Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)

Some Properties of Batch Value of Information in the Selection Problem (Extended Abstract*)

Shahaf S. Shperberg and Solomon Eyal Shimony
Dept. of Computer Science, Ben-Gurion University of the Negev
shperbsh@post.bgu.ac.il, shimony@cs.bgu.ac.il

Abstract

We examine theoretical properties of value of information (VOI) in the selection problem, and identify cases of submodularity and supermodularity. We use these properties to compute approximately optimal measurement batch policies, implemented on a "wine selection problem" example.

1 Introduction

Given a set of items of unknown utility (drawn from a known distribution), we need to select an item with as high a utility as possible. Measurements of item values prior to selection are allowed, at a known cost. The problem is to optimize the decision process of measurement and selection. The selection problem is intractable, but has numerous applications [Tolpin and Shimony, 2012; Radovilsky *et al.*, 2006; Radovilsky and Shimony, 2008]. We analyze cases where the value of information (VOI) is submodular, which guarantees that greedy algorithms achieve good approximate solutions.

The selection problem is applicable in meta-reasoning [Russell and Wefald, 1991b; 1991a; Hay *et al.*, 2012], as well as settings where the items to be selected are physical objects. Examples of the latter type are: oil exploration, finding a good set of parameters for setting up an industrial imaging system [Tolpin and Shimony, 2012], and selecting experiments to be performed [Azimi *et al.*, 2016].

A widely adopted scheme for selecting measurements (sensing actions in some contexts, or deliberation steps in meta-reasoning) uses value of information (VOI) [Russell and Wefald, 1991b; 1991a]. Optimizing value of information is intractable in general, thus both researchers and practitioners often use various forms of myopic VOI estimates [Russell and Wefald, 1991b; 1991a; Hay *et al.*, 2012] coupled with greedy search. Submodularity is an important property, because if the VOI is submodular, simple, greedy algorithms result in provably near-optimal policies [Krause and Guestrin, 2009; 2011; Papachristoudis and Fisher III, 2012]. However, as stated in [Krause and Guestrin, 2009], the VOI is not submodular in general [Tolpin and Shimony, 2012].

Specifically, the selection problem analyzed in this paper is as follows. We have a set of items $\mathcal{I}$, each of which has some unknown value (or utility). The item utilities are random variables with a known joint distribution. It is possible to perform measurements on an item, thereby obtaining information about its utility. Measurements have a cost, specified by a known cost function $\mathcal{C}$, which is usually an additive cost function. After performing measurements, the decision-maker selects one item. We assume a risk-neutral decision-maker, thus the decision maker always selects an item that has the highest expected utility given the observations. The problem is to find a policy of performing measurements such that the utility of the selected item minus the cost of measurements has a maximum expected value. In some settings [Tolpin and Shimony, 2012; Azimi *et al.*, 2016], a *measurement budget* is also specified, and a policy is considered valid only if this budget is not exceeded. Some *budgeted* applications [Azimi *et al.*, 2016] optimize just the expected value of the selection (not factoring in the measurement costs), subject to the budget constraint.

There are two common selection problem settings: batch, and online (also called sequential, or conditional). In the online setting, the decision-maker may decide on measurements based on information from previous measurements. In the batch setting, the decision-maker the decision-maker does not get to perform additional measurements after receiving observations. In this paper we consider only the batch setting, wherein the VOI is the expected value of the best item given the observations, minus the expected value of the best item according to the initial (prior) distribution. That is, before receiving the information, there is some item that has the best expected value, which we call the "current best" item α. After receiving the observations O, some other item $\beta(O)$ may have the highest expected value. The expected value of the difference $u_{\beta(O)} - u_\alpha$ is the value of information (VOI). Note that both the identity of the resulting best item β and its utility depend on O. In the batch setting with perfect observations, the distribution over the observed values is equal to the utility distribution of the respective item. Thus finding an optimal policy can be done by finding a set of measurements S to perform that has the highest expected VOI minus cost (also called the *net VOI*). In this paper, we consider mostly the case of perfect observations, i.e. where as a result of performing a measurement on an item, its precise utility value becomes known. In general, measurements can generate noisy (imperfect) observations. We briefly point out the cases where our results can be extended beyond perfect observations.

