

# Intentions and Attention in Exploratory Health Search

Marc-Allen Cartright  
Department of Computer Science  
University of Massachusetts Amherst  
Amherst, MA 01003  
irmarc@cs.umass.edu

Ryen W. White  
Microsoft Research  
One Microsoft Way  
Redmond, WA 98052  
ryenw@microsoft.com

Eric Horvitz  
Microsoft Research  
One Microsoft Way  
Redmond, WA 98052  
horvitz@microsoft.com

## ABSTRACT

We study information goals and patterns of attention in exploratory search for health information on the Web, reporting results of a large-scale log-based study. We examine search activity associated with the goal of diagnosing illness from symptoms versus more general information-seeking about health and illness. We decompose exploratory health search into *evidence-based* and *hypothesis-directed* information seeking. Evidence-based search centers on the pursuit of details and relevance of signs and symptoms. Hypothesis-directed search includes the pursuit of content on one or more illnesses, including risk factors, treatments, and therapies for illnesses, and on the discrimination among different diseases under the uncertainty that exists in advance of a confirmed diagnosis. These different goals of exploratory health search are not independent, and transitions can occur between them within or across search sessions. We construct a classifier that identifies medically-related search sessions in log data. Given a set of search sessions flagged as health-related, we show how we can identify different intentions persisting as foci of attention within those sessions. Finally, we discuss how insights about foci dynamics can help us better understand exploratory health search behavior and better support health search on the Web.

## Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval – *search process, information filtering.*

## General Terms

Experimentation, Human Factors.

## Keywords

Health search, medical search, diagnosis, cyberchondria.

## 1. INTRODUCTION

Search engines are an important resource in locating medical information online. A December 2009 poll found that 66% of Internet users have pursued online information about health or medicine [20]. The Web has become a primary source of information about illnesses or treatments and a first stop for many when concerning symptoms emerge. Exploratory search describes a class of search activities that move beyond fact retrieval toward fostering learning, investigation, and information use [19]. Previous studies have shown that people use search engines to perform *exploratory health search* (EHS), where they diagnose medical conditions or perform informational health-related searches [23][28].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGIR '11, July 24–28, 2011, Beijing, China.

Copyright 2011 ACM 978-1-4503-0757-4/11/07...\$10.00.

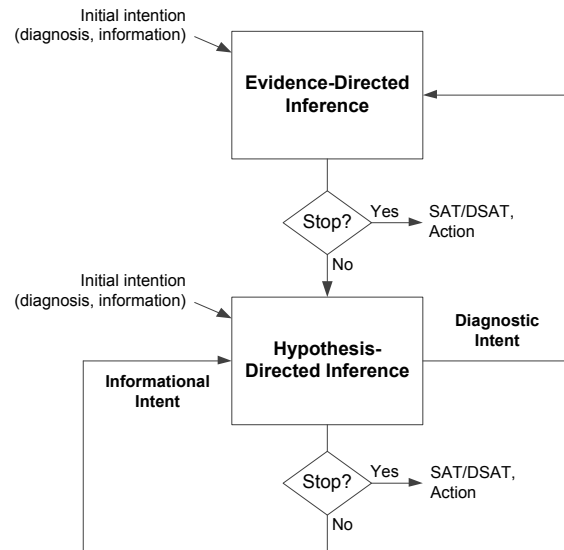


Figure 1. Intentions and flow of attention in exploratory health search.

We consider different informational goals associated with exploratory health search. In particular, we consider *diagnostic search* versus more general information-seeking about health and illness. We further partition diagnostic search into *evidence-directed* and *hypothesis-directed* information seeking. Evidence-based search focuses on the relevance of signs and symptoms to the potential existence of one of more medical disorders. With hypothesis-directed search, information is sought specifically about one or more illnesses. Hypothesis-directed information seeking plays a role in diagnostic search, where searchers pursue information about confirming or ruling out of potential disorders. However, hypothesis-directed search also describes the pursuit of information such as risk factors and therapies for illnesses, including the search for information on disorders that have been previously diagnosed by professionals.

We borrow the concepts of evidence-directed and hypothesis-directed phases of analysis from studies of diagnostic problem solving in medicine, which have demonstrated that physicians often employ an approximation of hypothetico-deductive reasoning [10], a method that involves a volley between two phases of analysis. In an *evidence-directed* phase, findings are fused to construct a list of potential explanatory diagnoses ranked by likelihood. In a *hypothesis-directed* phase, the current list of diagnoses ranked by likelihood is used to guide the collection of additional evidence, centering on discriminating among the competing hypotheses. The additional evidence collected via this phase is then considered, along with the evidence collected earlier, in a revised evidence-directed phase. Several automated decision-support systems for diagnosis and triage have been developed that employ the hypothetico-deductive cycle [11][13][17].

