

Gradient-Based Optimization for Bayesian Preference Elicitation

Ivan Vendrov, Tyler Lu, Qingqing Huang, Craig Boutilier

Google Research, Mountain View, California

{ivendrov, tylerlu, qqhuang, cboutilier}@google.com

Abstract

Effective techniques for eliciting user preferences have taken on added importance as recommender systems (RSs) become increasingly interactive and conversational. A common and conceptually appealing Bayesian criterion for selecting queries is *expected value of information (EVOI)*. Unfortunately, it is computationally prohibitive to construct queries with maximum EVOI in RSs with large item spaces. We tackle this issue by introducing a *continuous formulation of EVOI as a differentiable network* that can be optimized using gradient methods available in modern machine learning computational frameworks (e.g., TensorFlow, PyTorch). We exploit this to develop a novel Monte Carlo method for EVOI optimization, which is much more scalable for large item spaces than methods requiring explicit enumeration of items. While we emphasize the use of this approach for pairwise (or k -wise) comparisons of items, we also demonstrate how our method can be adapted to queries involving subsets of item attributes or “partial items,” which are often more cognitively manageable for users. Experiments show that our gradient-based EVOI technique achieves state-of-the-art performance across several domains while scaling to large item spaces.

1 Introduction

Rapid advances in AI, machine learning (ML), speech and language technologies have enabled recommender systems (RSs) to become more conversational and interactive. Increasingly, users engage RSs using language-based (speech or text) and multi-modal interfaces that have the potential to increase communication bandwidth with users compared to passive systems that make recommendations based only on user-initiated engagement. In such contexts, actively *eliciting* a user’s preferences with a limited amount of interaction can be critical to the user experience.

Preference elicitation has been widely studied in decision analysis and AI (Keeney and Raiffa 1976; Salo and Hämäläinen 2001; Chajewska, Koller, and Parr 2000; Boutilier 2002), but has received somewhat less attention in the recommender community (we discuss exceptions below). Bayesian methods have proven to be conceptually appealing for elicitation; in particular, *expected value of in-*

formation (EVOI) offers a principled criterion for selecting preference queries and determining when to make recommendations. EVOI is relatively practical in domains with small decision spaces and, as such, can be applied to certain types of content-based RSs, e.g., those with small “catalogs” of attribute-based items. However, maximizing EVOI is generally computationally intractable and often unable to scale to realistic settings with millions of items.

Our work aims to make Bayesian elicitation more practical. Our first contribution is a novel formulation of the problem of selecting a *slate query with maximum EVOI* as a continuous optimization in a differentiable network. This allows the problem to be solved using gradient-based techniques available in modern ML computational frameworks such as TensorFlow or PyTorch. The key to our approach is relaxing the set of items from which recommendations and queries are generated by allowing attributes to vary continuously without requiring they correspond to available items (these could be interpreted as “hypothetical” items in attribute- or embedding-space). Once optimized in this relaxed fashion, any hypotheticals are projected back into the nearest actual items to generate suitable recommendations or queries. We also propose regularizers that can be used during optimization to keep hypothetical items close to some actual item. We show empirically that this approach achieves comparable performance to state-of-the-art discrete methods, and also offers several key advantages: (a) it leverages highly optimized ML software and hardware for ease of implementation, performance and parallelization; (b) it generalizes to a variety of query types, including those involving items not present in the dataset, or partially specified items; (c) it accommodates continuous item attributes; and (d) it can use arbitrary differentiable metrics and non-linear utilities, allowing end-to-end optimization using gradient ascent.

Our second contribution leverages this flexibility—we propose a novel elicitation strategy based on *partial comparison queries*. In multi-attribute domains where items have large numbers of attributes, asking a user to compare complete items can be impractical (e.g., with voice interfaces) or cognitively burdensome. Instead we ask the user to compare partially instantiated items, with only a small number of attributes specified, as is common in decision analysis

(Keeney and Raiffa 1976). Finding EVOI-optimal partial queries is a difficult combinatorial problem, but we develop a simple, efficient continuous optimization method using the ideas above that performs well in practice.

The remainder of the paper is organized as follows. We outline related work in Sec. 2. In Sec. 3 we lay out the framework and preliminary concepts upon which our contributions are based. We introduce our basic continuous relaxation for Bayesian recommendation and EVOI computation using choice or comparison queries involving fully specified items in Sec. 4, and several algorithmic variants based on it. In Sec. 5 we apply the same principles to elicitation using queries that use only subsets of attributes. We provide detailed empirical evaluation of our methods in Sec. 6, analyzing both recommendation quality as a function of number of queries, and the computational effectiveness of our methods. We conclude with brief remarks on future research directions in Sec. 7. A longer version of the paper is available on arXiv, which includes an appendix with further algorithmic and experimental details.¹

2 Related Work

Preference elicitation has long been used to assess decision-maker preferences in decision analysis (Keeney and Raiffa 1976; Salo and Hämäläinen 2001; Holloway and White 2003), marketing science (Toubia, Hauser, and Simester 2004) and AI (Chajewska, Koller, and Parr 2000; Boutilier 2002), often exploiting the multi-attribute nature of actions/decisions. In RSs, elicitation has primarily been studied in domains where items have explicit attributes and the recommendable space has a relatively small number of items (e.g., content-based as opposed to collaborative filtering (CF) RSs), which makes such settings qualitatively similar to other multi-attribute settings (Pu and Chen 2008).

Most work on elicitation assumes some uncertainty representation over user preferences or utility functions, some criterion for making recommendations under such uncertainty, and uses one or more types of query to reduce that uncertainty. A variety of work uses *strict uncertainty* in which (the parameters of) a user’s utility function is (are) assumed to lie in some region (e.g., polytope) and queries are used to refine this region. Various decision criteria can be used to make recommendations in such models (e.g., using a centroid (Toubia, Hauser, and Simester 2004)), with *minimax regret* being one popular criterion in the AI literature (Boutilier et al. 2006). Queries can be selected based on their ability to reduce the volume or surface of a polytope, to reduce minimax regret, or other heuristic means.

A variety of query types can be used to elicit user preferences. It is common in multi-attribute utility models to exploit utility function structure to ask users about their preferences for small subsets of attributes (Keeney and Raiffa 1976; Fishburn 1967; Braziunas and Boutilier 2005). Techniques based on multi-attribute utility theory—the focus of our work—have been applied to RSs (see Chen & Pu (2004) for an overview). For example, critiquing methods (Burke 2002; Viappiani, Faltings, and Pu 2006) present candidate

items to a user, who critiques particular attributes (e.g., “I’d prefer a *smaller* item”) to drive future recommendations toward a preferred point, while unconstrained natural language conversations offer further flexibility (Radlinski et al. 2019). We focus on elicitation methods that use *set-wise/slate choice queries*: a user is presented with a slate of k (often multi-attribute) items and asked to state which is their most preferred. If $k = 2$, this is a classic *pairwise comparison*. Such queries are common in decision support, conjoint analysis and related areas.

In this work, we focus on Bayesian elicitation, in which a distribution over user utility functions (or parameters) is maintained and updated given user responses to queries. A fully Bayesian approach requires one to make recommendations using expected utility w.r.t. this distribution (Chajewska, Koller, and Parr 2000; Boutilier 2002; Viappiani and Boutilier 2010), but other criteria can be used for reasons of computational efficiency, e.g., minimizing maximum expected regret (loss) of a recommendation w.r.t. the utility distribution (Bourdache, Perny, and Spanjaard 2019).

