Leveraging User Interaction Signals and Task State Information in Adaptively Optimizing Usefulness-Oriented Search Sessions

Jiqun Liu
The University of Oklahoma
Norman, OK, USA
jiqunliu@ou.edu

Chirag Shah
University of Washington
Seattle, WA, USA
chirags@uw.edu

ABSTRACT

Current information retrieval (IR) systems still face plenty of challenges when applied in addressing complex search tasks (CSTs) that trigger multi-round search iterations. Existing relevance-oriented optimization algorithms and metrics are limited in helping users find documents that are *useful* for completing CSTs, rather than merely topically relevant. To address this gap, our work aimed to characterize CSTs from a process-oriented perspective and develop a state-based adaptive approach to simulating and evaluating search path recommendations. Based on the data collected from 80 journalism search sessions, we first extracted intention-based task states from participants' annotations to characterize temporal their temporal cognitive changes in searching and validated the state labels with expert assessments. Built upon the state labels and state distribution patterns, we then developed a simulated adaptive search path recommendation approach, aiming to help users find needed useful documents quicker. The results demonstrate that 1) different types of CSTs can be differentiated based on their distinct state distribution and transition patterns; 2) After a small number of iterative training, our adaptive recommendation algorithm can consistently outperform the best possible performance from individual participants in terms of the useful-based search efficiency across all CSTs. Going beyond traditional static viewpoint of task facets and relevance-focused evaluation approach, our work characterizes CSTs with a dynamic perspective and develops a domain-specific adaptive search algorithm that can help users find useful documents quicker and learn from online search logs. Our findings can facilitate future exploration of adaptive search path adjustments for similar types of CSTs in other domains and work task scenarios.

CCS CONCEPTS

• **Information systems → Users and interactive retrieval**.

KEYWORDS

Task state; usefulness; adaptive search recommendation

ACM Reference Format:
Jiqun Liu and Chirag Shah. 2022. Leveraging User Interaction Signals and Task State Information in Adaptively Optimizing Usefulness-Oriented Search Sessions. In *The ACM/IEEE Joint Conference on Digital Libraries in 2022 (JCDL '22), June 20–24, 2022, Cologne, Germany*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3529372.3530926

JCDL '22, June 20–24, 2022, Cologne, Germany

ACM ISBN 978-1-4503-9345-4/22/06…$15.00
https://doi.org/10.1145/3529372.3530926

1 INTRODUCTION

Current search technologies and digital libraries still face plenty of challenges when applied in addressing *complex search tasks* that trigger multi-round interactions between users and information (e.g., finding useful information for applying for PhD programs or deciding investment portfolios) [8, 35, 36, 44]. Part of the complexity of this problem is reflected in the cognitive variation, transitions of sub-goals, as well as significant behavioral changes during the process of search interactions. In prolonged search sessions, while most of explicit behavioral variations (e.g., query reformulation, search result browsing, changes of eye fixations) can be directly observed and recorded, the implicit change in users' *search task states* (e.g., exploration, know-item focused search, learning and evaluation) is difficult to monitor or predict. Learning about changes in task states would allow researchers to better understand 1) how a user navigates through an evolving problem space via information searching and 2) how the dynamic nature of a complex search task is manifested through search processes [23]. Also, with respect to practical applications, it is believed that the knowledge about users' cognitive variations in search interactions can help tailor search paths to adaptively supporting task performances.

A large body of existing interactive information retrieval (IR) research have conceptualized tasks as static problems or predefined goals that motivate users interact with search systems. Empirical studies built upon this static definition have explored multiple facets of search tasks (e.g., task complexity, task topic) and their associations with search actions. However, as a search session proceeds, the impacts of predefined facets on users' sub-goals and local search steps gradually fade away [22]. With very little research exploring the dynamic nature of complex search tasks, it is unclear how these tasks are unfolded during search interactions, and how we can adaptively improve and optimize search recommendations according to task state transitions. Consequently, an unfortunately significant proportion of search interactions is repetitive, inefficient, or unpleasant. Left unchecked, this problem has far-reaching effects on people's productivity, learning performance, as well as the quality of critical decision-making.

To address this research gap, we re-conceptualize complex search tasks from a process-oriented perspective and define them as search tasks that involve uncertain, broad solution space and evolve over time during search processes. Built upon this definition, we characterize the dynamic nature of complex search tasks and utilize explict search interaction signals and the knowledge of task states in building state-aware adaptive search recommendations, with the goal of helping users find necessary useful information quicker. Specifically, we aim to answer three research questions:

- **RQ1**: How do users' task states vary during search interactions in complex search tasks of different types?
- **RQ2**: How can we leverage task state information and user interaction signals in developing adaptive search supports for users engaging in complex search tasks?

In interactive IR community, many researchers have explored user traits, search task characteristics [14, 20], as well as their associations with search sessions in varying study settings. Although the theoretical contributions are clearly conveyed, the practical value of the knowledge about users and their tasks often remains unclear without a series of evaluations of the performance of the algorithms or systems that actually utilize the knowledge. In our research, we learn the knowledge about users' task states and state transitions from the empirical evidences collected in user studies and then apply the knowledge in simulating adaptive search recommendations and evaluating the effects of usefulness-oriented recommendations on multiple facets of search interactions. Our study has following two main contributions:

- This study proposes a novel framework that characterizes the dynamic nature of complex search tasks and empirically demonstrates how the knowledge about users' evolving task states can be leveraged in building adaptive search path recommendations and improving users' search efficiency.
- Going beyond traditional relevance-centric measures, our evaluation experiment focuses on the extent to which the simulated search paths could help users gather needed amount of useful pages sooner. This usefulness-oriented approach echoes the actual need of users (finding useful information for task completion, instead of gathering topically relevant documents) and has the potential to achieve better task performance and user satisfaction.

This article is structured as follows. We start by introducing the background and reviewing related works, with the goal of clarifying the rationale and motivations behind our study. Then, this article explains the user study design and experimental setup. Next, we present the results and explain how our results answer the proposed research questions. After that, we discuss the connections between our findings with that of previous research and explains the knowledge gaps addressed by this study. We conclude by discussing the contributions and limitations of our study.

2 RELATED WORKS

This section covers three topics of research that are directly related to our research questions: *Complex Search Tasks*, *Process Models of Search Tasks*, and *Adaptive Search Recommendations*.

2.1 Complex Search Tasks

Task complexity is a critical task facet in IIR research. Researchers from information seeking and IR communities have developed multiple frameworks to describe complex search tasks and examine the impacts of task complexity on various aspects of search interactions. For instance, Byström and Järvelin explored multiple task levels and investigated how work task complexity affects people's interactions with information sources. Built upon Bloom et al. [4]'s taxonomy, Kelly et al. [16] explored the cognitive aspect of task complexity and classified search tasks into different categories according to the level of complexity of associated learning goals. Urgo et al. [42] expanded on Kelly et al. [16]'s typology of task complexity and integrated knowledge dimension with learning dimension in the context of complex task design and learning assessments. Capra et al. [5] studied the extent to which prior determinability of task outcomes is indicative of task complexity.

The existing frameworks discussed above define complex search tasks from a static perspective and focus on the predefined aspects of search tasks (e.g. complexity of learning goals, pre-determinability of task outcomes), without explaining the dynamic dimensions. Thus, it is not clear how the nature of a complex search task is manifested during the process of search interactions.