We believe that it can be useful to view diagnostic health search as a coarse form of hypothetico-deductive reasoning, where searchers harness Web search and browsing to guide the iterative cycle of hypotheses about wellness and disease being formulated from evidence and those hypotheses individually or together then guiding the collection of additional discriminating evidence. People performing diagnostic search may branch out of diagnostic goals and focus on general information about particular illnesses or symptoms. Also, people may initiate search without intention to perform diagnosis and flow into a diagnostic session. We consider non-diagnostic health search as being largely hypothesis-directed, focused on learning more about the nature and treatment of illnesses. Figure 1 captures key intentions and flow of attention in exploratory health search, where a searcher's initial intention may be diagnostic or informational. For diagnostic search, the user may cycle between evidence-directed and hypothesis-directed phases as their information needs evolve. At some point the searching will come to an end or be suspended, and the searcher will be either satisfied (SAT) or dissatisfied (DSAT) with the information acquired. At times, the searching will lead to actions in the world. Such actions include seeking healthcare services, evidence of which may be observable in log data as healthcare utilization intent (HUI) [29].

We explore in this paper intentions and the dynamics of attention in exploratory health search. Our work relates to prior work on understanding how users' cognitive states and information needs evolve during search as they discover new information [3]. An improved understanding of the processes by which people use Web search to explore health issues online and act on the information gathered can guide refinements of medical search systems.

To better understand exploratory health search, we analyze a set of filtered logs from a toolbar deployed by the Windows Live search engine. To perform the filtering, we first construct a statistical classifier that can identify EHS sessions from general search and browsing logs. Then, we extract terms in queries representing potential causes and symptoms, treating each of these distinct entities as a *focus*, capturing the current focus of attention of the searcher. Finally, we investigate two applications of medical foci for information retrieval (IR): (i) clustering EHS sessions based on the extracted attributes, helping us better understand foci dynamics during EHS, and (ii) predicting a searcher's likely next focus in EHS sessions given their recent foci.

The remainder of the paper is structured as follows. In Section 2, we describe related work on medical IR, with a particular focus on studies of health-related information seeking and medical search. Section 3 describes the data set that we use, including the definitions of evidence-directed and hypothesis-directed activities that we adopt, and a classifier trained to automatically label diagnostic search sessions given human labeled data. Section 4 describes our work on extracting and analyzing foci dynamics from search sessions. In Section 5, we describe work on clustering sessions to identify common patterns of EHS behavior and in predicting future foci given recent search history. We discuss our findings and their implications in Section 6 and conclude in Section 7.

## 2. RELATED WORK

Web interaction logs have been used previously to study medical Web search behavior. Bhavnani *et al.* [6] demonstrated that term co-occurrence counts for medical symptoms and disorders on Web pages can reasonably predict the degree of influence on search behavior. Spink *et al.* [23] characterized healthcare-related

queries issued to Web search engines, and showed that users were gradually shifting from general-purpose search engines to specialized Web sites for medical- and health-related queries. Ayers and Kronenfeld [1] also utilized log data on Web use, and perform a multiple regression analysis to explore the relationship between chronic medical conditions and frequency of Web use, as well as changes in health behavior due to frequency of Web use. Their findings suggest that it was not the presence of one particular chronic illness, but rather the total number of chronic conditions that determines the nature of Web use. They also found that the more frequently someone uses the Web as a source of health information, the more likely that person is to alter their health behavior. White and Horvitz [28] used a log-based methodology to study escalations in medical concerns during Web search from rudimentary symptoms to serious ailments, a behavior that they call *cyberchondria*. They found that users frequently escalated from basic medical concerns to more serious concerns based on exposure to Web content. In a follow-up study, White and Horvitz [29] also examined the occurrence of escalation from symptoms to searches for real-world medical resources, informational goals (HUIs, defined earlier). We seek to more broadly characterize the dynamics of focus of attention to a larger set of actions, including escalations from queries on basic symptoms to rare and frightening diseases and to searches for local medical expertise.

Researchers have investigated the search behavior of medical domain experts [5][15][16][31] with a view to better understanding the search behavior of those with specialist domain knowledge. Hersch *et al.* [15] review research in the medical informatics and information science literature on how physicians use IR tools to support clinical question answering and decision making. They found that retrieval technology was inadequate for this purpose and generally retrieved less than half of the relevant articles on a given topic. They follow up this review with a study of how medical and nurse practitioner students use MEDLINE to gather evidence for clinical question-answering [16]. Their findings show that these users were only moderately successful at answering clinical questions with the help of literature searching. Bhavnani [5] observed healthcare and online shopping experts while they performed search tasks inside and outside their domains of expertise, and identified domain-specific search strategies in each domain, and that such search knowledge is not automatically acquired from general-purpose search engines. Willemuth [31] performed a longitudinal study examining the tactics of medical students searching a database in microbiology. Findings showed that over the course of the study changes in students' search tactics were observed as domain knowledge increased.