*Originally published as [Shperberg and Shimony, 2017] (JAIR).

The theoretical results in this paper (Section 2) examine the conditions under which the batch value of perfect information is submodular, and provide counterexamples for some attempts to generalize the submodularity conditions. We also show that optimizing VOI is NP-hard even under very restrictive assumptions. Finally, we capitalize on the submodularity results by suggesting a simple "compound" greedy scheme in Section 3 for near-optimal solution of the selection problem, and compare its performance to the standard greedy algorithms on a wine quality dataset.

2 Main Results

Definition 1 (perfect information batch selection setting). *Let* $\mathcal{I} = \{I_0, I_1, ..., I_n\}$ *be a set of* $n+1$ *items of uncertain utility, represented by r.v.s* $X_0, ..., X_n$. *We assume w.l.o.g. that the current best item* α *is item* I_0. *For a cost* C_i *we can measure* I_i. *We select a subset* $S \subseteq \mathcal{I}$ *to be measured as a batch, for a total cost of* $\sum_{I_i \in S} C_i$, *observe the results* O *(the true utilities of the items in* S*), and select a final item* $I_f(O)$ *that has the highest expected utility given the observations.*

The *optimization version* of the perfect information batch selection problem is: under the perfect information batch selection setting, find the set that achieves:

$$\max_{S \in \mathcal{I}} (E_S[I_f(O)] - \sum_{I_i \in S} C_i) \quad (1)$$

Optionally, in the *budget limited* version of the selection problem, we are given a budget limit C and need to optimize S under the additional constraint: $\sum_{I_i \in S} C_i \leq C$.

Denoting the expected value of X_i by μ_i, note that by construction $\mu_\alpha \geq \mu_i$ for all $i > 0$. For a set of items $S \subseteq \mathcal{I}$, denote the expected value of information of a (perfect) observation of the utility of all these items by $VPI(S)$, defined as the expected value $E_S[I_f(O)] - \mu_\alpha$ (with expectation taken over all possible observations on S). Denote by p_i the PDF of random variable X_i.

Example 1. *Consider a wine selection problem with quality distributions similar to Figure 2. Suppose that one wine case* α *that we wish to purchase has a known quality of* $u_\alpha = 8$. *We have been offered an additional option, with a quality distribution* X_1 *uniformly distributed in* $\{4, 10\}$, *so* $\mu_1 = 7$. *Suppose that our utility scale is linear in the quality, and that all wines cases cost the same. Testing a wine case is possible, thereby revealing its true quality. If we make no tests, we should rationally pick the* α *wine for a quality of 8. If we do test the unknown wine prior to the purchase, then with probability 0.5 its quality is revealed as 10, and we purchase it, thereby gaining 2. Otherwise, stick with* α, *and gain nothing. On the average we gain 1, so* $VPI(\{X_1\}) = 1$.

2.1 Batch VOI for Known α

Theorem 1. *For a perfect information batch selection setting with independent item utility distributions, where the utility* u_α *of the currently best item* α *is known, the value of perfect information* $VPI(S)$ *is a submodular set function.*

Corollary 1. *Theorem 1 also holds given only a distribution over* u_α, *if there is no way to obtain additional information about* u_α. *That is because an optimal (risk neutral) decision maker would have to act as if* $u_\alpha = \mu_\alpha$.

A similar argument leads to a generalization to noisy observations: although Theorem 1 is stated in terms of perfect information, this is not an inherent limitation. Consider a more general setting where measurements are noisy, but the value of each item can be measured only once. In this setting, one can simply use the expected posterior value instead of the actual value when making the decision, and our results still apply. However, in settings where the measurement types on an item are allowed to vary (e.g. allow a choice between one and two conditionally independent measurements, or a choice of measurements that reveal different features of an item), it is well known that submodularity does not hold [Frazier and Powell, 2010; Tolpin and Shimony, 2012].

Theorem 1 is relevant to additional settings. First, consider the special case where $u_\alpha = 0$. In this case, action α can be re-cast as making no selection at all, and the conditions of the theorem hold if all items have a non-positive prior expected value. This is actually reasonable when items with uncertain value are being sold to our decision-making agent, as the seller wishes to gain from the sale, and presumably would not wish to sell an item for less than its expected value.

