather further information, which information would be the most useful? When one wishes some sort of statistical guarantees on the (local) optimality of the choice and/or the guarantee of rationality, a statistical decision theoretic framework is useful. To summarize, this problem of decision–making with incomplete information and information costs can be analyzed in two parts:

A1: How much information is enough? At what point do we have adequate information to select one of the alternatives?

A2: If one wishes to acquire more information, which information will allow us to make the best possible decision at hand while minimizing information costs?

Possible solutions to this decision–making quandary depend on the context in which the decision is being made. This paper focuses upon the decision–making problems involved in adaptive problem–solving. Adaptive problem–solving occurs when a system has a number of control points which affect its performance over a distribution of problems. When the system solves a problem with a given set of settings for control points, it produces a result which has a corresponding utility. The goal of adaptive problem–solving is: given a problem distribution, find the setting of control points which maximizes the expected utility of the result of applying the system to problems in the distribution.

More rigorously, the adaptive problem–solving problem can be describes as follows. Given a flexible performance element PE with control points $CP_1 ... CP_n$, where each control point CP_i corresponds to a particular control decision and for which

there is a set of alternative decision methods $M_{i,1}...M_{i,k}$,[1] a control strategy is a selection of a specific method for every control point (e.g., STRAT = $<M_{1,3},M_{2,6},M_{3,1},...>$). A control strategy determines the overall behavior of the scheduler. It may effect properties like computational efficiency or the quality of its solutions. Let PE(STRAT) be the problem solver operating under a particular control strategy. The function U(PE(STRAT), d) is a real valued utility function that is a measure of the goodness of the behavior of the scheduler over problem d. The goal of learning can be expressed as: given a problem distribution D, find STRAT so as to maximize the expected utility of PE. Expected utility is defined formally as:

$$\sum_{d \in D} U(PE(STRAT), d) \times probability(d)$$

For example, in a planning system such as PRODIGY [Minton88], when planning to achieve a goal, control points would be: how to select an operator to use to achieve the goal; how to select variable bindings to instantiate the operator; etc. A method for the operator choice control point might be a set of control rules to determine which operators to use to achieve various goals plus a default operator choice method. A strategy would be a set of control rules and default methods for every control point (e.g., one for operator choice, one for binding choice, etc.). Utility might be defined as a function of the time to construct a plan, cost to execute the plan, or some overall measure of the quality of the plan produced.

Given the context of adaptive problem–solving, the question A1 from above "how much information is enough" can be further elaborated. In adaptive problem–solving, the expected utility can be interpreted as average utility from the performance element on a per problem basis. This means that "how much information is enough" generally depends upon the amortization period for learned knowledge, the relative cost of acquiring additional information and interactions between successive choices. These issues are the focus of other work [Gratch and DeJong 93], and are not directly addressed by the work described in this abstract.

Question A2 from above is "find which information will allow us to make the best possible decision at hand while minimizing information costs". This question can be cast in two ways: 1) for a given amount of resources, which information should I get to maximize the expected utility of of the outcome my decision; and 2) given that I want to have a certain confidence level in the goodness of my decision (e.g., local optimality), how can I achieve this while expending the least resources possible.

This abstract focuses on question A2, namely: "find which information will allow us to make the best possible decision at hand while minimizing information costs". Specifically, our approach draws upon techniques from statistics in the of area of parametric statistical models to model the uncertainty in utility and information cost estimates. In parametric statistical models, one presumes that data is distributed according to some form of model (e.g., the normal distribution, the poisson distribution, etc.). This distribution can be described in terms of a fixed set of parameters (such as mean, variance, etc.). If one can infer the relevant parameters for the distribution underlying the data (so–called parameter estimation), then because the uncertainty in utility estimates is explicitly modelled in the statistics, three types of questions regarding utility can be answered:

B1: which alternative has the highest expected utility;

B2: how certain are we of this ranking of the alternatives; and

B3: how much is this uncertainty likely to change if we acquire additional information.

Correspondingly, the same approach can be used in modelling the cost distribution. Hence the expected cost of acquiring additional information can also be estimated.

Of course, the accuracy of all of these estimates is dependent upon the goodness of fit of the parametric model used to model the uncertainty and cost estimates. Generally, we use the normal (gaussian) distribution model as our parametric model. Fortunately, the normal distribution has the strong property that it is a good approximation to many forms of distributions for reasonably large sample sizes (due to the Central Limit Theorem).