Expected value of information (EVOI) (Howard and Matheson 1984) provides a principled technique for elicitation in Bayesian settings. EVOI requires that one ask queries such that posterior decision (or recommendation) quality (i.e., user utility) is maximized in expectation (w.r.t. to possible user responses). Such queries are, by construction, directly relevant to the decision task at hand. In RSs, “value” is usually defined as the utility of the top recommended item. Guo and Sanner (2010) select high EVOI queries assuming a diagonal Gaussian distribution over user utilities, but their algorithms evaluate all $O(n^2)$ pairwise queries (or $O(n)$ in their “greedy” algorithm), neither of which are practical for real-time elicitation over millions of items. The most direct predecessor of our work is that of Viappiani and Boutilier (2010), who propose an approximate iterative algorithm for EVOI optimization that only considers a small number of queries. We provide details of this approach in Sec. 3.4.

Other approaches to elicitation using distributions over user preferences include various forms of *maximum information gain*. These ask the query whose expected response offers maximal information according to some measure. Canal et al. (2019) select pairwise comparisons by maximizing mutual information between the user response and the user preference point in some high-dimensional space. Rokach et al. (2012) select the pairwise query that minimizes “weighted generalized variance,” a measure of posterior spread once the query is answered, while Zhao et al. (2018) ask the “least certain” pairwise comparison constructed from the top k items (or between the top k and the rest). While often more tractable than EVOI, information gain criteria suffer (compared to EVOI) from the fact that not all information has the same influence on recommendation quality—often only a small amount of information is decision-relevant. Other non-Bayesian query selection criteria are possible; e.g., Bourdache et al. (2019) use a variant of the current solution strategy (Boutilier et al. 2006) for generating pairwise comparison queries.

The continuous optimization method we propose in this work can handle arbitrary combinatorial items with both

¹Please see arXiv:1911.09153.

discrete and continuous attributes. Such a setting is studied by Dragone et al. (2017), where point estimates of utility are used, with a structured perceptron-like update after each query. Dragone et al. (2018) use partially specified items for elicitation, but consider a critiquing model of interaction whereas we focus on set-wise comparisons.

While content-based RSs model user preferences w.r.t. item attributes, in large, commercial RSs, user preferences are usually estimated using user interaction data collected passively. Preferences are modeled using CF (Konstan et al. 1997; Salakhutdinov and Mnih 2008), neural CF (Yi et al. 2019)) or related models, which often construct their own representations of items, e.g., by embedding items in some latent space. Among the earliest approaches to elicitation with latent item embeddings are active CF techniques (Rashid et al. 2002; Boutilier, Zemel, and Marlin 2003) that explicitly query users for item ratings; more recently Canal et al. (2019) similarly elicit over item embeddings. In this work, we also generate comparison queries using learned item embeddings, though our methods apply equally to observable item attributes.

Finally, we note that elicitation has strong connections to work on Bayesian (or blackbox) optimization (Shahriari et al. 2015), where the aim is to find the optimum of some function—analogueous to finding a good recommendation for a user—while minimizing the number of (expensive) function evaluations—analogueous to minimizing the number/cost of user queries. In contrast with most work in preference elicitation, Bayesian optimization methods typically focus on direct function evaluation, equivalent to asking a user for their utility for an item, as opposed to asking for a qualitative comparisons, though recent work considers dueling bandit models using pairwise comparisons in Bayesian optimization (González et al. 2017; Sui et al. 2018). Such techniques are generally derivative-free, though recent work considers augmenting the search for an optimum with (say, sampled) gradient information (e.g., Wu et al. 2017)). We use gradient information when computing EVOI, directly exploiting the linear nature of user utility in embedding space.

3 Background

We begin by outlining the basic framework, notation, and prior results on which our methods are based.

3.1 Bayesian Recommendation

We adopt a Bayesian approach to recommendation and elicitation in which the RS maintains a *probabilistic belief* about a user’s utility function over recommendable items. It uses this to generate both recommendations and elicitation queries. A *recommendation problem* has six main components: an *item set*, a user *utility function* drawn from a utility space, a *prior belief* over utility space, a *query space*, a *response space*, and a *response model*. We outline each in turn.

We assume a set of N recommendable items $\mathcal{X} \subseteq \mathbb{R}^d$, each an instantiation of d (for ease of exposition, real-valued) attributes (categorical attributes can be converted in standard fashion). These attributes may be dimensions in some latent space, say, as generated by some neural CF

model (see Sec. 6.) A *user*, for whom a recommendation is to be made, has a *linear utility function* u over items, parameterized as a vector $\mathbf{u} \in \mathcal{U}$, where $\mathcal{U} \subseteq \mathbb{R}^d$; i.e., $u(\mathbf{x}; \mathbf{u}) = \mathbf{x}^T \mathbf{u}$ for any $\mathbf{x} \in \mathcal{X}$.² A user’s most preferred item is that with greatest utility: $\mathbf{x}_\mathbf{u}^* = \operatorname{argmax}_{\mathbf{x}} u(\mathbf{x}; \mathbf{u})$. The *regret* of any recommendation $\mathbf{x} \in \mathcal{X}$ is $\operatorname{Regret}(\mathbf{x}; \mathbf{u}) = u(\mathbf{x}; \mathbf{u}) - u(\mathbf{x}_\mathbf{u}^*; \mathbf{u})$, and is a natural measure of recommendation quality. Finally, we assume the RS has some prior belief P_0 over \mathcal{U} . This reflects its uncertainty over the true utility function u of the user.

The prior P_0 is typically derived from past interactions with other users, and reflects the heterogeneity of user preferences in the domain. While we explicate our techniques in the *cold-start* regime where the RS has no information about the user in question, P_0 may also incorporate past interactions or other information about that user—this has no impact on our methodology. The prior will be updated as the RS interacts with the user, and we use P generically to denote the RS’s current belief.

Given a belief P , the *expected utility* of an item (or recommendation) \mathbf{x} is:

$$EU(\mathbf{x}; P) = \mathbb{E}_P[\mathbf{x}^T \mathbf{u}] = \mathbf{x}^T \left[\mathbb{E}_P \mathbf{u} \right] = \mathbf{x}^T \left[\int_{\mathcal{U}} \mathbf{u} P(\mathbf{u}) d\mathbf{u} \right]. \quad (1)$$

The optimal recommendation given P is the item with maximum expected utility:

$$\mathbf{x}_P^* = \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}} EU(\mathbf{x}; P), \quad EU^*(P) = EU(\mathbf{x}_P^*; P). \quad (2)$$

The RS can refine its belief and improve recommendation quality by asking the user questions about her preferences. Let \mathcal{Q} be the *query space*. For any query $q \in \mathcal{Q}$, the *response space* \mathcal{R}_q reflects possible user responses to q . For example, a pairwise comparison query (e.g., “do you prefer \mathbf{x}_1 to \mathbf{x}_2 ?”) has a binary response space (yes, no). For any sequence of queries $\mathbf{q} = (q_1, \dots, q_n)$, $n \geq 0$, to simplify notation we assume that any corresponding sequence of user responses $\mathbf{r} = (r_1, \dots, r_n)$, where $r_i \in \mathcal{R}_{q_i}$, uniquely determines \mathbf{q} (e.g., through suitable relabeling so that a “yes” response to some q is encoded differently than a “yes” to a different q'). We also assume the RS has a *response model* that specifies the probability $R(r_i | q; \mathbf{u})$ of a user with utility function \mathbf{u} offering response $r \in \mathcal{R}_q$ when asked query q .