Complexity of the Selection Problem

The batch measurement selection problem was shown to be NP-hard [Reches *et al.*, 2013], in a setting where multiple noisy measurements per item are allowed. We show that the problem gives rise to an NP-hard decision problem even if the observations are perfect.

Definition 2 (perfect information budget-limited batch selection decision problem (PBSP)). *In the perfect information batch selection setting (Definition 1), is there a subset* $S \subseteq \mathcal{I}$ *which has a total measurement cost not greater than* C, *and such that the expected utility of the final item* I_f *selected after observing the utility of the items in* S, *is at least* U?

Theorem 2. *The PBSP is NP-hard.*

The proof is by reduction from Knapsack to a PBSP. We also show that the PBSP remains NP-hard under the conditions of Theorem 1.

VPI in the Presence of Dependencies

We now consider the perfect information batch selection setting, without the independence assumption. With dependencies the amount of information obtained by additional observations, having already made some observations, is ususally reduced. Intuition would suggest that the same would therefore occur for the VPI as well. Indeed, for $n = 2$ the VPI is still subadditive.

Theorem 3. *For a batch selection setting with 3 items, where the utility of the currently best item* α *is known, the value of perfect information is subadditive, i.e.* $VPI(\{1\}) + VPI(\{2\}) \geq VPI(\{1, 2\})$.

Unfortunately, this submodularity result has no *practical* use, as it does not generalize to $n \geq 3$, as is evident from the following counterexample. Let $u_\alpha = 10$, and we have 3 additional items with utility distributed as binary variables, with values $\{L, H\}$. The dependency is "parity", that is, exactly an even number of the items have value H, and the rest have value L. The distribution over the 4 possible legal configurations is uniform, i.e. each has probability 0.25. The utility values are: $u_{1_L} = u_{2_L} = 5$, and $u_{1_H} = u_{2_H} = 13$, so that $\mu_1 = \mu_2 = 9 < u_\alpha$. For the 3rd item, we have: $u_{3_L} = 0$, and $u_{3_H} = 18$, so that $\mu_3 = 9 < u_\alpha$. It is straightforward to show: $VPI(\{1, 2\}) > VPI(\{1\}) + VPI(\{2\})$.

2.2 Batch VOI with Unknown α

Consider now that we are given a distribution over u_α, but unlike corollary 1, additional information about u_α can be obtained. Now the VPI is neither necessarily submodular, nor necessarily supermodular: in the well-known case exhibited in figure 1, the VPI of each individual item is 0, but $VPI(\{\alpha, \beta\}) > 0$.

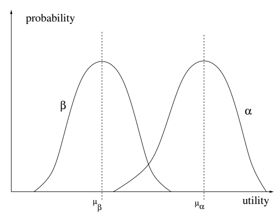

Figure 1: Utility distributions with supermodular VPI

Now consider: u_α is evenly distributed: $P(u_\alpha = 0) = P(u_\alpha = 10) = 0.5$, and u_β is distributed as $P(u_\beta = 1) = 0.7$, $P(u_\beta = 11) = 0.3$. We get $(\mu_\alpha = 5) > (\mu_\beta = 4)$. It is straightforward to show that we have:

$$VPI(\{\alpha, \beta\}) < VPI(\{\alpha\}) + VPI(\{\beta\})$$

An interesting question is about the VPI among sets of observations that *must* include an observation of the currently best item. In general, the VPI of such sets is neither submodular nor supermodular. Suppose we have 3 items, with distributions as follows. Current best item α, distributed: $P(u_\alpha = 20) = P(u_\alpha = 0) = 0.5$. Second best item β, distributed: $P(u_\beta = 9) = P(u_\beta = 5) = 0.5$, and third item γ, distributed $P(u_\gamma = 6) = P(u_\gamma = 2) = 0.5$. This gives us: $\mu_\alpha = 10$, $\mu_\beta = 7$ and $\mu_\gamma = 4$. We can show in this case:

$$VPI(\{\alpha, \beta, \gamma\}) > VPI(\{\alpha, \beta\}) + VPI(\{\alpha, \gamma\}) - VPI(\{\alpha\})$$

which clearly violates submodularity for sets containing observations of item α. However, if u_α is always sufficiently high, we can get submodularity as follows. Denote:

Condition C 1. $P(u_\alpha < \mu_i) = 0$ *for all* i *other than* α.