1. Note that a method may consist of smaller elements so that a method may be a set of control rules or a combination of heuristics. Note also that a method may also involve real–valued parameters. Hence, the number of methods for a control point may be infinite, and there may be an infinite number of strategies.

Thus, using techniques from parameter estimation and analysis of the data, we can answer questions B1, B2, & B3 from above regarding utility and estimate expected cost of information. The remaining task is to apply decision theoretic methods to determine the answers to questions A1 and A2 from this information. Working towards this goal we have developed two general approaches to addressing this problem: dominance–indifference and expected loss.

The first set of methods, called interval–based methods, involve quantifying the uncertainty in competing hypotheses by computing the statistical confidence that one hypothesis is better than another hypothesis. Once this confidence has been computed, the system attempts to allocate examples to efficiently show that one hypothesis dominates all the other hypotheses with some specified confidence. These methods also rely upon an indifference parameter, which allows it to show that the preferred hypothesis is quite similar in expected performance to another hypothesis with some confidence, thus making the preferred hypothesis is acceptable.

The second set of methods uses the decision theoretic concept of expected loss [Russell & Wefald 89, Russell & Wefald 93], which measures the probability that a less preferable decision is made weighted by the lost utility with respect to the alternative choice. More specifically, the expected loss of utility from adopting option H_i over option H_j with corresponding utility distributions U_i and U_j to be the integral of $u_i \in U_i$, $u_j \in U_j$, over all the regions where $u_j > u_i$, of $P_{u_i u_j}(u_i, u_j)(u_j - u_i)$. In the expected loss approach, the system acquires information until the expected loss is reduced below some specified threshold. This approach has the added benefit of not attempting to distinguish among two hypotheses with similar means and low variances (e.g., it recognizes indifference without a separate indifference parameter).

For both the interval–based and expected loss approaches, when selecting a best hypothesis, one must base this selection upon comparisons of the utility of the "best" hypothesis to the other possible hypotheses. Because there are multiple comparisons, the estimate for the overall error (or confidence) in a conclusion of selection of a best hypothesis, is based upon multiple smaller conclusions. For example, if we wish to show that H3 is the best choice among H1, H2, H3, H4, and H5, in the interval–based approach, we might show that H1 and H3 are indifferent, that H3 dominates H2, H3 dominates H4, and H3 dominates H5. Thus if we wish a confidence level of 95%, with a straight sum error model and equal allocation of error, each of the individual hypotheses would need a 98.75% confidence level (since 4 x 1.25% = 5%).

The exact form of the relationship depends upon the particular error model used. However, sampling additional examples to compare relative utilities of various alternatives can have varying effect upon the overall error estimate, based upon parameters of the particular distribution. Also, selecting an example from different distributions can have widely varying cost. Thus, in many cases, it may be desirable to allocate the error estimates unequally. While we have implemented default strategies of equal allocation of errors, we have also developed algorithms which allocate the error unequally. In this approach, the system estimates the marginal benefit and marginal cost of sampling to extract another data point to compare two competing hypotheses. The marginal benefit is estimated by computing the decrease in the error estimate (or increase in confidence estimate) due to acquiring another sample presuming that the statistical parameters remain constant (e.g., in the case of the normal distribution, that the sample mean and variance do not change). The marginal cost is estimated using estimated parameters on the cost distribution for the relevant hypotheses. The system then allocates additional examples preferring the highest ratio of marginal benefit to marginal cost.

For the expected loss approach, a comparable combination of expected losses is used. In this approach, the expected loss of choosing H_i is the sum of the pair-wise expected losses from choosing Hi over each of the other H_j's. Again, depending upon the exact parameters of the utility distributions, allocating examples to each of the pair-wise expected loss computations will have varying effects on the sum of expected losses. Additionally, exactly as with the interval–based case, the cost of sampling also varies. Thus, allocating examples to the pair-wise computations unequally can be advantageous. Consequently, we have also implemented an expected loss algorithm which uses the estimated marginal cost and estimated marginal benefit of sampling to allocate sampling resources. Thus, in all, there are four new algorithms, interval–based

equal error allocation (STOP1), interval–based unequal error allocation (STOP2), expected loss equal allocation (EL1), and expected loss unequal allocation (EL2)[2].

2 EMPIRICAL PERFORMANCE EVALUATION

We now turn to an empirical evaluation of the hypothesis selection techniques. This evaluation lends support to the techniques by: (1) using synthetic data to demonstrate that the techniques perform as predicted when assumptions are met and (2) using a real-world hypothesis selection problem to demonstrate the robustness of the approaches. As the interval-based and expected loss approaches use different parameters which are incomparable, we first test the interval–based and expected loss approaches separately and then compare them head–to–head in a real–world comprehensive test.