We focus on *slate (comparison) queries*, $q = (\mathbf{x}_1, \dots, \mathbf{x}_k)$, $k \geq 2$, in which a slate of k items is presented to the user, who is asked which is most preferred.³ We consider two response models for slate queries. The *noiseless response model* assumes a user responds with the utility maximizing item: $R(\mathbf{x}_i | q; \mathbf{u}) = \mathbf{1}[i =$

²Linearity of utility may be restrictive if the attributes are observable or “catalog” properties of items as opposed to reflecting some learned embedding. Our methods can be extended to other utility representations in such a case, for example, UCP- (Boutilier, Bacchus, and Brafman 2001) or GAI-networks (Gonzales and Perny 2004; Braziunas and Boutilier 2005), which offer a linear parameterization of utility functions without imposing linearity w.r.t. the attributes themselves.

³The response set $\mathcal{R}_q = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ can be augmented with a “null” item to account for, say, a “none of the above” response.

$\text{argmax}_{j=1}^k u(\mathbf{x}_j; \mathbf{u})$]. The *logistic response model* assumes $R(\mathbf{x}_i|q; \mathbf{u}) = e^{u(\mathbf{x}_i; \mathbf{u})/\tau} / \sum_{j=1}^k e^{u(\mathbf{x}_j; \mathbf{u})/\tau}$, where temperature $\tau > 0$ controls the degree of stochasticity/noise in the user’s choice. Other choice models could be adopted as well (Louviere, Hensher, and Swait 2000).

We assume the response to any query is conditionally independent of any other given \mathbf{u} . Let $R(r|q; P) = \mathbb{E}_{\mathbf{u} \sim P} R(r|q; \mathbf{u})$ be the expected probability of response r given belief P . Given any response sequence \mathbf{r} (which determines the generating query sequence \mathbf{q}) and current belief P , the *posterior belief* of the RS is given by Bayes rule:

$$P_r(\mathbf{u}) = P(\mathbf{u}|\mathbf{r}) \propto R(\mathbf{r}|\mathbf{q}; P) = R(\mathbf{r}|\mathbf{q}; \mathbf{u})P(\mathbf{u}).$$

3.2 Expected Value of Information

While computing optimal query strategies can be cast as a sequential decision problem or POMDP (Boutillier 2002; Holloway and White 2003), we adopt a simpler, commonly used approach, namely, *myopic* selection of queries using the well-founded *expected value of information* (EVOI) criterion (Howard and Matheson 1984; Chajewska, Koller, and Parr 2000). Given belief P , we first define the expected utility (w.r.t. possible responses) of the best recommendation after the user answers query q :

Definition 1. The posterior expected utility (PEU) of q is:

$$PEU(q; P) = \sum_{r \in \mathcal{R}_q} R(r|q; P) EU^*(P_r). \quad (3)$$

The (myopic) expected value of information is $EVOI(q; P) = PEU(q; P) - EU^*(P)$.

A query with maximum PEU maximizes the expected utility of the best recommendation conditioned on the user’s response. EVOI, which is maximized by the same query, measures the *improvement* in expected utility offered by the query relative to the prior belief P and can serve as a useful metric for terminating elicitation.

3.3 EVOI Optimization with Particle Filtering

We use a Monte Carlo approach to optimize EVOI as in (Viappiani and Boutillier 2010). Given belief P , we sample m points $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_m)$ from P and maximize the sample average to approximate the integral within EVOI:

$$\begin{aligned} PEU(q; P) &= \sum_{r \in \mathcal{R}_q} R(r|q; P) \max_{\mathbf{y}_r \in \mathcal{X}} \left[\int_{\mathbf{u}} u(\mathbf{y}_r; \mathbf{u}) P_r(\mathbf{u}) d\mathbf{u} \right] \\ &\approx \sum_{r \in \mathcal{R}_q} \max_{\mathbf{y}_r \in \mathcal{X}} \left[\frac{1}{m} \sum_{j=1}^m \mathbf{y}_r^T \mathbf{u}_j R(r|q; \mathbf{u}_j) \right]. \end{aligned}$$

For slate queries under logistic response, an optimal query $q^* = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ w.r.t. EVOI satisfies:

$$\max_{\mathbf{x}_1, \dots, \mathbf{x}_k} \sum_{i=1}^k \max_{\mathbf{y}_i \in \mathcal{X}} \mathbf{y}_i^T \left[\frac{1}{m} \sum_{j=1}^m \mathbf{u}_j \frac{e^{\mathbf{x}_i^T \mathbf{u}_j / \tau}}{\sum_{\ell=1}^k e^{\mathbf{x}_\ell^T \mathbf{u}_j / \tau}} \right]. \quad (4)$$

Computing EVOI for a single query using this approach requires $O(Nmk^2)$ time. Consequently, search over all $\binom{N}{k}$ queries to maximize EVOI is prohibitively expensive, even in the pairwise case ($k = 2$), when N is moderately large.

3.4 Query Iteration

While finding optimal EVOI queries is intractable, effective heuristics are available. Notice that computing EVOI for a query $q = (\mathbf{x}_1 \dots \mathbf{x}_k)$ requires identifying the items with greatest *posterior* expected utility conditioned on each potential user response \mathbf{x}_i , i.e., the maximizing items \mathbf{y}_i^* in Eq. 4. We refer to this operation as a *deep retrieval* for query q , and write $(\mathbf{y}_1^* \dots \mathbf{y}_k^*) = \text{DeepRetr}(\mathbf{x}_1 \dots \mathbf{x}_k)$. We sometimes refer to $(\mathbf{x}_1 \dots \mathbf{x}_k)$ as the *query slate* and $(\mathbf{y}_1^* \dots \mathbf{y}_k^*)$ as the (*posterior*) *recommendation slate*.

The following result tells us that replacing the query slate with the induced recommendation slate increases EVOI:

Theorem 2 (Viappiani and Boutillier 2010). *Let $q = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ be a slate query with $q' = (\mathbf{y}_1^* \dots \mathbf{y}_k^*) = \text{DeepRetr}(q)$. Under noiseless response, $EVOI(q'; P) \geq EVOI(q; P)$; while under logistic response, $EVOI(q'; P) \geq EVOI(q; P) - \Delta$, where*

$$\Delta = \sum_{i=1}^k \left[\int \left(1 - \frac{e^{\mathbf{x}_i^T \mathbf{u} / \tau}}{\sum_{\ell=1}^k e^{\mathbf{x}_\ell^T \mathbf{u} / \tau}} \right) P_{\mathbf{x}_i}(\mathbf{u}) d\mathbf{u} \right] EU(\mathbf{y}_i^*; P_{\mathbf{x}_i}).$$

This suggests a natural iterative algorithm QueryIteration: start with a query slate, replace it with its induced recommendation slate, repeat until convergence. Viappiani and Boutillier (2010) find that, in practice, QueryIteration converges quickly and finds high EVOI queries in settings with small item sets.

4 Continuous EVOI Optimization

While QueryIteration is often an effective heuristic, it cannot scale to settings with large item spaces: each iteration requires the computation of the item with maximum PEU for each of the k responses; but computing this maximum generally requires EU computation for each candidate item.