Example 2. *Consider the same wine selection problem instance as in example 1, except that the quality of the* α *wine case is no longer known to be 8: instead its quality is uniformly distributed among* $\{7, 8, 9\}$. *With the quality of* X_1 *independent of wine* α *and distributed as in example 1, this example obeys condition C1.*

Denote by $VPI^\alpha(S)$ the value of information of perfectly observing the utility of all items in S, as well as that of α.

Theorem 4. *For a batch selection setting with jointly independent items where condition C1 holds, the value of perfect information* $VPI^\alpha(S)$ *is a submodular set function of* S.

3 Application of Results

Consider a batch setting selection problem where the measurement cost function $\mathcal{C}$ is supermodular. If we know the utility of item α, then $VPI(S) - \mathcal{C}(S)$ is submodular due to Theorem 1. We can thus use a standard greedy algorithm that starts with an empty candidate set S, and repeatedly adds to S items that have the highest net gain (best (marginal) VPI minus cost), until no item has a positive net gain. We call this method the **(additive) greedy** algorithm. According to a fundemental result in [Nemhauser *et al.*, 1978], the greedy algorithm already guarantees an expected utility that is close to optimal: submodularity is a guarantee against "premature stopping" in the greedy algorithms.

If u_α is unknown, but obeys condition C1, run the greedy algorithm twice: once for sets that contain measurements of item α, and once for sets that do not. Compare the expected value of both resulting measurement sets, and return the better of the two. We call this method the **compound greedy** algorithm. Theorems 1 and 4 imply that the functions we optimize in both cases are submodular, thus the greedy algorithms return sets that are near-optimal.

3.1 Example Setting: Wine Selection

A comparison of algorithm performance on a typical dataset indicates the type of results one can observe with greedy optimization algorithms for the selection problem.

Definition 3 (Wine selection problem)**.** *Given a set of wine types* $\mathcal{I} = \{I_0, ..., I_n\}$, *each wine has an unknown quality, but a quality distribution is known for each type. In addition, for a known cost* C_i, *we can purchase and send a bottle of each wine type to a sommelier for analysis and quality determination. Which subset* S *of the wines (if any) needs to be sent to the sommelier in order to maximize the expected utility of testing and final decision? (That is, maximize the expected quality of the final selection, minus the sum of costs* C_i *of wines in* S, *i.e. the net VPI).*

The setting for the tests was based on the UCI white wine quality dataset [Cortez, 2009; Cortez *et al.*, 2009], from which we constructed for each wine a quality distribution, assumed to be independent. Figure 2 shows these distributions as a scatter plot where darker color indicates higher probability. Each value on the X axis indicates a specific wine type, with wines sorted by expected quality value.

Using the above distribution, the following experiments were conducted. Each experiment was on a set $\mathcal{I}$ of $n + 1$ randomly picked wines from the dataset, where n was an experimental parameter, and for each wine a random cost C_i was drawn uniformly between 0.01 to 0.1 (assumed to be on the same scale as quality values). The wine with the best expected value from $\mathcal{I}$ is the α wine, the prior best. We then used 4 different methods to find the measurement policy (i.e. batch of wines to be tested).

Exhaustive: Every possible subset S of $\mathcal{I}$ (both with and without the alpha wine) was examined. The S which maximized the net VPI was retured. S here is the optimal (batch) measurement policy.

Greedy (additive). Wines are kept sorted by myopic expected net VPI w.r.t. the current batch. A batch S is in-