2.1 SYNTHETIC DATA

Synthetic data is used to show that the techniques perform as expected when the underlying assumptions are valid. For interval-based approaches we show that the technique will choose the best hypotheses, or one ε–close to the best, with the requested probability. For the expected loss approach we show that the technique will exhibit no more that the requested level of expected loss.

The statistical ranking and selection literature uses a standard class of synthetic problems for evaluating the statistical error of hypothesis evaluation techniques called the *least favorable configuration of the population means* [Bechhofer54]. This is a parameter configuration that is most likely to cause a technique to choose a wrong hypothesis (one that is not ε–close) and thus provides the most severe test of the technique's abilities. Under this configuration, $k-1$ of the hypotheses have identical expected utilities, μ, and the remaining hypothesis has expected utility $\mu+\varepsilon$. The last hypothesis has the highest expected utility and should be chosen by the technique. The costs and variances of all hypotheses are equal.

We test each technique on the least favorable configuration under a variety of control parameter settings. The least favorable configuration becomes more difficult (requires more examples) as the confidence γ^*, the number of hypotheses k, or common utility variance σ^2, increases. It becomes easier as the indifference interval ε, increases. In the standard methodology a technique is evaluated varying values for k, γ^*, and σ/ε. The last term combines the variance and indifference interval size into a single quantity which as it increases makes the problem more difficult. For our experiments, n_0 is set to seven, μ is fifty, σ^2 is sixty-four, and all other parameters are varied as indicated in the results. All experimental trials are repeated 5000 times to obtain statistically reliable results.[3] The results from these experiments, in Table 1 below, show that as the requested confidence increases, the STOP1 algorithm correctly takes more training examples to ensure that the requested accuracy level is achieved.

The expected loss approach EL1 was also evaluated in the least favorable configuration to test the algorithms ability to bound expected loss. EL1 was tested on various loss thresholds, H^*, over this problem. For this evaluation, μ is fifty, all hypotheses share a common utility variance of sixty–four, and ε is two, k = 3,5,10, H* = 1.0, 0.75, 0.5, 0.25. The sample size results and observed loss values are summarized below in Table 1. The results illustrate that as the loss threshold is lowered EL1 takes more training examples to ensure the expected loss remains below the threshold. Again, all trials are repeated 5000 times.

γ^*	STOP1		H*	EL1
0.75	1253 (0.85)		1.0	307 (0.67)
0.90	1976 (0.93)		0.75	332 (0.67)
0.95	2596 (0.95)		0.5	399 (0.17)
			0.25	497 (0.07)

Table 1: Least Favorable Configuration Results

Both systems satisfied their statistical requirements. In each configuration STOP1 selected the correct hypothesis with at least as much confidence as requested. The expected loss exhibited by EL1 was

2. For details of these algorithms see [Chien et al 94].

3. Another important question with the interval-based approach is its efficiency when all hypotheses are indifferent to each other. Evaluations in this configuration give comparable efficiency to the least favorable configuration, but space precludes presentation of those results.

no greater than the loss threshold. As expected the number of examples required increased as the acceptable error level or loss threshold decreased. The number of examples required by the two techniques should not be compared directly as each algorithm is solving a different task. Later we will discuss the relationship between the efficiency of the approaches.

2.2 NASA SCHEDULING DATA

The test of real-world applicability is based on data drawn from an actual NASA scheduling application. This data provides a strong test of the applicability of the techniques. All of the statistical techniques make some form of normality assumption. However the data in this application is highly non-normal – in fact most of the distributions are bi-modal. This characteristic provides a rather severe test of the robustness of the approaches.

Our particular application has been adaptive learning of heuristics for a lagrangian relaxation [Fisher81] scheduler developed by Colin Bell called LR–26 [Bell93] to schedule communications between NASA/JPL antennas and low earth orbiting satellites. LR–26 formulates the scheduling problem as an integer linear programming problem and a typical problem in this domain has approximately 650 variables and 1300 constraints.

The goal of these evaluations was to choose a heuristic that solved scheduling problems quickly on average. This is easily seen as a hypothesis evaluation problem. Each of the heuristics corresponds to a hypothesis. The cost of evaluating a hypothesis over a training example is the cost of solving the scheduling problem with the given heuristic. The utility of the training example is simply the negation of its cost. In that way, choosing a hypothesis with maximal expected utility corresponds to choosing a scheduling heuristic with minimal average cost.