To address this gap and effectively support complex search tasks, some latest works in IIR research have identified *states* or *stages* as essential, dynamic components of a task and studied how users' information seeking intentions and encountered search problems vary across different stages of task-based search interactions [13, 23]. Given the knowledge learned about state transition patterns, we could describe the dynamic nature of complex search tasks within search sessions and differentiate tasks of different types based on the actual process of performing tasks, rather than predefined static characteristics (e.g., task goal, task product, task topic). This research direction emphasizes the value of understanding task processes and offers a new perspective for conceptualizing complex search tasks. However, it is still unclear how the process-oriented dynamic frameworks could be applied in producing adaptive search supports for users engaging in complex search tasks.

2.2 Process Models of Search Tasks

There has been increasing recognition that information seeking under search tasks of varying types should be construed as sessions or episodes, rather than a set of individual single-query-single-response segments [3]. During these sessions, users often issue several queries, review multiple search engine result pages (SERPs), and engage in a number of different behaviors in interactions, with the ultimate goal of completing a search task as well as the associated work task. Thus, learning the features of a search session would allow us to better understand the nature of the associated search task and inform the design of adaptive IR systems.

To develop a fine-grained model for characterizing whole session IIR, recent IIR research have explored users' information seeking goals or intentions in local search steps as well as the search problems they encountered at different moments of search [23, 30, 32, 33]. In addition, researchers have also developed a series of classifiers for predicting intentions and in-situ problem types based on observable search behaviors [29, 34]. These studies enriched our understanding about multidimensional transitions in task processes. Nevertheless, how the knowledge about these changes could be applied in supporting users is still an open question to the IIR community.

To support model training and scalable application, researchers have also developed a variety of formal models to depict and optimize search processes from different perspectives. For instance, Fuhr [10] extended the classic Probability Ranking Principle (PRP) from traditional IR to IIR with the development of the IIR-PRP model, which defines search sessions as situation transition processes. The IIR-PRP model is a critical step towards building a computational model for supporting the functional design of IR systems. Inspired by Microeconomics theory, Azzopardi [1] developed a search economy model to determine and describe optimal search behavior in interactive sessions. Zhang and Zhai [53] proposed a novel formal model that frames the task of an IR system as to select a sequence of "interface cards" to present to and support the user and enables adaptive optimization of navigational interfaces in an IR system. Despite the computational benefits, these formal models abstract out a variety of user traits via making simplified assumptions and do not have effective representations of task states and task-based document usefulness.

2.3 Adaptive Search Recommendations

Given the dynamic nature of task-based search interactions, IIR researchers have explored several ways to characterize the sequence and transitions of users' search phases and tactics. One widely-used approach is to simplify the process by making it memoryless (assuming that the current state is dependent only on its previous state) [9, 27, 41]). In IIR context, this assumption is reasonable in the sense that users' often decide their search tactics in local, small steps (i.e. query segments), rather than making global search plans beforehand [39, 49]. This is known as the Markov Property [12]. A Markov Decision Process (MDP) is a stochastic decision process that has the Markov property.

Developing adaptive recommendations is not new in IR [51, 54]. For instance, Luo et al. [27] applied partially observable MDP framework in characterizing and optimizing task-based search processes. Specifically, they proposed four hidden decision making states based on users' query term selection and page visiting behaviors. The four states were defined by two dimensions: (1) relevance dimension: whether the user thinks the returned documents are relevant. Luo et al. [27] argued that if the set of previously retrieved documents leads to at least one satisfied (SAT) click (dwell time longer than 30s on the clicked page) [6], then the current state is likely to be relevant. (2) exploration dimension: whether the user would like to explore another subtopic or keep searching within the current information patch. If the newly added query term(s) appear in previously retrieved documents, then it means the user stays at the same sub information need and thereby is likely to continue exploitation. Within the reinforcement learning framework, Luo et al. [27] mathematically modeled dynamics in search sessions and simulated optimized ranking based on the estimated states in the interaction process. Similarly, Zhai [50] proposed a game-theoretic framework of text retrieval, which frames IR process as a cooperative game between a user and search engine. The optimization goal of this cooperative game is to help the user complete their search task with minimum overall effort (e.g., number of steps, dwell time needed) and minimum operating cost for the search engine. Losada et al. [26] applied multi-armed bandits model in pooling-based evaluation and proposed a novel formal adjudication method that can optimize the number of judgments required to identify a predefined number of relevant documents and thereby reduce the cost of evaluation. Wei et al. [46] focused on the issue of learning to rank and developed a MDP-based learning model, namely *MDPRank*, aiming to optimize relevance-based measures.

The aforementioned research jointly offer a starting point for applying state-based approach in addressing IR research problems. One major limitation of their works is that we do not know what exactly users were trying to accomplish at different moments and to what extent can we help users find documents that are actually *useful* for accomplishing the goal or task, rather than just *topically relevant*. Therefore, we still need a dynamic state-based framework which can (1) clarify the connection between *how* people search (i.e. behaviors) and *why* people search in such ways (e.g. motivations, intentions, in-situ problems and supports needed) and then (2) leverage the knowledge about the *how* (i.e. user interactions) and *why* (intention-based task states) parts learned from authentic users in developing usefulness-oriented adaptive recommendations.

3 METHODOLOGY

To answer the two RQs, we analyzed search behavior and user annotation data collected through a controlled lab study, information seeking intention (ISI) study. The ISI study looked at users' local intentions and search activities in different steps of the search session. Our goal is to leverage the knowledge about intention states extracted from user annotation data in developing and evaluating adaptive recommendations to support users engaging in CSTs. Full descriptions about the user study procedure were reported separately in our previously published preliminary works [23]. This section aims to provide enough methodological details for readers to properly assess the validity of our research design without covering redundant contents published before.

3.1 Complex search tasks and participants

In the ISI study, we recruited 40 undergraduate students majoring in journalism at Rutgers university, with the goal of minimizing the potential effects of disparate knowledge backgrounds on the study. To simulated search scenarios of varying types, we developed four tasks within the journalism domain to encourage realistic, complex searches: copy editing (CPE), story pitch (STP), relationships (REL), and interview preparation (INT). The four types of complex search tasks are designed based on two task facets extracted from Li and Belkin [18]'s classification scheme: Product and Goal. In terms of task product, an *intellectual* task encourages participants to develop new ideas. A *factual* task is a search for known or objective information. Tasks with *specific* goals have goals that are unambiguous and quantifiable. Tasks with *amorphous* goals have no predetermined strategy, target, or result. In ISI study, each participant was instructed to complete two search tasks of varying types, using Latin Square design (pairing CPE with INT and STP with REL to ensure tasks offer a balanced perspective of topics and facet values). We had eight task-topic combinations (i.e., four task types and two topics: Coelacanths and Methane Clathrates) in total. The descriptions of assigned search tasks are provided in Table 2.

3.1.1 *Information seeking intentions and study procedure.*

The study began by introducing participants to the features offered by the browser plug-in through a tutorial and asking participants to complete a demographic survey. Participants were given the choice to explore anywhere on the web as long as searches were performed

Table 1: Tasks assigned in the ISI study.