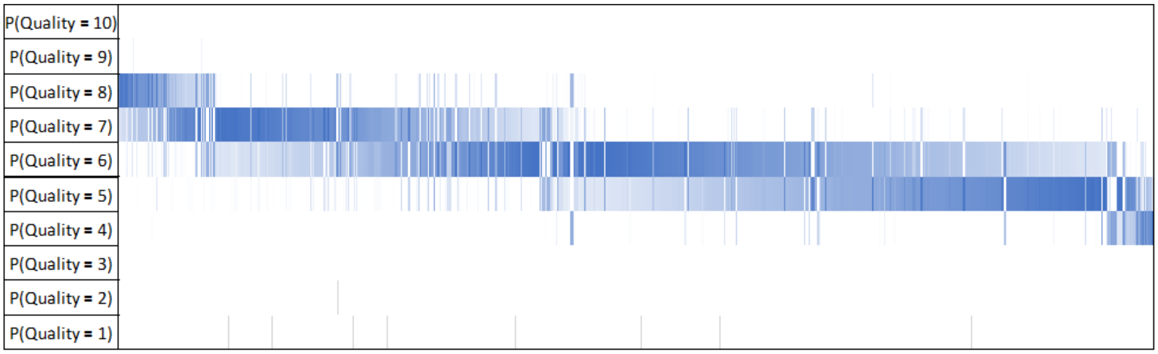

Figure 2: Wine quality distributions

crementally constructed, starting from the empty set: every iteration, the best candidate wine from $\mathcal{I} - S$ is added to S, as long as the net myopic VPI for adding this wine is positive.

Greedy (rate). Same as additive greedy, but with wines sorted according to expected VPI **divided** by cost of measurement, as in [Azimi *et al.*, 2016].

Compound greedy. Run the greedy (additive) algorithm twice: once for sets that do not contain the α wine, and once for sets that do. Compare the expected net VPI of both resulting measurement sets, and return the better of the two.

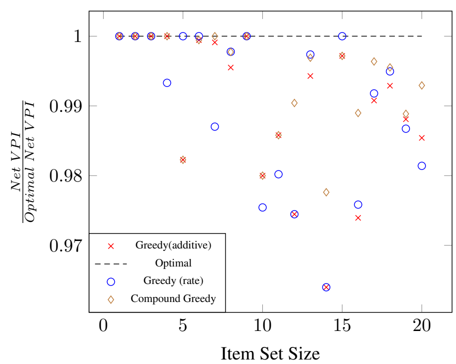

Figure 3: Comparison of net VPI for various item set sizes

The net VPI, averaged over 15 random item sets for each item set size, with n varying from 1 to 20, is shown in Figure 3. Both standard greedy algorithms averaged 0.99 of the optimal net VPI, while compound greedy averaged slightly better at 0.993, all considerably better than the theoretical bound.

Evaluating the VPI of a batch is non-trivial, but can be approximated by sampling [Azimi *et al.*, 2016]. In the wine selection problem we have independent discrete r.v.s with greatly overlapping domains, so we can cheaply compute the distribution of the maximum, and from there evaluate the expectation of the maximum exactly. Runtimes for the algorithms appear in Figure 4. Clearly, the exhaustive method delivers the best net VPI, but its runtime is prohibitive. Both the additive and rate greedy were the fastest, with compound greedy roughly a constant factor slower. In fact, despite the improved VPI computation, this part still dominates the runtime, and adding caching of computations of random variable maximizations resulted in the compound greedy algorithm being only a few percent slower than the other greedy algorithms (not shown). Therefore, although the improvement due to the compound greedy algorithm appears small, it comes essentially for free and is thus worthwhile. The greedy algorithms appear to be scalable: an experimental run with $n = 100$ wines resulted in runtimes of approximately 200 seconds for each of the greedy algorithms (including compound greedy, with caching).

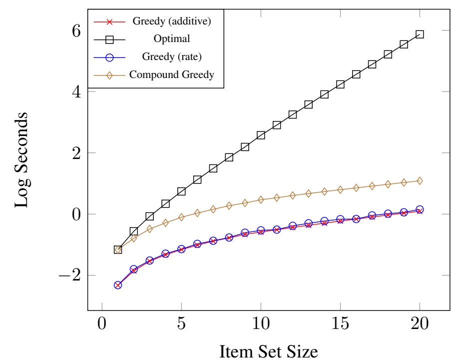

Figure 4: Runtime comparison: time vs. item set size

4 Conclusion

Cases where the batch value of perfect information is submodular in the selection problem were examined. We have shown that a resulting optimization problem is NP-hard, even in such restricted cases. Greedy optimization algorithms seem to achieve good results in practice. The theoretical results suggest that greedy algorithms should be supplemented by examining sets that include the currently best item, even if its individual VPI is zero, supported by empirical evidence.