Application of the COMPOSER method of adaptive problem–solving [Gratch93] to this domain has produced very promising results. For one version of the problem, the goal is to learn heuristics to allow for faster construction of schedules. For this application [Gratch et al. 93a, Gratch & Chien 93], learned heuristics produced the results of: 50% reduction in CPU time to construct a schedule for solvable problems, and a 15% increase in the number of solvable problems within resource bounds. For a second version of the problem, the goal is to learn heuristics to choose goal constraints to relax when it appears that the problems posed to the scheduler is unsolvable. In this application, the utility is negative the number of constraints relaxed in the final solution. Preliminary tests showed a range from –681 to –35 in average utility in the control space, with an average utility of –137. Over 6 trials, COMPOSER picked the best strategy in every trial. We are currently in the process of enriching the heuristic space for this application.

Currently, we are using data from this scheduling application to further test our adaptive problem–solving techniques. The scheduling application involved several hypothesis evaluation problems, four of which we use in this evaluation. Each problem corresponds to a comparison of some set of scheduling heuristics, and contains data on the heuristics' performance over about one thousand scheduling problems. An experimental trial consists of executing a technique over one of these data sets. Each time a training example is to be processed, some problem is drawn randomly from the data set with replacement. The actual utility and cost values associated with this scheduling problem is then used. As in the synthetic data, each experimental trial is repeated 5000 times and all reported results are the average of these trials.

We ran the interval–based hypothesis selection method (STOP1), the cost sensitive interval–based method (STOP2), the COMPOSER approach, and an existing statistical technique by Turnbull and Weiss [Turnbull & Weiss 87] over the four scheduling data sets. In each case the confidence level was set at 95%, and n_0 set to fifteen. These results, shown in Table 2, show that the interval–based techniques significantly outperformed the Turnbull and COMPOSER techniques and that STOP2 slightly outperformed STOP2. We also ran the expected loss techniques EL1 and EL2 (cost sensitive) over the four scheduling data sets. In each case the loss threshold was set at three and n_0 was fifteen. Table 3 summarizes the results along with the number of hypotheses and the relative difficulty (σ/ε) of each data set. These results in Table 3 show that the expected loss approaches bounded the expected loss as predicted.

2.3 Comparing STOP1 and EL1

It is difficult to directly compare the performance of STOP1 and EL1 as the algorithms are accomplishing different tasks. STOP1 is attempting to identify

	Parameters			STOP1	STOP2	TURNBULL	COMPOSER
	k	γ^*	σ/ε				
D1	3	0.95	34	908 (1.00)	648 (1.00)	26,691 (1.00)	78 (1.00)
D2	2	0.95	34	74 (1.00)	76 (1.00)	13,066 (1.00)	346 (1.00)
D3	7	0.95	14	2,371 (0.94)	2,153 (0.93)	94,308 (1.00)	2,456 (0.97)
D4	7	0.95	11	7,972 (0.96)	7,621 (0.94)	87,357 (1.00)	21,312 (0.89)

Table 2: Estimated expected total number of observations for scheduling data. Achieved probability of a correct selection is shown in parenthesis.

	Parameters		EL1		EL2	
	k	H^*	Observations	Loss	Observations	Loss
D1	3	3.0	78	0.1	49	1.0
D2	2	3.0	30	1.8	30	1.8
D3	7	3.0	335	3.0	177	3.9
D4	7	3.0	735	1.7	283	2.2

Table 3: Estimated expected total number of observations and expected loss of an incorrect selection for the scheduling data.

a nearly optimal hypothesis with high confidence while EL1 is attempting to minimize the cost of a mistaken selection. If the goal of the task is to identify the best hypothesis then clearly STOP1 should be use. If the goal is to simply improve expected utility as much as possible, either could be used and it is unclear which is to be preferred. This section attempts to assess this question on the NASA scheduling domain. We run each algorithm under a variety of parameter settings and compare the best performance of each algorithm.

To directly compare STOP1 and EL1, we ran a comprehensive adaptive scheduling test. In this test, STOP1 and EL1 all were given the task of optimizing several control parameters of an adaptive scheduler, with the goal of speeding up the schedule generation process. This task corresponds to the sequential solution of the problems D1 through D4 posed in the scheduling test (for more details on this application domain see [Gratch et al. 93]. In this test the STOP1 algorithm was run with confidence levels γ^*=0.75,0.90,0.95 and an indifference level of 1.0, and EL1 was run with loss bound L=5, 1, 0.5. The best settings found and averaged results resulting from 1000 runs are shown below in Table 4.