Task Type	Task Descriptions
Copy Editing/**CPE** (Factual Specific)	**Assignment**: You are a copy editor at a newspaper and you have only 20 minutes to check the accuracy of six italicized statements in the excerpt of a piece of news story below. **Task**: Please find and save an authoritative page that either confirms or disconfirms each statement.
Story Pitch/**STP** (Factual Amorphous)	**Assignment**: You are planning to pitch a science story to your editor and need to identify interesting facts about the coelacanth ("see-la-kanth"), a fish that dates from the time of dinosaurs and was thought to be extinct. **Task**: Find and save Web pages that contain the six most interesting facts about coelacanths and/or research about their preservation.
Relationship/**REL** (Intellectual Amorphous)	**Assignment**: You are writing an article about coelacanths and conservation efforts. You have found an interesting article about coelacanths but in order to develop your article you need to be able to explain the relationship between key facts you have learned. **Task**: In the following, there are five italicized passages, find an authoritative Web page that explains the relationship between two of the italicized facts.
Interview Preparation/**INT** (Intellectual Amorphous)	**Assignment**: You are writing an article that profiles a scientist and their research work. You are preparing to interview Mark Erdmann, a marine biologist, about coelacanths and conservation programs. **Task**: Identify and save authoritative Web pages for the following: Identify two (living) people who likely can provide some personal stories about Dr. Erdmann and his work. Find the three most interesting facts about Dr. Erdmann's research. Find an interesting potential impact of Dr. Erdmann's work.

on the browser with the plug-in turned on in order to record their search actions and allow participants to easily save webpages. Prior to beginning a search, participants went through the task description and were given up to 20 minutes to complete the first task by performing Web searches and saving relevant pages and documents.

Then, participants were asked to read the explanatory information about the intention annotation task and how to annotate information seeking intentions. Researchers utilized the search intention typology developed by Rha et al. [32] that is based on a subcategory of the classification scheme of dynamic intentions proposed by Xie [49]. For each participant, the entire search session was divided into separate query segments and replayed: during every query statement, participants were asked to select any number of intentions from the typology. Note that A *query segment* starts with a query, includes searching and browsing activities associated with the query, and ends before the next query. Query segment serves as the basic unit of analysis for both task state distribution analysis (RQ1) and the simulation of adaptive search recommendations (RQ2).

During each round of annotation, participants could select as many intentions as desired that applied to a given segment. The intention selection process was repeated for every query segment. The same procedure was used to complete the second task and the study concluded with an exit interview that provided open-ended questions for participants to offer insight into their search experience. The entire study procedure lasted about two hours.

3.2 Temporal Variations in Task States

Tasks of different types could be identified and disambiguated based on the temporal transitions within search sessions. To answer RQ1 and describe the connection between task states and search interactions, we examined the distribution of different task states across different moments and stages of complex search tasks.

For state identification, we employed *K-modes clustering analysis* for extracting clusters out of users' annotations. K-modes clustering as a unsupervised learning method extended the traditional K-means paradigm to cluster categorical data. In the intention study (ISI) dataset, task states are represented by the clusters of intention vectors (each of the vectors consists of twenty individual intentions as separate elements). It is worth noting that clustering algorithm itself did not generate meaningful state type or label. After the valid clusters were captured, we interpreted each cluster and define labels or names of task states based on the most frequent information seeking intention(s) in the corresponding clusters. After clustering, we also recruited two external assessors (PhD students majoring in IR) to do external assessments for the clusters, aiming to ensure that the states we obtained through clustering analysis is not only mathematically valid, but also meaningful to human assessors.

3.3 Experimental Setup

To answer RQ2 and examine the practical value of state-based framework, we employ Q-learning algorithm and simulate query segment recommendations based on users' task states and estimated rewards (i.e. useful pages retrieved). Then, we evaluate our model by comparing the performance of simulated search path with that of the average performance of participants' original search sessions and measuring the extent to which the simulated recommendations can reduce the efforts (i.e. number of query segments or steps needed) of completing assigned tasks.

The ISI study dataset, consisting of 80 sessions and 693 unique query segments, has following three characteristics that enable us to simulate and evaluate Q-learning-based adaptive recommendations: (1) ISI study has specified task completion requirements for all four tasks, which allows us to properly define rewards and determine the termination point for each iteration of learning and simulation; (2) ISI study involves four task types and two different topics, which enables us to test the state-based Q-learning approach in a variety of contexts; (3) A relatively small amount of states (four intention-based task states from the answer to RQ1) allows us to update the entire policy of Q-learning algorithm in a more efficient manner.

In following subsections, we introduce key components of our model: *states*, *actions*, *rewards* and *evaluation measures*. Then, we explain how the Q-learning simulation model operates on our dataset.

3.3.1 States and Actions. A Q-learning algorithm is trained to decide actions based on current state as well as the value of state-action pair [38]. In this work, we employ the task states reported in our response to RQ1 as the states in simulation. In the iterative learning process, the evolving Q-learning algorithm selects and evaluates different actions.

We define actions as rules (e.g., distinct ranking algorithms, different combinations of parameters, and weights) of selecting a specific search path (i.e., query segment) from a finite set of available options under a given task state. The output of Q-learning algorithm is a *policy* that defines the value of all state-action pairs (i.e., the estimated reward a user can obtain through taking a specific action under a given state). An optimized policy can determine the optimal action given a user's task state and recommend a query segment extracted from the collection of all search interactions. Possible actions of our algorithm are defined based on following query-level behavioral features which do not rely on relevance feedback:

- **Query similarity** = (number of overlapped unique terms between two adjacent queries)/(number of total unique terms from the two queries)

- **Number of pages clicked within the query segment**
- **Average dwell time on content page**
- **Dwell time on search engine result page (SERP)**

To better utilize interaction signals, we also take into consideration *content page clicking patterns* and extract related features using *Word Embedding* technique. Specifically, we employ *Word2vec* algorithm (in this case, continuous bag-of-words algorithm, or CBOW) for producing "word embeddings" with the unique content pages or URLs clicked in all query segments. We define each URL as a “word” and a query segment consisting of multiple clicked URLs as a “sentence”. Using CBOW technique, each URL can be turned into a unique vector that includes multiple dimensions. In Natural Language Processing (NLP) tasks, Word2vec technique can "understand" the semantic difference between words and group words with similar meanings together [7, 40]. In this study, similar or adjacent URLs (e.g., visited content pages that are frequently “co-clicked” or close to each other in multiple query segments) will get high weight values on same dimensions. Applying word embedding technique here can help us turn unique page clicking patterns into behavioral features for training models and may allow us to automatically separate highly useful pages from less useful pages. Therefore, we add the trained weights of dimensions from CBOW algorithm to the query segment ranking function. Since we are dealing with a relatively small dataset, we set the number of dimensions to 5 for the hidden layer of CBOW Neural Network.

Inspired by the model setups in previous reinforcement-learning-based IR studies [27], we rank every qualified query segment (i.e. query segment under the same topic and state) based on a linear combination of the weighted ranks of the query segment generated according to the value of different features. Given a defined action, the query segment which achieves the highest total weighted rank (smallest rank value) will be ranked on the top and thus will be selected as the recommended search path. Based on this definition, we can obtain different actions or rules of ranking by manipulating the weights of one or more features. The Q-learning recommendation process can be described as a "guessing game": Based on the knowledge learned from previous actions and feedback, the algorithm seeks to guess the "best action" that can maximize the chance of obtaining highest rewards under a given state.

To create a relatively broad action space for training, we begin with the "baseline action" which assigns the same weight to all features and then change the weight of one feature at a time by increasing the weight of the feature by a factor of $x = 1, 2, 3, 4, 5$. At the same time, we evenly decrease the weights of other features. We repeat this process for all features (four behavioral features and five Word2vec dimensions), which generates 37 unique actions in total. During the training and learning process, using different actions may lead to different query segment recommendations and thereby produce different rewards for the user.

3.3.2 Rewards. Reward operationalization is key to the iterative training and optimization. Since our ultimate goal is to support users engaging in complex search tasks, we define rewards of query segments based on their *actual contributions to completing the task at hand*. In our lab studies, we asked participants to search for information and bookmark pages that are useful for completing the tasks we assigned to them. To accurately measure the reward associated with different bookmarked pages and query segments, we annotated bookmarked pages and determined their respective rewards under a certain task context based on the extent to which they meet the requirements stated in our task descriptions.