To overcome this, we develop a continuous optimization formulation for EVOI maximization. Intuitively, we “relax” the discrete item space \mathcal{X} to obviate the need for enumeration, and treat the items in the query and recommendation slates as continuous vectors in \mathbb{R}^d . Once EVOI is optimized in the relaxed problem using gradient ascent, we project the solution back into the discrete space \mathcal{X} to obtain a slate query using only feasible items. Apart from scalability, this offers several advantages: we can exploit highly optimized and parallelizable ML frameworks (e.g. TensorFlow) and hardware for ease of implementation and performance; variants of our methods described below have common, reusable computational elements; and the methods are easily adapted to continuous item spaces and novel query types (see Sec. 5.)

4.1 A Reference Algorithm

We develop a basic continuous formulation by considering the EVOI objective in Eq. 4—we focus on logistic response for concreteness. In the discrete case, each item \mathbf{x}_i (in the query slate) and \mathbf{y}_i^* (recommendation slate) are optimized by enumerating the feasible item set \mathcal{X} . We relax the requirement that items lie in \mathcal{X} and treat these as continuous variables. Let \mathbf{X} (query slate) and \mathbf{Y} (recommendation slate) be

Algorithm 1 DeepRetrUniq. Inputs: optimized \mathbf{X}^* and \mathbf{U}

```

1: Let  $\mathcal{X}$  be an  $N \times d$  matrix of all feasible items.
2:  $S \leftarrow \emptyset$ 
3:  $\mathbf{p} = \text{softmax}(\mathbf{U}\mathbf{X}^*)$ 
4: Let  $\mathbf{T}$  be a  $k \times k$  matrix where the  $ij$ -th entry is the index
   of the  $j$ -th highest element in the  $i$ -th row of  $\mathbf{p}^T \mathbf{U} \mathcal{X}^T$ 
5: for  $i = 1..k$  do
6:   for  $j = 1..k$  do
7:     if  $T_{ij} \notin S$  then
8:        $S \leftarrow S \cup \{T_{ij}\}$ 
9:       break
10:    end if
11:   end for
12: end for
13: return slate of items indexed by  $S$ 

```

two $d \times k$ matrices whose i -th columns represent \mathbf{x}_i and \mathbf{y}_i^* , respectively. Let \mathbf{U} be an $m \times d$ matrix whose rows are the sampled utilities \mathbf{u}_i . We can express the softmax probabilities in the inner sum of Eq. 4 as $\text{softmax}(\mathbf{U}\mathbf{X})$, constructing a row vector of probabilities (each element is a logit).

Similarly, the dot products $\mathbf{y}_i^{*T} \mathbf{u}_j$ in the outer sum can be expressed as $\mathbf{U}\mathbf{Y}$. The Hadamard (element-wise) product gives:

$$[\mathbf{U}\mathbf{Y} \circ \text{softmax}(\mathbf{U}\mathbf{X})]_{ij} = \mathbf{y}_j^{*T} \mathbf{u}_i \frac{e^{\mathbf{x}_j^T \mathbf{u}_i / \tau}}{\sum_{\ell=1}^k e^{\mathbf{x}_\ell^T \mathbf{u}_i / \tau}}.$$

Summing over j and averaging over i , we obtain:

$$\frac{1}{m} \mathbf{1}_m^T \mathbf{U}\mathbf{Y} \circ \text{softmax}(\mathbf{U}\mathbf{X}) \mathbf{1}_k = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^k \mathbf{y}_j^{*T} \mathbf{u}_i \frac{e^{\mathbf{x}_j^T \mathbf{u}_i / \tau}}{\sum_{\ell=1}^k e^{\mathbf{x}_\ell^T \mathbf{u}_i / \tau}}, \quad (5)$$

where $\mathbf{1}_d$ is a d -vector of 1s. If we maximize \mathbf{Y} for any given \mathbf{X} , this is exactly the PEU of query slate \mathbf{X} . We can then apply gradient-based methods to optimize \mathbf{Y} (to compute PEU) and \mathbf{X} (to find the best query) as we detail below.

A solution \mathbf{X} may contain infeasible items depending on the optimization method used, in which case we must project the query slate \mathbf{X} into item space \mathcal{X} —we discuss mechanisms for this below. One approach is to leverage Thm. 2, and perform one deep retrieval to select feasible (optimal) recommendation items \mathbf{y}_i^* given \mathbf{X} . Thm. 2 ensures that $\{\mathbf{y}_i^*\}$, interpreted as a query, is a better feasible slate (or at least not much worse) than \mathbf{X} . To avoid duplicate \mathbf{y}_i^* 's, we deep retrieve the top- k items for each possible user response and only add distinct items to the query slate (see Alg. 1). Alg. 2 details this basic “reference” algorithm. We now discuss variations of the reference algorithm.

4.2 Taxonomy of Continuous EVOI Algorithms

The reference algorithm can be instantiated in different ways, depending on how we represent the (continuous) query and recommendation slates \mathbf{X} and \mathbf{Y} , and how we project these into feasible item space. These three choices form a taxonomy of continuous EVOI algorithms:

Algorithm 2 Continuous EVOI (reference algorithm)

```

1:  $\mathbf{X}^*, \mathbf{Y}^* \leftarrow \max_{\mathbf{X}, \mathbf{Y}} \frac{1}{m} \mathbf{1}_m^T \mathbf{U}\mathbf{Y} \circ \text{softmax}(\mathbf{U}\mathbf{X}) \mathbf{1}_k$  using
   gradient methods.
2: return DeepRetrUniq( $\mathbf{X}^*, \mathbf{U}$ ) as query slate

```

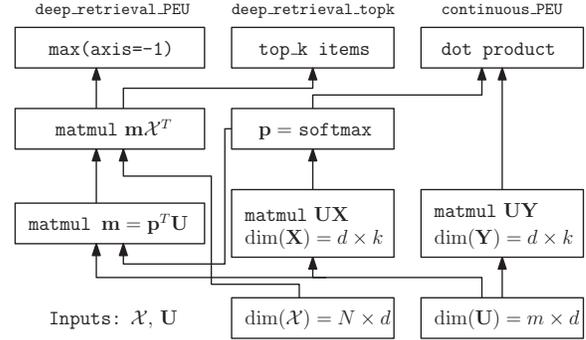


Figure 1: A common network architecture. Input items \mathcal{X} , utility particles \mathbf{U} are passed through dense layers with weights \mathbf{X} (columns are query items) and \mathbf{Y} (columns are rec. items *post*-user response). Output `continuous_PEU` is PEU with rec. slate \mathbf{Y} . If the slate is constrained to be feasible, `deep_retrieval_PEU` gives feasible PEU. `deep_retrieval_topk` outputs top- k EU items for each response to generate a query slate in DeepRetrUniq.

Query Slate Representation. The simplest way to represent queries are as “unconstrained” variables $\mathbf{x}_i \in \mathbb{R}^d$ uninfluenced by the feasible set \mathcal{X} (though possibly with some norm constraint). However, in most applications feasible items will be sparsely distributed in \mathbb{R}^d and such unconstrained optimization may yield queries far from any feasible item. To address this, we can incorporate “soft” feasibility constraints as regularization terms in the optimization, encouraging each \mathbf{x}_i to be near some feasible item. We note that this restriction can magnify local optimality issues.