	Cost (CPU Days)	Examp.	Utility
STOP1 (0.75,1.0)	1.94	4199	17.1
EL1(0.5)	1.84	2630	17.3

Table 4: Direct Comparison of STOP1 and EL1

These results show that the two algorithms produce roughly comparable utilities, the difference in utilities is smaller than the indifference interval specified to STOP1. However, the STOP1 algorithm requires nearly twice the examples on average to achieve this utility. While it is difficult to generalize this result to other applications, this and other empirical results has suggested to us that the expected

loss approach is more efficient at the task of improving expected utility.

3 CONCLUSION

This paper has described techniques for choosing among a set of alternatives in the presence of incomplete information and varying costs of acquiring information. In our approach, we model the cost and utility of various alternatives using parameterized statistical models. By applying techniques from an area of statistics called parameter estimation, statistical models can be inferred from data regarding the utility and information cost of each of the various alternatives. These statistical models can then be used to estimate the utility and cost of acquiring additional information and the utility of selecting specific alternatives from the possible choices at hand. We apply these techniques to adaptive problem–solving, a technique in which a system automatically tunes various control parameters on a performance element to improve performance in a given domain. We present empirical results comparing the effectiveness of these techniques on two applications in a real–world domain: speedup learning from a real–world NASA scheduling domain and schedule quality data from the same real–world NASA scheduling domain.

Acknowledgements

Portions of this work were performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration and portions at the Beckman Institute, University of Illinois under National Science Foundation Grant NSF–IRI–92–09394.

References

[Bell93] C. E. Bell, "Scheduling Deep Space Network Data Transmissions: A Lagrangian Relaxation Approach," **Proceedings of the SPIE Conference on Applications of Artificial Intelligence 1993: Knowledge–based Systems in Aerospace and Industry**, Orlando, FL, April 1993, pp. 330–340.

[Fisher81] M. Fisher, "The Lagrangian Relaxation Method for Solving Integer Programming Problems," ***Management Science 27***, 1 (1981), pp. 1–18.

[Gratch93] J. Gratch, "COMPOSER: A Decision–theoretic Approach to Adaptive Problem Solving," Technical Report UIUCDCS–R–93–1806, Department of Computer Science, University of Illinois, Urbana, IL, May 1993.

[Gratch et al. 93a] J. Gratch, S. Chien, and G. DeJong, "Learning Search Control Knowledge for Deep Space Network Scheduling," **Proceedings of the Tenth International Conference on Machine Learning**, Amherst, MA, June 1993, pp. 135–142.

[Gratch et al. 93b] J. Gratch, S. Chien, and G. DeJong, "Learning Search Control Knowledge to Improve Schedule Quality," **Proceedings of the of the IJCAI93 Workshop on Production Planning, Scheduling, and Control**, Chamberry, France, August 1993.

[Chien et al. 94] S. Chien, J. Gratch, and M. Burl, "On the Efficient Allocation of Resources for Hypothesis Evaluation in Machine Learning: A Statistical Approach," Forthcoming JPL Technical Report, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 1994.

[Gratch & Chien 93] J. Gratch and S. Chien, "Learning Search Control Knowledge for the Deep Space Network Scheduling Problem: Extended Report and Guide to Software," Technical Report UIUCDCS–R–93–1801, Department of Computer Science, University of Illinois, Urbana, IL, January 1993.

[Gratch & DeJong 93] J. Gratch and G. DeJong, "Rational Learning: A Principled Approach to Balancing Learning and Action," Technical Report UIUCDCS–R–93–1801, Department of Computer Science, University of Illinois, Urbana, IL, April 1993.

[Russell & Wefald 89] S. Russell and E. Wefald, "On Optimal Game Tree Search using Rational Meta–Reasoning," **Proceedings of the Eleventh International Joint Conference on Artificial Intelligence**, Detroit, MI, August 1989, pp. 334–340.

[Russell & Wefald 91] S. Russell and E. Wefald, "Do the Right Thing: Studies in Limited Rationality," MIT Press, Cambridge, MA 1991.

[Turnbull & Weiss 87] A class of sequential procedures for k–sample problems concerning normal means with unknown unequal variances, in **Statistical Ranking and Selection, Three Decades of Development**, R. Haseeb (ed.), American Sciences Press, Columbus, OH, 1987, pp. 225–240.