Specifically, for example, in the copy editing (CPE) task, we asked participants to search for and bookmark pages that either confirm or disconfirm one of the six statements we provided in task descriptions. Therefore, we annotated all bookmarked pages under this task and represent each bookmarked page with a vector consisting of six elements, corresponding to the six statements in CPE task. If a page confirms the first two of the six task statements, then we represent the page using vector $v = \{1, 1, 0, 0, 0, 0\}$. If another bookmarked page confirms the last statement, then the page can be represented using the vector $v = \{0, 0, 0, 0, 0, 1\}$. Thus, the total reward from these two pages can be represented by the sum vector $v = \{1, 1, 0, 0, 0, 1\}$. As the simulated search sessions proceed, we add the vectors of all bookmarked pages together and terminate a simulated session when all element values are larger than zero (which means all statements are confirmed or disconfirmed by at least one unique statement). Similarly, for the interview preparation (INT) tasks, we apply the same annotation and reward calculation methods and use corresponding reward vectors in training.

The story pitch (STP) and relationship (REL) tasks have more subjective and flexible criteria of search task completion. Specifically, the STP task asked participants to bookmark pages that contain six most interesting facts about coelacanths or methane clathrates. In REL task, participants were asked to find an authoritative web page that explains the relationship between two of the five listed facts. For these two tasks, we define the task completion (or simulation termination) point based on users’ bookmarking behavior: for the STP task, we terminate a simulated search session once the algorithm collects six unique bookmarked pages within the session. With respect to the REL task, we end a simulated session once the algorithm finds a potentially authoritative content page that was bookmarked by at least two different participants.

3.3.3 Q-learning Process. Suppose there is a target user u_t that the interactive search system is trying to help. The target user u_t initializes the search process with a starting task state s_1. Given the knowledge of the user’s task state, the search system goes back to the Q function, finds the corresponding state, and locates the current best available action a_1 under the given state. Q function defines the mapping from states to actions and determines the selection of action as it shows the expected value of the total reward starting from the current state. The actions are evaluated based on a weighted sum of the expected values of the rewards from all future search iterations starting from the current state. Note that at the initial state we may start the simulation process with a set of arbitrarily defined values of Q function for all state-action pairs.

Within the Q-learning framework, besides the immediate reward obtained, the response from search environment also includes the next state of the information searching episode. To keep the iterative simulation process moving forward in a reasonable way, we add a function $f_{state\_transition}$ which selects the next state s_1 based on the state transition probabilities we extracted from authentic search sessions in the ISI dataset. The selected action a_1 and the obtained reward r_1 through taking a_1 under the task state s_1 are used to update the values in Q function and improve the policy.

After that, another round of iteration starts, with the target user u_t moving to a new task state s_2. Similar to the previous round, the system will go back to the Q function again and select an action based on the current new state s_2. The action will point to another potentially good search path sp_n from a user u_n. Again, to keep the iterative simulation going, u_n's task state s_n is assumed to be determined by the function $f_{state\_transition}$ based on the corresponding state transition probabilities. The selected action a_n and the received reward r_n through taking a_n under the task state s_n will be utilized to update the values in Q function and further improving the policy (i.e., mapping current state to a potentially better action). We terminate a simulation episode and start another new simulation episode (initialized with a randomly selected state and query segment) once the current accumulated useful information fully satisfies all requirements of the search task at hand.

During each round of simulation, the value function of state-action pairs is updated as follows [38]:

$$Q^{new}(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha * (r_t + \gamma * max_a Q(s_{t+1}, a) - Q(s_t, a_t)) \tag{1}$$

where r_t is the reward the target user obtained when moving from s_t to s_{t+1}, and $\alpha \in (0, 1)$ is the learning rate. $max_a Q(s_{t+1}, a)$ is the estimate of optimal future value within the MDP. The discount factor $\gamma \in (0, 1)$ assigns higher value for the rewards received earlier than those received later.

Note that within the current greedy version of Q-learning framework, it is likely that the algorithm keeps repeating a local optimal action without exploring any alternative, potentially better actions. To address this problem, we will apply a ϵ-greedy policy here ($\epsilon = 0.15$). ϵ-greedy policy allows the Q-learning algorithm to (1) select an action with 1-ϵ probability that gives maximum expected reward in a given state or (2) select a random action with ϵ probability [48]. The parameter ϵ enables us to adjust the balance between exploitation (i.e. keep using a known action that gives local maximum reward) and exploration (i.e. go beyond local optimal action and investigate broader solution space under uncertainty).

*Algorithm*1 presents the Q-learning process for a given task in one round of iteration. It usually takes a relatively large amount of iterations (e.g., over 100 iterations) for a Q-learning algorithm to converge to good performance (in this case, a smaller number of steps for completing a task). In each round of *iteration*, we run ten *episodes*. This is because we have ten unique search sessions under each *task-topic combination*. In each simulated episode, we use a starting state of an authentic search session as the initial state. We terminate an Q-learning iteration and start a new iteration process once we exhaust all ten starting states. This simulation setup ensures that the simulated episodes and authentic search sessions share the same starting point and thereby allows us to reasonably compare the performance of simulated sessions with that of participants' actual search sessions.

3.3.4 Model Evaluation. To evaluate simulated search sessions, we employ multiple evaluation measures which address different aspects of task-based search interactions.

Algorithm 1: Q-learning process for each iteration.

Input: Set of all query segments M; States S_i associated with each segment; predefined task T; Set of all authentic search sessions under the task T: SE_T
Output: Number of steps for completing the search task T; Initialization;
Group query segments into different subsets M_{S_i} by state S_i;
foreach *authentic search session m in SE_T* **do**
  Take the first state S_0 of the session m as the initial state for simulation;
  while *Task requirement is NOT satisfied* **do**
    Take $action_n$ according to the current state S_i and $Q\{state, action\}values$;
    foreach *query segment sp_i in M_{S_i}* **do**
      Calculate the weight of each feature according to the $action_n$;
      Calculate the weighted rank value of sp_i under each feature;
      Calculate the total weighted rank value for sp_i;
    end
    Take the top ranked sp as the simulated search path for the current state;
    Receive the reward r_i (remove the reward from repeated pages);
    Update the policy/Q-table Q according to r_i under S_i and $action_n$;
    Move to the next state according to the state transition function $f_{state\_transition}$;
  end
  Calculate the number of steps used in the simulated session m';
end
Calculate the average number of steps used for all simulated sessions;
Save the updated policy Q for the next iteration of training;
Terminate current iteration;

- **Number of query segments**. Number of steps or query segments is the main measure for our evaluation. Using rewards manually annotated based on users' bookmarked pages to iteratively update the algorithm, we aim to reduce the steps (i.e., query segments) needed for completing the search task at hand and improve the efficiency of search.

We evaluate the effectiveness of our model by comparing the number of steps in simulated search sessions with two baselines:

- **average number of steps in original search sessions**.
- **average number of steps for reaching the completion point**, which refers to the actual number of steps each user took for satisfying the specified task requirements under a given task.