We suggest that deviations from submodularity indicate points where the greedy and myopic optimization schemes can be improved w.r.t. net VPI, at relatively little computational cost. As such, the simple method suggested in this paper complements the idea of “blinkered VOI” [Hay *et al.*, 2012]. Our motivation for this work comes from metareasoning in search, where information is gathered by search actions. Solving a selection problem is a way to proceed at the first level in the search tree. Generalization beyond the first level is a non-trivial task for future research.

Acknowledgments

Supported by ISF grant 417/13 and by the Frankel Center.

References

[Azimi *et al.*, 2016] Javad Azimi, Xiaoli Fern, and Alan Fern. Budgeted optimization with constrained experiments. *Journal of Artif. Intell. Research*, 56:119–152, 2016.

[Cortez *et al.*, 2009] Paulo Cortez, António Cerdeira, Fernando Almeida, Telmo Matos, and José Reis. Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, 47(4):547–553, 2009.

[Cortez, 2009] Paulo Cortez. Wine Quality Data Set. `https://archive.ics.uci.edu/ml/datasets/Wine+Quality`, 2009. [Online; UCI Machine Learning Repository].

[Frazier and Powell, 2010] Peter I. Frazier and Warren B. Powell. Paradoxes in learning and the marginal value of information. *Decision Analysis*, 7(4):378–403, 2010.

[Hay *et al.*, 2012] Nicholas Hay, Stuart J. Russell, David Tolpin, and Solomon Eyal Shimony. Selecting computations: Theory and applications. In Nando de Freitas and Kevin P. Murphy, editors, *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence, Catalina Island, CA, USA, August 14-18, 2012*, pages 346–355. AUAI Press, 2012.

[Krause and Guestrin, 2009] Andreas Krause and Carlos Guestrin. Optimal value of information in graphical models. *J. Artif. Intell. Res. (JAIR)*, 35:557–591, 2009.

[Krause and Guestrin, 2011] Andreas Krause and Carlos Guestrin. Submodularity and its applications in optimized information gathering. *ACM TIST*, 2(4):32, 2011.

[Nemhauser *et al.*, 1978] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions - I. *Mathematical Programming*, 14(1):265–294, 1978.

[Papachristoudis and Fisher III, 2012] Georgios Papachristoudis and John W. Fisher III. Theoretical guarantees on penalized information gathering. In *Statistical Signal Processing Workshop (SSP)*, pages 301–304, Aug 2012.

[Radovilsky and Shimony, 2008] Yan Radovilsky and Solomon Eyal Shimony. Observation subset selection as local compilation of performance profiles. In David A. McAllester and Petri Myllymäki, editors, *UAI 2008, Proceedings of the 24th Conference in Uncertainty in Artificial Intelligence, Helsinki, Finland, July 9-12, 2008*, pages 460–467. AUAI Press, 2008.

[Radovilsky *et al.*, 2006] Yan Radovilsky, Guy Shattah, and Solomon Eyal Shimony. Efficient deterministic approximation algorithms for non-myopic value of information in graphical models. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Taipei, Taiwan, October 8-11, 2006*, pages 2559–2564. IEEE, 2006.

[Reches *et al.*, 2013] Shulamit Reches, Ya'akov (Kobi) Gal, and Sarit Kraus. Efficiently gathering information in costly domains. *Decision Support Systems*, 55(1):326–335, 2013.

[Russell and Wefald, 1991a] Stuart J. Russell and Eric Wefald. *Do the right thing - studies in limited rationality*. MIT Press, 1991.

[Russell and Wefald, 1991b] Stuart J. Russell and Eric Wefald. Principles of metareasoning. *Artif. Intell.*, 49(1-3):361–395, 1991.

[Shperberg and Shimony, 2017] Shahaf S. Shperberg and Solomon Eyal Shimony. Some properties of batch value of information in the selection problem. *J. Artif. Intell. Res. (JAIR)*, 58:777–796, 2017.

[Tolpin and Shimony, 2012] David Tolpin and Solomon Eyal Shimony. Semimyopic measurement selection for optimization under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 42(2):565–579, 2012.