Recommendation Slate Representation. As above, we can also treat the recommendation slate items \mathbf{y}_i^* as unconstrained variables or regularize them to be near feasible items. We consider two other approaches, allowing a trade-off between computational ease and fidelity of EVOI computation. The first *equates* the query and recommendation slates. Thm. 2 implies that there exist optimal EVOI queries where the two coincide, i.e. $\mathbf{y}_i^* = \mathbf{x}_i$, under noiseless response, and do so approximately under logistic response. Using a single set of slate variables in Eq. 5 for both gives the following optimization:

$$\operatorname{argmax}_q \text{EVOI}(q) \approx \operatorname{argmax}_{\mathbf{X}} \frac{1}{m} \mathbf{1}_m^T \mathbf{U}\mathbf{X} \circ \text{softmax}(\mathbf{U}\mathbf{X}) \mathbf{1}_k. \quad (6)$$

A more computationally expensive, but more accurate, approach for EVOI avoids relaxing the query slate, instead computing exact EVOI for \mathbf{X} . This requires k deep retrieval operations (one per item) over \mathcal{X} at each optimization step. In practice, we can apply deep retrieval only periodically, giving an alternating optimization procedure that optimizes query variables given the recommendation slate at one stage

and deep retrieving the recommendation items given the query variables the next. One could also compute the max over a small subset of prototype items.

Projecting to Feasible Items. The \mathbf{x}_i (query) values after optimization will not generally represent feasible items. We could recover a feasible query by projecting each \mathbf{x}_i to its nearest neighbour in \mathcal{X} , but this might give poor queries, e.g., by projecting an informative infeasible pairwise comparison to two feasible items that differ in only a dimension that offers no useful information. Instead, we use a single deep retrieval of \mathbf{X} to obtain a set of \mathbf{y}_i^* to serve as the query.

4.3 Variants of Continuous EVOI Algorithms

We identify four natural algorithms motivated by the taxonomy above. The modularity of our approach allows the design of a single network architecture, shown in Fig. 1, that supports all of these algorithms due to its reusable elements.

- ContFree. Equate query and recommendation slates; unconstrained except for an ℓ_2 norm bound on variables.
- ContReg. Like ContFree but with an added ℓ_2 regularization $\lambda \sum_{i=1}^k \min_{\mathbf{z} \in \mathcal{X}} \|\mathbf{x}_i - \mathbf{z}\|_2^2$ ensuring the variables stay close to feasible items.
- ContDeepRetr. Query variables are unconstrained, deep retrieval occurs at every optimization step.
- ContAlter. Query variables are unconstrained, and we use the alternating optimization described above.

All four algorithms use one deep retrieval at completion to ensure the resulting slate query uses feasible items in \mathcal{X} .

4.4 Initialization Strategies

Since the optimization objective is non-convex, we can (and do, in practice) encounter multiple local optima in each algorithm variant. Hence, the initialization strategy can impact results, so we consider several possibilities. The first strategy, Random, initializes the optimization with random vectors. The second follows (Viappiani and Boutilier 2010), initializing the slate query with the highest-utility items for k randomly selected utility particles (which they find performs well in conjunction with QueryIteration). We call this strategy RandUserTopItem. A third strategy, Balanced, initializes using a query slate that splits the sampled utility particles into k evenly weighted groups in which each of the k responses results in roughly m/k utility particles that are (mostly) consistent with the response. Such balanced queries may use infeasible items. In practice, we use multiple restarts to mitigate the effects of local maxima.

5 Partial Comparison Queries

In many domains, slate queries with fully specified items are impractical: they can impose a significant cognitive burden on the user (e.g., comparing two car configurations with dozens of attributes); and often a user lacks the information to determine a preference for specific items (e.g. comparing two movies the user has not seen). It is often more natural to use *partial comparison queries* with items partially specified using a small subset of attributes (e.g., truck vs.

coupe, comedy vs. drama). Finding partial queries with good EVOI requires searching through the combinatorial space of queries (attribute subsets). Unlike full queries, with partial queries the query and item spaces are distinct. However, we can readily adapt our continuous EVOI methods.

5.1 Semantics of Partial Comparisons

The most compelling semantics for partial comparisons may generally be domain specific, but in this work we adopt the simple *ceteris paribus* (all else equal) semantics (Fishburn 1977). Let \mathcal{X} be defined over d attributes A . Given $S \subseteq A$, a *partial item* \mathbf{x}_S is a vector with values defined over attributes in S . We assume there is some $\mathbf{y} \in \mathcal{X}$ s.t. $\mathbf{y}_S = \mathbf{x}_S$ (i.e., any partial item has a feasible completion). Given partial query slate $q = (\mathbf{x}_{1,S} \dots \mathbf{x}_{k,S})$, a *uniform completion* of q is any full query $q' = (\mathbf{x}_1 \dots \mathbf{x}_k)$ such that $\mathbf{x}_{1,A \setminus S} = \dots = \mathbf{x}_{k,A \setminus S}$ (i.e., each item is completed using the same values of attributes $A \setminus S$). *Ceteris paribus responses* require that, for *any* uniform completion q' of q :

$$P_r(\mathbf{x}_{i,S} | q; \mathbf{u}) = P_r(\mathbf{x}_i | \mathbf{u}, q') .$$

This condition holds under, say, additive independence (Fishburn 1967) and induces an especially simple response model if we assume utilities are linear. For instance, with logistic responses, we have:

$$P_r(\mathbf{x}_{i,S} | q; \mathbf{u}) = \text{softmax}(\mathbf{u}_S^T \mathbf{x}_{1,S}, \dots, \mathbf{u}_S^T \mathbf{x}_{k,S}) .$$

Thus our response model for partial queries is similar to those for full queries. The optimal EVOI problem must find both attributes and partial items:

$$\operatorname{argmax}_{S, \mathbf{x}_{1,S} \dots \mathbf{x}_{k,S}} \text{EVOI}(\mathbf{x}_{1,S} \dots \mathbf{x}_{k,S}) .$$

5.2 Continuous EVOI for Partial Comparisons

As in Sec. 4, we need a continuous slate representation and projection method. For query representation, we relax partial items to fall outside \mathcal{X}_S . W.l.o.g., let attribute values be binary with a relaxation in $[0, 1]$ and a partial item vector be a point in $[0, 1]^d$. Rather than representing the recommendation slate explicitly, we compute exact EVOI at each step of optimization (deep retrieval). We project to partial queries by only specifying a small number of attributes p for each item: we discretize by setting the p highest attribute values for each item to 1 and the rest to 0. This projection could cause arbitrary loss in EVOI, so we use ℓ_1 -regularization to encourage each \mathbf{x}_i to have p values close to 1 and the rest close to 0. The optimization objective becomes:

$$\operatorname{argmax}_{\mathbf{x}_1 \dots \mathbf{x}_k \in [0,1]^d} \text{EVOI}(\mathbf{x}_1 \dots \mathbf{x}_k) - \lambda \sum_i \|\text{sort}(\mathbf{x}_i) - \mathbf{j}\|_1 \quad (7)$$

where $\text{sort}(\cdot)$ sorts elements of a vector in ascending order, \mathbf{j} is a d -dimensional vector with the last p coordinates set to 1 and the rest to 0, and λ is the regularization constant. We call this the ContPartial algorithm.

6 Experiments

We assess our continuous EVOI algorithms by comparing their speed and the quality of queries they generate to several baseline methods in two sets of experiments: the first uses full-item comparison queries, the second uses our partial comparison queries. We are primarily interested in comparing elicitation performance using (true) regret over complete elicitation runs—in a real scenario, one would generally evaluate the quality of recommendations after each round of elicitation. Since all methods attempt to maximize EVOI, we also report EVOI.