The second measure was computed at that point in original search sessions where pages containing enough answers to the search task had been retrieved and displayed to the user. In other words, the task-based reward requirement is satisfied from the "information supply" side: the values in all elements of the corresponding task-based cumulative vector are greater than zero, regardless of users' knowledge states (e.g. whether a user fully understood the information presented on a visited page). Thus, number of steps needed for actual completion is essentially the minimal number of query segments users took for presenting "just enough" relevant information within a search session. This minimal path length for task completion varies across different task types and thus were operationalized using different vectors. Note that the minimal-length or truncated search sessions may still contain irrelevant or unsatisfactory results as it was difficult for individual users to ensure that every step can contribute to task completion when searching under a unfamiliar topic. Also, users often continue searching for information when the task requirements were already met. This may be because they were not sure whether the task goals were actually accomplished or if all requirements were met. Given the feedback from post-search interview and the specific predefined task goals, it is very unlikely that participants continued searching simply because they wanted to explore more and not for the purpose of finishing the search task sooner.

4 RESULTS

4.1 RQ1: Temporal variations in task states

To answer RQ1, we identified four states and interpreted each extracted task state based on the most frequent information seeking intentions annotated within the query segments under the corresponding state. To test the validity of the task states extracted by K-modes clustering algorithm with Elbow method, the two invited expert assessors did manual task state annotation. Based on the results from clustering and assessors' annotations, we calculated the Cohen's Kappa coefficients κ for all three pairs:

- Annotator A and annotator B: 0.782
- Annotator A and clustering algorithm: 0.716
- Annotator B and clustering algorithm: 0.744

The Cohen's Kappa agreements exceed 0.65 in all pairs, indicating *substantial* agreement [17]. This result shows that the task states identified through clustering is meaningful and can be used for addressing RQ1 and RQ2. To better illustrate the meaning of the task state, we list the most frequent information seeking intentions with associated percentages under each state (see Table 2). For instance, in the Exploratory state, the percentage of the intention of identifying something to start searching is 100%, indicating that this intention occurred in all query segments under this state.

Table 2: Intention states from ISI dataset.

Task state	Frequency	Interpretation
Exploitation	54.3%, 376 query segments	The two most frequent intentions are find specific information (39.4%) and identify something more to search (40.4%). Meanwhile, the intention of identifying something to start searching never appears. Under this state, users are likely to have a certain topic in mind and they try to follow the current explored search path, keep exploiting the associated information patches and search for more relevant pages or other information items.
Known-item	18.2%, 126 query segments	The two most frequent intentions are find specific information (100%) and obtain specific information items (100%). In this state, users may have well-defined information need(s) or goals in mind to guide their search sessions. These specific targets may come from previous interactions and are not necessarily in users' minds when search sessions are initiated.
Exploratory	16.6%, 115 query segments	The most frequent intention in this state is identify something to start searching (100%). In this state, users may try to adopt new search strategies, explore unknown subtopics, or open new search paths.
Learn and Evaluate	10.9%, 76 query segments	In this state, most intentions under the Evaluate category (above 60%) and the intentions of learning domain knowledge and keeping useful links (both above 80%) occurred frequently.

To answer RQ1, we analyzed how the relative frequencies of different task states vary across different stages of search. Comparing the temporal variations of the relative frequencies of intention-based task states can illustrate how the difference between search tasks of different types is manifested at cognitive level.

The frequency densities of intention states are presented in Figure 1. In these figures, each dot represents a *unique query segment* occurring at a particular point of the associated search session. To capture the temporal changes in task state, we employed *query percent* to characterize the sequential aspect of sessions. Query percent here equals to n/N, where n represents the n-th query segment within a session, and N is the total number of query segments within the search session. Query percent was measured at the end of each query segment and shows the stage a user was in within a task-based search session. For instance, in a four-query search session, the query percent value of the first query segment is 0.25, whereas in a five-query search session, the query percent value of the first segment is 0.2. Using query percent measure allows us to reasonably compare search sessions of varying lengths.

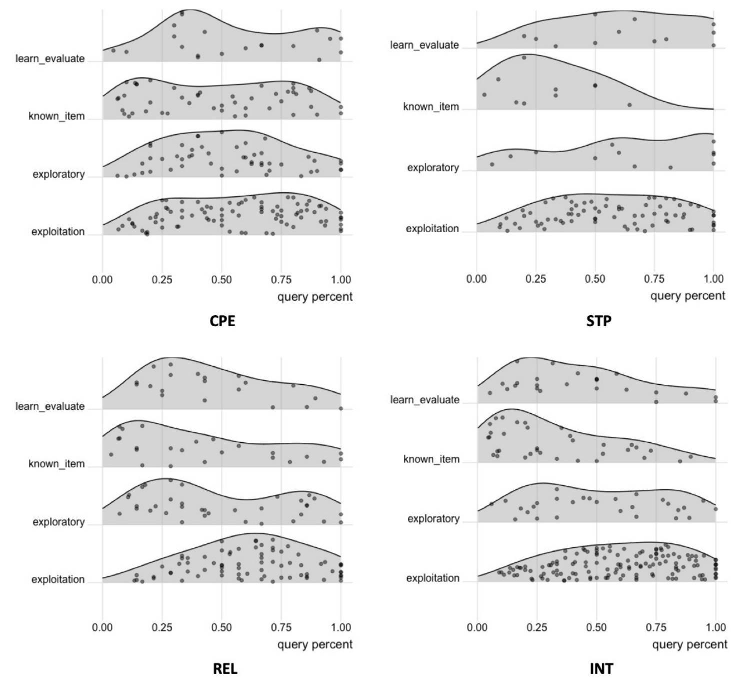

Figure 1: Frequency density of intention states at different search stages measured by query percent.

Overall, participants tended to use more queries in the factual specific task (i.e., copy editing) compared to the three goal-amorphous tasks. In the two cognitively challenging, intellectual-amorphous tasks, the density plot peak points of known-item state and learn-and-evaluate state show up at early stages of search sessions (query percent < 0.3). This result indicates that in these two tasks, participants frequently searched for easily accessible, known items and sought to evaluate information at the beginning of search sessions. In contrast, in the two factual tasks, the curves of frequency density distribution are relatively flat and the density plot peak points appear a bit later compared to the two intellectual tasks. In addition, in the two intellectual amorphous tasks, due to the lack of clear information cues, participants tended to do exploratory search at early stages, with the peak point of exploratory-state frequency density plot occurring around the 0.25 point.

In addition to the temporal changes in task states within search sessions, we also investigated the transition patterns between different task states in search tasks of varying types. Overall, we found that the process of doing a complex search task is often nonlinear, and that the difference in task type can be inferred based on the variations of task state transition probabilities. For example, our results demonstrate that participants engaging with goal amorphous tasks were motivated to do more open-ended, exploratory search (i.e., 50% chance of remaining in and repeating the exploratory state). In contrast, in factual specific tasks (e.g., copy editing task), participants usually search for known information items more frequently and spent much less time in learning and evaluation. More detailed results regarding state transition probabilities in different task types are reported in our preliminary work [23].

Next section presents our simulation of state-aware adaptive recommendations (RQ2), which was built upon the descriptive results regarding temporal state distributions from RQ1.

4.2 RQ2: Simulated Adaptive Recommendations

Aligned with RQ2, the main goal of our state-based search simulation is to reduce steps (query segments) needed for completing the given task (satisficing all the requirements specified in task descriptions) and thereby improving the efficiency of search interaction.