Several researchers have also noted a need for improved information services to the medical information consumer. Cline and Haynes [8] present a review of work in this area that suggests that public health professionals should be concerned about online health seeking, consider potential benefits, synthesize quality concerns, and identify criteria for evaluating online health information. Eysenbach and Kohler [12] used focus groups and naturalistic observation to study users attempting assigned search tasks on the Web. They found that the credibility of Web sites (in terms of source, design, scientific or official appearance, language used, and ease of use) was important in the focus group setting, but in practice, many participants largely ignored the source of their medical information. Baker *et al.* [2] examined the effects that Web and email use has on users' knowledge about health care matters and their use of the health care system. They found that

users rarely use email to communicate with physicians and that the Web’s influence of healthcare utilization is uncertain. Sillence and colleagues [22] studied the influence of design and information content on the trust and mistrust of online health sites. Through observational study of a small number of subjects engaged in structured and unstructured search sessions they found that aspects of design appeal engendered mistrust, whereas the credibility of information and personalization of content engendered user trust. Eastin and Guinsler [10] investigated the relationship between online health information seeking and healthcare utilization such as visiting a general practitioner. Their findings suggest that an individual’s level of health anxiety moderates the relationship between online health information seeking and health care utilization decisions. Lewis [18] performed a qualitative study into young peoples’ use of the Web for health material that showed that in fact they are often skeptical consumers of the material they encounter. The findings of these studies demonstrate some of the conflicting opinions around the effect of healthcare information on human behavior. This may be attributable to differences in the goals of the studies, the samples used, and the experimental methodologies. In response to these issues, some organizations have already started efforts to improve general quality of medical information on the Web, such as the Health on the Net Foundation [14]. Others have studied augmenting Web pages or search results with tools to support credibility assessment [21].

Previous work in this area has focused on general medical search behavior, and either performed relatively simple analyses of query logs, or more in-depth studies across small focused groups. The large-scale log-based methodology that we employ enables observation of many searchers’ naturalistic information-seeking behaviors at the *session level*, rather than at the query level as in previous studies [23], affording a detailed analysis of how people perform exploratory health search on the Web.

### 3. DATA SET

We now describe the data that we used in this study and the classifier we developed to identify exploratory health search sessions.

#### 3.1 Sessions

We began with anonymized logs of URLs visited by users who consented to provide data through a widely-distributed browser toolbar. We gathered data over six months during early 2009. Log entries include a user identifier, a timestamp for each page view, and the URL of the page visited. We excluded intranet and secure (https) URL visits at the source. Only entries generated in the English speaking regions of the United States were included. From these logs we extracted search sessions on Google, Yahoo, and Microsoft’s Live Search via a methodology similar to White and Drucker [26]. Sessions comprised queries, hyperlink clicks on search engine results, and pages visited during post-result-page navigation. Sessions began with a query to one of the three search engines and terminated following a period of user inactivity of over 30 minutes, suggesting that the current search episode has ended. We filtered these logs to retain only sessions that had at least one query containing a term in a wordlist extracted from a list of symptoms from the Merck medical dictionary. To improve session coverage, synonyms of Merck symptoms were also used. Synonyms for each symptom were identified through a two-step walk on the search engine click graph with an approach similar to [4]. Statistics about the constructed data set are shown in Table 1. *Action* is used for a query or URL (page view) within a session.

**Table 1. Characterization of the data set.**

| Statistic                                  | Value          |
|--|----------------|
| Number of unique users                     | 660,219        |
| Number of sessions                         | 1,237,737      |
| Total number of actions                    | 26,983,597     |
| Average number of actions per session      | 21.8           |
| Average number of queries per session      | 10.4           |
| Min / Max / StDev num. actions per session | 1 / 968 / 27.7 |

We randomly sampled 0.1% of the sessions in the data set (1,238 instances) and annotated them manually. The annotation procedure is described in more detail in the next section.

#### 3.2 Labeling Exploratory Health Sessions

We sought to construct and use a statistical classification following the filtering based on medical keywords to remove sessions that were tagged as medical erroneously (e.g., those containing queries such as *[saturday night fever]* and *[red eye xbox]*). The goal was to classify each session into one of two classes: *EHS* (positive class), and *non-EHS* (negative class). An exploratory health search session is one where the searcher performed, at some point in the session, one or more actions indicating that they were either engaged in some evidence-directed (i.e., to understand the relevance of a set of observed symptoms) or hypothesis-directed strategy. We focused in this study on the pursuit of information on the causes or remedies for a set of observed symptoms. A session may contain activities suggesting that a searcher is engaged in one or both of these types of pursuits. We now describe the criteria that we used to manually label evidence-directed and hypothesis-directed events in search sessions.