In all experiments, we sample utility vectors \mathbf{u} from some prior to serve as “true users”, and simulate their query responses assuming a logistic response model. Regret is the difference in (true) utility of the best item under \mathbf{u} and that of the item with greatest EU item given the posterior. To reduce the effect of local EVOI maxima, we run our continuous methods and QueryIteration with 10 initialization restarts (unless otherwise stated) using one of random, Balanced or RandUserTopItem initializers.

6.1 Slate Comparisons with Complete Items

For full-item slates, we compare to the following baselines:

Random: Queries composed of randomly selected items.

Top5Exhaustive: Exhaustive search for the best slate among the top 5 EU items.

Greedy: The greedy algorithm of Viappiani and Boutilier (2010), which incrementally appends to the query slate the item that maximizes slate utility (given items already on the slate).

QueryIteration: The iterative algorithm from Sec. 3.4.

RandUserTopItem: The “sampling initialization” heuristic of Viappiani and Boutilier (2010), which uses the item with greatest utility for each of k randomly selected utility particles (which we find performs well on its own).

Exhaustive: Brute force an optimal EVOI query.

Synthetic. We consider a synthetic dataset where we sample both items and utility particles with dimension $d = 10$ from $\mathcal{N}(\mathbf{0}, \mathbf{I})$. We run 20 trials of simulated elicitation, each sampling 5000 items and 100 utility vectors. Results are plotted for pairwise comparisons in Fig. 2 (averaged regret is an empirical estimate of “expected loss” as defined by Chajewska et al. (2000)). Of the continuous methods, only ContAlter is plotted to reduce clutter in the figure. We first consider regret. The initial regret is 10.45 (no queries); and to reach regret of 1.0, ContAlter takes 5 queries (other continuous methods perform similarly) while QueryIteration, RandUserTopItem and Greedy require 7 queries. EVOI performance mirrors that of regret. Note that the EVOI of different methods is not generally comparable after the first query, since different queries result in different posteriors. The best performing strategies for the first query (EVOI) are QueryIteration (3.325), RandUserTopItem (3.269), ContAlter (3.266), and ContFree (3.161). Greedy performs worst with EVOI 3.078. Overall, these methods are competitive with Exhaustive, which achieves an average EVOI of 3.361 on the first query.

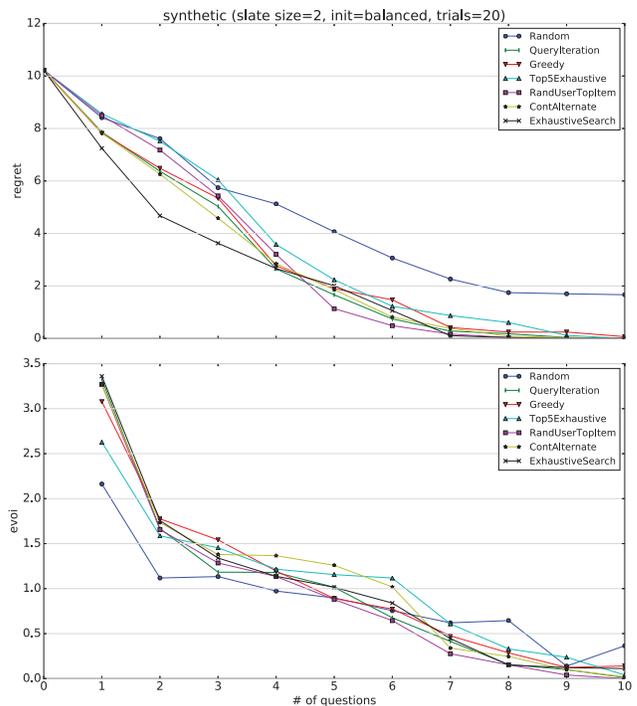


Figure 2: Simulated elicitation using synthetic data.

In terms of regret, ContAlter is quite competitive: achieving among the least regret across the first four queries. Because Exhaustive is myopic, it achieves the lowest regret only for the first few queries.

MovieLens. Using the MovieLens 100-K dataset, we train user and movie embeddings with dimension $d = 10$. Regret plots are shown in Fig. 3 for slate sizes $k = 2$ and $k = 5$, averaged over 100 elicitation trials of 10 queries (using 1682 items and random selections of 943 user embeddings in each trial). Again, ContAlter is the only continuous method plotted to reduce clutter. Regret starts at slightly below 0.3. For pairwise comparisons, RandUserTopItem takes 5 queries to reduce regret to 0.03 ($\approx 10\%$ of the original regret) while ContAlter (best continuous method), QueryIteration, Greedy and Exhaustive each require 6 queries. Interestingly, while Exhaustive reduces regret the most with the first two queries, it does not sustain its performance over a longer trajectory, likely due its myopic objective. While Top5Exhaustive performs best on the first query (both w.r.t. regret and EVOI) it does not maintain the same regret performance (or high EVOI) in subsequent rounds and requires 9 queries to reach regret 0.03. This demonstrates one weakness of myopic EVOI and the importance of non-myopic elicitation. For slates of size of 5, as expected, the same top methods require about half as many queries (3 to 4) to reduce regret to the same point (0.03). In terms of EVOI on the first query, for pairwise comparisons, the (in general, intractable) Top5Exhaustive method performs best with EVOI of 0.084, while ContDeepRetr and ContAlter are nearly as good, with 0.084 and 0.083, respectively. Next best is

	MovieLens			Goodreads		
(N, m, k)	(1682, 943, 2)	(1682, 943, 5)	$(2 \cdot 10^5, 100, 2)$	$(10^6, 100, 2)$	$(10^6, 500, 2)$	$(10^6, 500, 5)$
Greedy	0.01 (0.00)	0.04 (0.00)	0.14 (0.00)	0.77 (0.08)	4.26 (0.29)	17.62 (0.52)
ContAlter	2.56 (0.17)	2.71 (0.09)	2.24 (0.06)	2.54 (0.06)	2.96 (0.06)	3.14 (0.06)
ContFree	0.88 (0.04)	0.95 (0.05)	0.78 (0.04)	0.85 (0.04)	0.94 (0.04)	1.00 (0.04)
Exhaustive	1+ day	N/A	N/A	N/A	N/A	N/A

Table 1: Wall clock runtimes for EVOI algorithms. Avg. times (sec.) along with standard deviation, in parenthesis, are shown.

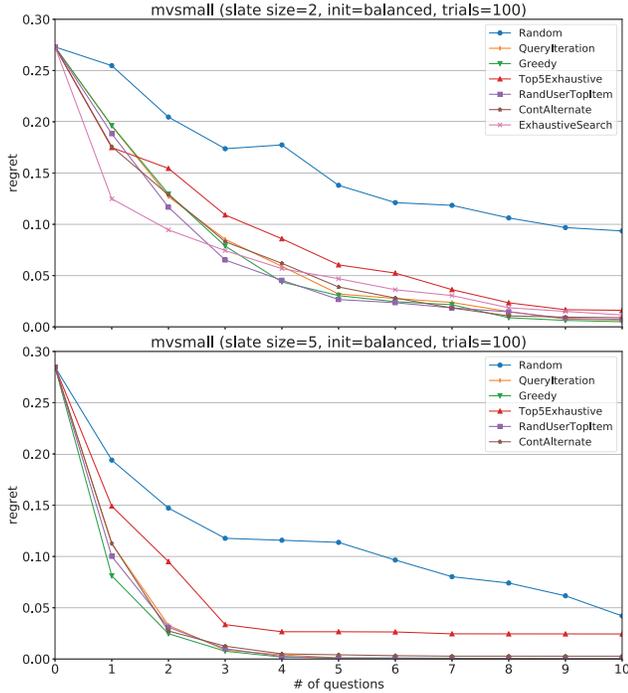


Figure 3: Elicitation using MovieLens-100k.