Recall that in the ISI user study, participants were asked to search for information that might be useful for completing the predefined task. Participants were not required or encouraged to do anything beyond the search task requirements. However, we observed varying levels of divergence between the length of original search session and the length of actual completion sessions in all task-topic combinations: participants tended to continue their search processes and issue more queries when they already met the requirements of task completion, due to four possible reasons: 1) participants were not sure if the bookmarked pages include enough information for meeting the minimum requirement of search tasks; 2) there was still plenty of time left before the ongoing search session hit the predefined termination point (20 minutes), which encourages the participant to revisit the bookmarked pages and verify the saved information; 3) in intellectual amorphous tasks, due to the difficulty of search tasks and the ambiguity of search goals, participants might need to explore more pages and learn more about the associated topic(s) in order to fully understand the information gathered in previous query segments; 4) part of the useful information on bookmarked pages (which was captured and included in manual reward annotation) was missed by participants in their original search sessions. Consequently, they might continue searching for this part of information (e.g. for confirming some of the statements in task description) that was already presented in their previously bookmarked pages. For instance, on a Google books page, a participant might be searching for information about the size of Coelacanth (which can be inferred based on the query stored in the URL). However, the participant might not notice that the bookmarked Google books page also contains useful information about the process of discovering Coelacanth in Indonesia (which comes from another statement that needs confirmation).

As it is explained in *Methodology* section, for every task-topic pair (e.g. CPE-Coelacanth), in each iteration, we run ten episodes or sessions of simulation and use the starting states from all ten authentic search sessions as the initial states of our ten simulated sessions. For evaluation, we compare the average number of steps in every round of iteration with the average numbers of steps from both the actual completion baseline and original sessions.

Figure 2 includes the performance of simulated search sessions in copy editing (CPE) tasks. We observe that the state-based simulated search path outperforms both original session and actual completion baselines before the 75th iteration in both topics. At the last two iterations, the simulated search path statistically significantly outperforms both baselines (t-test, p-value < 0.01) by reducing two to four steps on average for search task completion. This result indicates that through learning the connection between rewards received and the context of recommendation (i.e., state-action pairs), our state-based Q-learning algorithm finds shorter search paths for finding enough useful documents and completing CPE tasks.

Regarding the story pitch (STP)-Coelcantah task, participants' average number of steps needed for completing the task is 2.583, which does not leave much room for Q-learning improvement. Nevertheless, our state-based model still achieves slightly better performances than the actual completion baseline, with the number of steps being reduced by 0.1 to 0.48 steps on average after the 60th iteration. Compared to the STP-Coelacanth task, STP-Methane-Clathrates task appeared to be more difficult, with participants taking significantly more steps for finding and saving pages that contain six most interesting facts about the topic. This hypothesis is partially confirmed by our data about participants' post-search task perception: According to the ratings, STP-Coelacanth has a lower average level of perceived difficulty (Mean difficulty_Coelacanth = 1.2, Mean difficulty_Methane Clathrates = 2.1; $p<0.05$) and a higher level of perceived task success (Mean perceived success_Coelacanth = 6.5, Mean perceived success_Methane Clathrates = 4.8; $p<0.05$). As a result, in the Methane Clathrates task, the Q-learning algorithm achieves a much higher improvement and reduces the average length of session by 2 to 3.6 steps after the 60th training iteration.

Compared to the two factual tasks (CPE and STP), the two intellectual tasks placed greater challenges on both participants and Q-learning algorithm. As a result, the simulated sessions achieve relatively smaller improvement in search efficiency and takes more rounds of learning iteration to converge to a better performance (except for the REL-Mathane-Clathrates task where the simulation algorithm significantly improves the performance of search interaction by reducing 3.7 steps on average after the 100th iteration). This may be because these two intellectual amorphous tasks required participants to engage in more high-level cognitive activities (e.g., examine the relationship between two facts, prepare useful materials for different aspects of an interview). Thus, participants were not able to gather all needed information and save useful pages within only one or two queries. This heightened requirement at cognitive level also makes it difficult for our Q-learning algorithm to identify potentially efficient search paths with high rewards and thereby limits the performance of simulated search episodes. In the REL-Coelacanth and two INT tasks, the final improvements achieved in the last two rounds of iterations range from 0.9 to 1.7 query segments. In addition, compared to CPE and STP tasks, in the two intellectual amorphous tasks, it takes more rounds of iterations for the algorithm to converge to good-performing policies.

Table 3 compares the results from the last two rounds of Q-learning iterations for each task-topic combination and that of the two baselines, namely actual completion session and original search session. The results demonstrate that the Q-learning algorithm significantly outperforms the two baselines to different extents in all

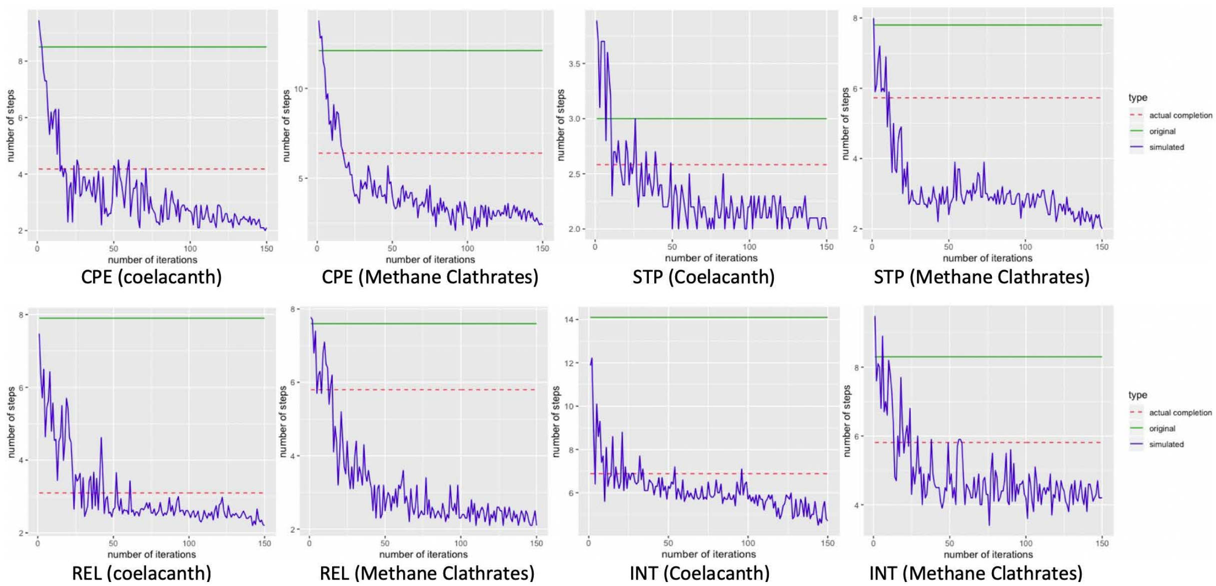

Figure 2: Number of steps needed for task completion. dashed line: average number of steps needed for actual completion (all needed information collected for fulfilling the task requirement); solid line: average number of steps in original sessions.

Table 3: Search efficiency: average number of steps.

Task-Topic	Simulated	Actual	Original
CPE-C (factual specific)	**2.05****	4.18	8.5
CPE-MC (factual specific)	**2.45****	6.4	12.1
STP-C (factual amorphous)	**2.05***	2.583	3
STP-MC (factual amorphous)	**2.1****	5.727	7.8
REL-C (intellectual amorphous)	**2.21****	3.1	7.9
REL-MC (intellectual amorphous)	**2.3****	5.8	7.6
INT-C (intellectual amorphous)	**4.75****	6.88	14.1
INT-MC (intellectual amorphous)	**4.2****	5.81	8.3

Note: C: Coalacanth; MC: Methane Clathrates; Significant values indicate whether the predictor is significantly better than the best baseline (i.e., actual completion baseline) ($*: p < .05$, $**: p < .01$). Statistically significant results are boldfaced.

task types. The simulated sessions consisting of query segments from different users reduce the number of steps needed for task completion and improve the efficiency in collecting useful documents. Due to the cognitive challenges behind the two intellectual amorphous tasks, our Q-learning algorithm achieves significant but smaller improvements compared to the performances in CPE and STP-Methane-Clathrates tasks. This may be because there were less high-reward search paths available in associated solution spaces and thus longer search sessions became inevitable. Nevertheless, in general, we are still able to support users and improve their search efficiency in REL and INT tasks after more rounds of iterations.

5 DISCUSSION

The contribution of our work is unique in 1) understanding the distribution of intention-based task states across varying task types, and 2) simulating and developing an usefulness-oriented dynamic approach for recommending search paths extracted from a collective solution space. Regarding the RQs, we have following answers.

The finding related to RQ1 extend the existing descriptive and computational models of task-based search process [2, 10, 43] by revealing the temporal cognitive changes in users' exploration of uncertain, evolving task-related informational solution space. Extracting and validating states based on users' segment-level annotations enabled us to collect direct empirical evidences on task-related variations at cognitive level and thereby better understand stepwise motivations behind the transition of search tactics [11, 49].

Then, to answer RQ2, we simulated state-aware adaptive recommendations using embedding techniques and reinforcement learning approach and evaluated their contributions in producing usefulness-focused search recommendations. Specifically, we built a Q-learning algorithm based on the knowledge we learned about state transitions and applied the model in simulating search sessions. Q-learning as a Reinforcement Learning method enabled us to run adaptive, step-by-step learning and utilize task state information and interaction signals in deciding the strategies of selecting query segments for recommendation. Our results demonstrate that the simulated search episodes can improve usefulness-oriented search efficiency to varying extents in different types of tasks.

As always, there are lessons learned from this study, limits to our work, as well as needs for future research efforts. A key implication is that Q-learning algorithm as a reinforcement learning method fits well with the research problem of developing state-based adaptive supports. Differing from traditional ML methods, Q-learning algorithm allows systems to learn continuously when new data streams keep flowing in and enables IR systems to iteratively update themselves according to the changing rewards associated with different ways of recommendations [45]. Thus, with a recommendation model built upon Q-learning algorithm, we would be able to support users in an online fashion.

The flexibility of Q-learning is not unconditional. To facilitate policy updates, we need to have clear definitions and measures regarding the benefits associated with each unit of action. In this study, we clearly defined the requirements of search task completion and represented the actual contribution of each bookmarked page using a unique vector. However, it is very difficult, if not entirely impossible, to accurately measure the "benefit" associated with each action in naturalistic search tasks. This is because (1) many search tasks in natural contexts are open-ended in nature and have no clear task completion point (e.g., find useful information about treating COVID patients); (2) users' perceived benefits in search sessions are very subjective and are often significantly affected by the gap between remembered utility and experienced utility [15, 21]; (3) Overall, users' perceived level of search success and actual search performance are not always aligned with each other. Smith and Rieh [37] argued that people often confuse the feeling of being able to find information with their own actual knowledge. Therefore, to develop generalizable intelligent search supports, researchers need to further explore how to build reliable and reproducible usefulness-based benchmark collections for task-based IR evaluations [28].

Moreover, the findings reported here from a journalism search study need to be tested based on the datasets collected under different domains, task contexts and study settings. It is also critical to investigate how user traits (e.g., existing belief [47], knowledge state [52], emotional state [25]) affect task progresses and user interactions at multiple levels, and how users evaluate their search gains, efforts and experiences differently under varying search states [24]. To address this issue, researchers need to design reliable measures that can capture the nature of these factors in various scenarios. For instance, in Web search, it is often difficult to differentiate the impact of topic knowledge from that of topic-independent search skills. In particular, how to accurately measure a user's level of search skills is still an open question [19, 31].

6 CONCLUSION

In summary, this work aimed to: 1) enhance our knowledge about the dynamic nature of CSTs, especially in terms of task state distributions; 2) leverage the knowledge learned about task states and state transition patterns in developing a computational model that produces useful adaptive search recommendations.

As a general matter, these goals were largely met. Our work connects the theoretical frameworks of search processes with the computational models of interactive IR and illustrates an innovative analytical approach that is both theoretically meaningful and practically applicable to the anatomy of and support for CSTs. Also, the Q-learning-based recommendation algorithm demonstrated here can potentially be applied in real-time search supports for CSTs from varying information-intensive decision-making scenarios.

Beyond simulations and experiments, a more complete development of state-aware, adaptive IR systems would require 1) a more psychologically realistic model of task processes built upon deeper knowledge about users' cognitive abilities, knowledge states, as well as systematic biases, 2) a comprehensive recommendation algorithm that incorporates more task features and user characteristics as parameters, 3) evaluation metrics that are more closely associated with users' in-situ and whole-session experiences, and 4) robust task-aware, usefulness-based benchmark collections. Besides, with respect to system design, future research also needs to explore new affordances and innovative interaction modalities in order to support more sophisticated intentions, task states as well as the associated search tactics that are not well supported by current IR systems and query-driven interaction paradigm.

7 ACKNOWLEDGMENTS

This work is supported by National Science Foundation (NSF) grants IIS-1423239 and IIS-1717488.

REFERENCES

[1] Leif Azzopardi. 2014. Modelling interaction with economic models of search. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*. 3–12.
[2] Marcia J Bates et al. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online Review* 13(5), 5 (1989), 407–424.
[3] Nicholas J Belkin. 2015. Salton award lecture: People, interacting with information. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1–2.
[4] Benjamin S Bloom et al. 1956. Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York, NY: McKay* 20 (1956), 24.
[5] Robert Capra, Jaime Arguello, Heather O'Brien, Yuan Li, and Bogeum Choi. 2018. The effects of manipulating task determinability on search behaviors and outcomes. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 445–454.
[6] Ye Chen, Ke Zhou, Yiqun Liu, Min Zhang, and Shaoping Ma. 2017. Meta-evaluation of online and offline web search evaluation metrics. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval*. 15–24.
[7] Kenneth Ward Church. 2017. Word2Vec. *Natural Language Engineering* 23(1), 1 (2017), 155–162.
[8] J Shane Culpepper, Fernando Diaz, and Mark D Smucker. 2018. Research frontiers in information retrieval: Report from the third strategic workshop on information retrieval in lorne (swirl 2018). In *ACM SIGIR Forum*, Vol. 52. ACM New York, NY, USA, 34–90.
[9] Sebastian Dungs and Norbert Fuhr. 2017. Advanced hidden markov models for recognizing search phases. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*. ACM, 257–260.
[10] Norbert Fuhr. 2008. A probability ranking principle for interactive information retrieval. *Information Retrieval* 11(3), 3 (2008), 251–265.
[11] Chathra Hendahewa and Chirag Shah. 2015. Implicit search feature based approach to assist users in exploratory search tasks. *Information Processing & Management* 51(5), 5 (2015), 643–661.
[12] Ronald A Howard. 1960. *Dynamic programming and Markov processes.* Hoboken, NJ: John Wiley & Sons.
[13] Hugo C Huurdeman and Jaap Kamps. 2014. From multistage information-seeking models to multistage search systems. In *Proceedings of the 5th Information Interaction in Context Symposium*. 145–154.
[14] Jiepu Jiang, Daqing He, and James Allan. 2014. Searching, Browsing, and Clicking in a Search Session: Changes in User Behavior by Task and over Time. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '14)*. ACM, 607–616.
[15] Daniel Kahneman. 2003. Maps of bounded rationality: Psychology for behavioral economics. *American economic review* 93, 5 (2003), 1449–1475.
[16] Diane Kelly, Jaime Arguello, Ashlee Edwards, and Wan-ching Wu. 2015. Development and evaluation of search tasks for IIR experiments using a cognitive complexity framework. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval*. 101–110.
[17] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* 33 (1977), 159–174.
[18] Yuelin Li and Nicholas J Belkin. 2008. A faceted approach to conceptualizing tasks in information seeking. *Information processing & management* 44, 6 (2008), 1822–1837.
[19] Jiqun Liu. 2021. Deconstructing search tasks in interactive information retrieval: A systematic review of task dimensions and predictors. *Information Processing & Management* 58, 3 (2021), 102522.
[20] Jingjing Liu, Michael J Cole, Chang Liu, Ralf Bierig, Jacek Gwizdka, Nicholas J Belkin, Jun Zhang, and Xiangmin Zhang. 2010. Search behaviors in different task types. In *Proceedings of the 10th annual joint conference on Digital libraries*. 69–78.
[21] Jiqun Liu and Fangyuan Han. 2020. Investigating Reference Dependence Effects on User Search Interaction and Satisfaction: A Behavioral Economics Perspective. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1141–1150.
[22] Jiqun Liu, Matthew Mitsui, Nicholas J Belkin, and Chirag Shah. 2019. Task, information seeking intentions, and user behavior: Toward a multi-level understanding of web search. In *Proceedings of the 2019 conference on human information*