ContReg with 0.08, and the remaining methods achieve 0.0725 or below. For slate size 5, Greedy performs best with 0.167, followed by ContAlter and ContDeepRetr with 0.157. RandUserTopItem, QueryIteration and ContReg are also competitive with EVOI from 0.153 – 0.155; meanwhile Top5Exhaustive performs weaker with 0.134 (and is worse on all 100 trials vs. all other methods).

Goodreads. The last experiments use the much larger Goodreads dataset. We train a $d = 50$ user and item embedding model. Each trial consists of a random set of 2×10^5 items (books) and a random set of 100 user embeddings. Due to the large item space, we did not run ContDeepRetr and ContReg (it is possible to run them on subsampled item sets). Using Balanced initialization, the regret profile of continuous methods is not as competitive as above; e.g., in Fig. 4 we see it reaches regret 0.1 after 5 queries. Random performs significantly worse with this large item space. This is likely because most item pairs are not significantly different or the user embeddings tend not to vary much. If we instead use RandUserTopItem to initialize (as QueryIteration does), we obtain competitive results w.r.t. both regret and

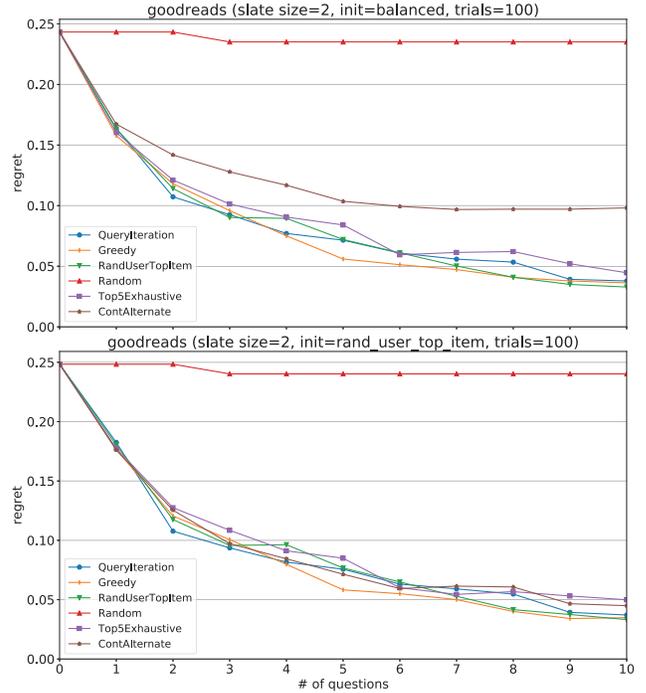


Figure 4: Elicitation on Goodreads, different initialization.

EVOI. We suspect that there could be significant structure in the (high-dimensional) item space that volumetric initialization does not capture well, resulting in poor local maxima. The choice of initialization has a similar impact on the EVOI of the first query. Our continuous methods achieve EVOI of about 0.076 compared to 0.089 for other methods.

Wall clock runtimes. We benchmark the algorithmic runtimes on a workstation with a 12-core Intel Xeon E5-1650 CPU at 3.6GHz, and 64GB of RAM. We measure the performance of Greedy, ContAlter, and ContFree (with Balanced initialization) since the other algorithms that are competitive in terms of EVOI are not computationally scalable. Results are shown in Table 1. We run 10 trials with 10 queries per trial for each algorithm and under each parameter setting. Note that Exhaustive is simply not scalable beyond the smallest problem instances. We implement Greedy using primarily matrix operations. For relatively small problems Greedy is fast, requiring at most 0.14s for MovieLens and Goodreads (where the number of items, $N \leq 2 \cdot 10^5$). However, and as expected, scaling is an issue for Greedy as it takes 4.26s to solve for pairwise comparisons with 10^6

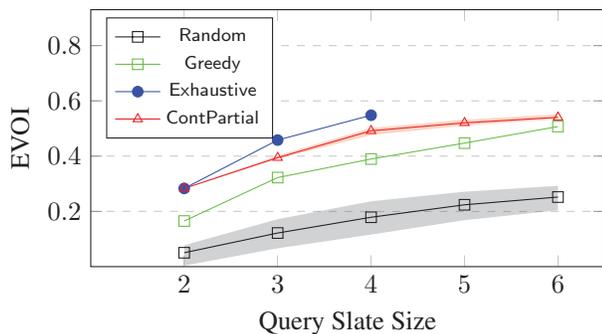


Figure 5: Single-attribute EVOI of queries found by ContPartial vs baselines. For methods with randomness we show the mean and interquartile range over 100 trials. Computational constraints only allow us to compute Exhaustive for $k \leq 4$.

items and 500 particles/users. For a larger slate size of 5 on Goodreads, Greedy becomes even less practical, requiring 17.62s to generate a query. Continuous methods scale better, with ContAlter taking only 3.14s on the largest problem instance while ContFree is even faster, taking at most 1s to generate a query (and is more consistent over all problem sizes)

6.2 Slate Comparisons with Partial Items

For partial comparison queries, we assess the quality of queries found by ContPartial by comparing the EVOI of the cold-start query it finds vis-a-vis three natural reference algorithms: (1) random queries; (2) a natural extension of Greedy for the partial setting—start with the attribute having highest utility variance across users, then greedily add attributes that result in the highest EVOI; and (3) exhaustive search for the best EVOI query.

We use the MovieLens-20M (Harper and Konstan 2015) dataset and represent each movie with 100 binary attributes from the Tag Genome (Vig, Sen, and Riedl 2012). We evaluate EVOI on the first query by randomly selecting 10^5 user embeddings as the prior, and running ContPartial for 100 restarts. We initialize query embeddings to random uniform values in $[0, 1]^{100}$, then run gradient ascent on Eq. 7 for 100 steps, initializing the regularization weight λ at 0.01 and multiplying λ by 1.1 each iteration.

As Figs. 5 and 6 show, ContPartial outperforms greedy and comes close to finding the best query for smaller slates and single attributes. With larger slates and multiple attributes, greedy performs better.

7 Conclusion and Future Work

We have developed a continuous relaxation for EVOI optimization for Bayesian preference elicitation in RSs that allows scalable, flexible gradient-based approaches. We also leveraged this approach to develop an EVOI-based algorithm for partial comparison queries. Our methods readily handle domains with large item spaces, continuous item attributes, and can be adapted to other differentiable metrics.

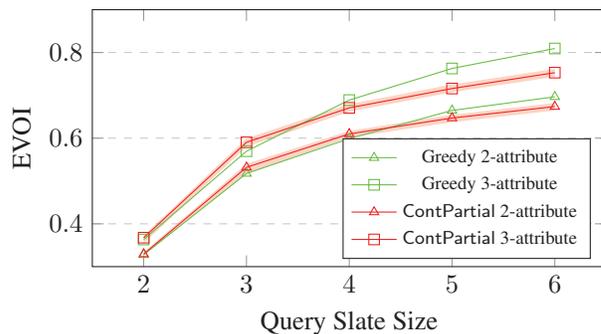


Figure 6: Multi-attribute EVOI of queries found by ContPartial vs our greedy baseline.