interaction and retrieval. 123–132.

[23] Jiqun Liu, Shawon Sarkar, and Chirag Shah. 2020. Identifying and predicting the states of complex search tasks. In *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*. 193–202.

[24] Jiqun Liu and Ran Yu. 2021. State-Aware Meta-Evaluation of Evaluation Metrics in Interactive Information Retrieval. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 3258–3262.

[25] Irene Lopatovska. 2014. Toward a model of emotions and mood in the online information search process. *Journal of the Association for Information Science and Technology* 65(9), 9 (2014), 1775–1793.

[26] David E Losada, Javier Parapar, and Alvaro Barreiro. 2017. Multi-armed bandits for adjudicating documents in pooling-based evaluation of information retrieval systems. *Information Processing & Management* 53(5), 5 (2017), 1005–1025.

[27] Jiyun Luo, Sicong Zhang, and Hui Yang. 2014. Win-win search: Dual-agent stochastic game in session search. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. 587–596.

[28] Jiaxin Mao, Yiqun Liu, Ke Zhou, Jian-Yun Nie, Jingtao Song, Min Zhang, Shaoping Ma, Jiashen Sun, and Hengliang Luo. 2016. When does Relevance Mean Usefulness and User Satisfaction in Web Search?. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 463–472.

[29] Matthew Mitsui, Jiqun Liu, Nicholas J Belkin, and Chirag Shah. 2017. Predicting information seeking intentions from search behaviors. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1121–1124.

[30] Matthew Mitsui, Chirag Shah, and Nicholas J Belkin. 2016. Extracting information seeking intentions for web search sessions. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 841–844.

[31] Neema Moraveji, Daniel Russell, Jacob Bien, and David Mease. 2011. Measuring improvement in user search performance resulting from optimal search tips. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. 355–364.

[32] Eun Youp Rha, Matthew Mitsui, Nicholas J Belkin, and Chirag Shah. 2016. Exploring the relationships between search intentions and query reformulations. *Proceedings of the Association for Information Science and Technology* 53, 1 (2016), 1–9.

[33] Tuukka Ruotsalo, Jaakko Peltonen, Manuel JA Eugster, Dorota Głowacka, Patrik Floréen, Petri Myllymäki, Giulio Jacucci, and Samuel Kaski. 2018. Interactive intent modeling for exploratory search. *ACM Transactions on Information Systems (TOIS)* 36, 4 (2018), 1–46.

[34] Shawon Sarkar, Matthew Mitsui, Jiqun Liu, and Chirag Shah. 2020. Implicit information need as explicit problems, help, and behavioral signals. *Information Processing & Management* 57, 2 (2020), 102069.

[35] Chirag Shah and Ryen W White. 2020. Tutorial on Task-Based Search and Assistance. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2436–2439.

[36] Chirag Shah and Ryen W White. 2021. Task intelligence for search and recommendation. *Synthesis Lectures onSynthesis Lectures on Information Concepts, Retrieval, and Services* 13, 3 (2021), 1–160.

[37] Catherine L Smith and Soo Young Rieh. 2019. Knowledge-Context in Search Systems: Toward Information-Literate Actions. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. ACM, 55–62.

[38] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. Boston, MA: MIT Press.

[39] Jaime Teevan, Christine Alvarado, Mark S Ackerman, and David R Karger. 2004. The perfect search engine is not enough: a study of orienteering behavior in directed search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 415–422.

[40] Nguyen Huy Tien, Nguyen Minh Le, Yamasaki Tomohiro, and Izuha Tatsuya. 2019. Sentence modeling via multiple word embeddings and multi-level comparison for semantic textual similarity. *Information Processing & Management* 56(6), 6 (2019), 102090.

[41] Vu Tran, David Maxwell, Norbert Fuhr, and Leif Azzopardi. 2017. Personalised search time prediction using Markov chains. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*. 237–240.

[42] Kelsey Urgo, Jaime Arguello, and Robert Capra. 2019. Anderson and krathwohl's two-dimensional taxonomy applied to task creation and learning assessment. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*. 117–124.

[43] Pertti Vakkari. 2001. A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation* 57(1), 1 (2001), 44–60.

[44] Yiwei Wang and Chirag Shah. 2021. Authentic versus synthetic: An investigation of the influences of study settings and task configurations on search behaviors. *Journal of the Association for Information Science and Technology* (2021).

[45] Christopher Watkins and Peter Dayan. 1992. Q-learning. *Machine learning* 8, 3-4 (1992), 279–292.

[46] Zeng Wei, Jun Xu, Yanyan Lan, Jiafeng Guo, and Xueqi Cheng. 2017. Reinforcement learning to rank with Markov decision process. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 945–948.

[47] Ryen White. 2013. Beliefs and biases in web search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3–12.

[48] Michael Wunder, Michael L Littman, and Monica Babes. 2010. Classes of multiagent q-learning dynamics with epsilon-greedy exploration. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 1167–1174.

[49] Iris Xie. 2008. *Interactive information retrieval in digital environments*. Hershey, PA: IGI Global.

[50] ChengXiang Zhai. 2016. Towards a game-theoretic framework for text data retrieval. *IEEE Data Eng. Bull.* 39(3), 3 (2016), 51–62.

[51] Weinan Zhang, Xiangyu Zhao, Li Zhao, Dawei Yin, Grace Hui Yang, and Alex Beutel. 2020. Deep Reinforcement Learning for Information Retrieval: Fundamentals and Advances. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2468–2471.

[52] Yao Zhang and Chang Liu. 2020. Users' Knowledge Use and Change during Information Searching Process: A Perspective of Vocabulary Usage. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020*. 47–56.

[53] Yinan Zhang and Chengxiang Zhai. 2015. Information retrieval as card playing: A formal model for optimizing interactive retrieval interface. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 685–694.

[54] Sijin Zhou, Xinyi Dai, Haokun Chen, Weinan Zhang, Kan Ren, Ruiming Tang, Xiuqiang He, and Yong Yu. 2020. Interactive recommender system via knowledge graph-enhanced reinforcement learning. In *Proceedings of the 43rd ACM SIGIR Conference on Research and Development in Information Retrieval*. 179–188.