They can also leverage modern ML frameworks to exploit various algorithmic, software and hardware efficiencies. Our experiments show that continuous EVOI achieves state-of-the-art results in some domains.

There are various avenues for future work. Further exploration of different forms of partial comparisons is of interest, including the use of *latent* or high-level conceptual features while using continuous elicitation methods to generate informative queries from a much larger, perhaps continuous, query space. Methods that incorporate user knowledge, natural language modeling and visual features, together with explicit or latent attributes during elicitation would be of great value. Finally, evaluating recommendations using traditional ranking metrics and conducting user studies will play a key role in making elicitation more user-friendly.

References

- Bourdache, N.; Perny, P.; and Spanjaard, O. 2019. Incremental elicitation of rank-dependent aggregation functions based on bayesian linear regression. In *Proceedings of the Twenty-eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*, 2023–2029.
- Boutilier, C.; Bacchus, F.; and Brafman, R. I. 2001. UCP-Networks: A directed graphical representation of conditional utilities. 56–64.
- Boutilier, C.; Patrascu, R.; Poupart, P.; and Schuurmans, D. 2006. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence* 170(8–9):686–713.
- Boutilier, C.; Zemel, R. S.; and Marlin, B. 2003. Active collaborative filtering. 98–106.
- Boutilier, C. 2002. A POMDP formulation of preference elicitation problems. 239–246.
- Braziunas, D., and Boutilier, C. 2005. Local utility elicitation in GAI models. 42–49.
- Burke, R. 2002. Interactive critiquing for catalog navigation in e-commerce. *Artificial Intelligence Review* 18(3–4):245–267.
- Canal, G.; Massimino, A.; Davenport, M.; and Rozell, C. 2019. Active embedding search via noisy paired compar-

- isons. In *International Conference on Machine Learning (ICML)*.
- Chajewska, U.; Koller, D.; and Parr, R. 2000. Making rational decisions using adaptive utility elicitation. 363–369.
- Chen, L., and Pu, P. 2004. Survey of preference elicitation methods. Technical Report TR IC/2004/67, EPFL.
- Dragone, P.; Teso, S.; Kumar, M.; and Passerini, A. 2018. Decomposition strategies for constructive preference elicitation. In *Proceedings of the 32st AAAI Conference on Artificial Intelligence*.
- Dragone, P.; Teso, S.; and Passerini, A. 2017. Constructive preference elicitation over hybrid combinatorial spaces. *CoRR* abs/1711.07875.
- Fishburn, P. C. 1967. Interdependence and additivity in multivariate, unidimensional expected utility theory. *International Economic Review* 8:335–342.
- Fishburn, P. C. 1977. Multiattribute utilities in expected utility theory. In Bell, D. E.; Keeney, R. L.; and Raiffa, H., eds., *Conflicting Objective in Decisions*. Wiley. 172–196.
- Gonzales, C., and Perny, P. 2004. GAI networks for utility elicitation. In *Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR2004)*, 224–234.
- González, J.; Dai, Z.; Damianou, A.; and Lawrence, N. D. 2017. Preferential Bayesian optimization. In *Proceedings of the 34th International Conference on Machine Learning (ICML-17)*, 1282–1291.
- Guo, S., and Sanner, S. 2010. Real-time multiattribute Bayesian preference elicitation with pairwise comparison queries. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS-10)*.
- Harper, F. M., and Konstan, J. A. 2015. The movielens datasets: History and context. *ACM Trans. Interact. Intell. Syst.* 5(4):19:1–19:19.
- Holloway, H. A., and White, III, C. C. 2003. Question selection for multiattribute decision-aiding. *European Journal of Operational Research* 148:525–543.
- Howard, R. A., and Matheson, J. E., eds. 1984. *Readings on the Principles and Applications of Decision Analysis*. Menlo Park, CA: Strategic Decision Group.
- Keeney, R. L., and Raiffa, H. 1976. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. New York: Wiley.
- Konstan, J. A.; Miller, B. N.; Maltz, D.; Herlocker, J. L.; Gordon, L. R.; and Riedl, J. 1997. Grouplens: Applying collaborative filtering to usenet news. *Communications of the ACM* 40(3):77–87.
- Louviere, J. J.; Hensher, D. A.; and Swait, J. D. 2000. *Stated Choice Methods: Analysis and Application*. Cambridge: Cambridge University Press.
- Pu, P., and Chen, L. 2008. User-involved preference elicitation for product search and recommender systems. *AI Magazine* 29(4):93–103.
- Radlinski, F.; Balog, K.; Byrne, B.; and Krishnamoorthi, K. 2019. Coached conversational preference elicitation: A case study in understanding movie preferences. In *Proceedings of the Annual SIGdial Meeting on Discourse and Dialogue*.
- Rashid, A. M.; Albert, I.; Cosley, D.; Lam, S. K.; McNee, S. M.; Konstan, J. A.; and Riedl, J. 2002. Getting to know you: Learning new user preferences in recommender systems. In *International Conference on Intelligent User Interfaces*, 127–134.
- Rokach, L., and Kisilevich, S. 2012. Initial profile generation in recommender systems using pairwise comparison. *Trans. Sys. Man Cyber Part C* 42(6):1854–1859.
- Salakhutdinov, R., and Mnih, A. 2008. Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. In *Proceedings of the International Conference on Machine Learning*, volume 25.
- Salo, A., and Hämmäläinen, R. P. 2001. Preference ratios in multiattribute evaluation (PRIME)—elicitation and decision procedures under incomplete information. *IEEE Trans. on Systems, Man and Cybernetics* 31(6):533–545.
- Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; and De Freitas, N. 2015. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* 104(1):148–175.
- Sui, Y.; Zoghi, M.; Hofmann, K.; and Yue, Y. 2018. Advancements in dueling bandits. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-18)*, 5502–5510.
- Toubia, O.; Hauser, J. R.; and Simester, D. I. 2004. Polyhedral methods for adaptive choice-based conjoint analysis. *Journal of Marketing Research* 41(1):116–131.
- Viappiani, P., and Boutilier, C. 2010. Optimal Bayesian recommendation sets and myopically optimal choice query sets. In *Advances in Neural Information Processing Systems 23 (NIPS)*.
- Viappiani, P.; Faltings, B.; and Pu, P. 2006. Preference-based search using example-critiquing with suggestions. 27:465–503.
- Fig, J.; Sen, S.; and Riedl, J. 2012. The tag genome: Encoding community knowledge to support novel interaction. *ACM Trans. Interact. Intell. Syst.* 2(3):13:1–13:44.
- Wu, J.; Poloczek, M.; Wilson, A. G.; and Frazier, P. 2017. Bayesian optimization with gradients. In *Advances in Neural Information Processing Systems (NIPS17)*, 5267–5278.
- Yi, X.; Yang, J.; Hong, L.; Cheng, D. Z.; Heldt, L.; Kumthekar, A.; Zhao, Z.; Wei, L.; and Chi, E. 2019. Sampling-bias-corrected neural modeling for large corpus item recommendations. In *Proceedings of the Thirteenth ACM Conference on Recommender Systems (RecSys19)*, 269–277.
- Zhao, Z.; Li, H.; Wang, J.; Kephart, J.; Mattei, N.; Su, H.; and Xia, L. 2018. A cost-effective framework for preference elicitation and aggregation. *arXiv preprint arXiv:1805.05287*.