##### 3.2.1 Evidence-Directed

During evidence-directed exploratory health search, users seek to understand the relevance of symptoms. A canonical example of an evidence-directed search is the query *[what causes dizziness and nausea?]*. Observing such a query in a log of search activity suggests that the user seeks to understand the relevance to disease of observing these two symptoms. We identified several criteria for identifying exploratory health queries as being evidence-directed. These criteria required that queries contain the following:

- combinations of symptoms via connectives, e.g., *[back pain, nausea, dizziness]*, *[dry cough with chest pain]*;
- the word “symptom” and/or the word “pain”, and;
- symptom or condition in conjunction with a specific entity, i.e., *[diarrhea in 2 year olds]*.

Although this list is not exhaustive, it provides a simple set of rules for identifying instances of evidence-directed search that can be easily applied by human judges as they inspect a session.

##### 3.2.2 Hypothesis-Directed

As displayed in Figure 1, hypothesis-directed search includes both diagnostic and informational intents. We now describe each type of informational goal.

###### 3.2.2.1 Diagnostic Intent

During the hypothesis-directed phase of diagnostic search, users seek information that can help to refine uncertainty about the hypotheses under consideration, including information that can rule out, raise or lower the likelihood of one or more hypotheses, or to identify new hypotheses. We use the presence of queries with reference to benign or serious medical conditions (using the list of conditions from [28]) as evidence of diagnostic intent.

### 3.2.2.2 Informational Intent

Hypothesis-directed searches also include the non-diagnostic pursuit of information about a condition, such as seeking details on potential hypotheses, including treatments and outcomes, e.g., pages on exercises to alleviate back pain. The primary focus of attention is obtaining information about conditions or treatment. We use the following features as evidence of informational intent:

- queries about treatment, including those with advanced syntax, e.g., [*whooping cough more:condition\_treatment*];
- use of ‘for’ in query as the connective between entities: [*ointments for eczema*];
- query contains the terms ‘treatment’ or ‘cure’;
- queries for exercises or medications, particularly in reference to a specific condition, e.g., [*back stretching exercises*];
- user performs an HUI action [29], e.g., searching for an in-world medical resource such as a clinic or hospital.

We focus on the search for treatments as a class of hypothesis-directed informational queries that could be distinguished easily from evidence-directed diagnostic search. Queries on treatments are clearly a subset of the larger space of hypothesis-directed search with informational goals, and do not address hypothesis-directed search with diagnostic goals. We pursue the basic partition of intentions into hypothesis-directed search for treatments and evidence-directed diagnostic search as an initial analysis of the study of intentions and attention in EHS.

One of authors (MAC) reviewed the 1,238 sessions, applied the criteria listed above, and assigned the labels *evidence-directed*, *hypothesis-directed*, or *both* to each of the sessions.

## 3.3 Classifying Exploratory Health Sessions

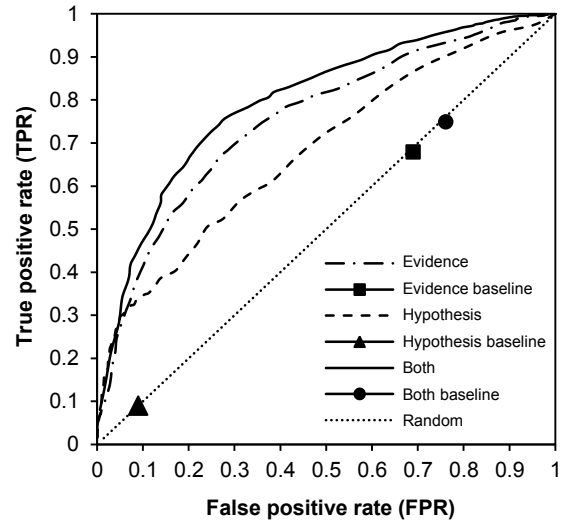
Search sessions may contain actions that are both evidence- and hypothesis-directed. We developed separate classifiers for identifying evidence-directed, hypothesis-directed, as well as the joint appearance in sessions of both phases of analysis.

Table 2 shows the features used in our classification. We identified a set of features about each of the sessions that captures elements of evidence-directed and hypothesis-directed phases (e.g., *HasCausePatterns*, *HasMedications*), as well as features that might help with selecting medical sessions and filtering out noise (e.g., *HasMedicalSites*, *HasSocialSites*). Many of the feature names listed in Table 2 are self-explanatory. Of the others, *HasListElements* is true if a query appears to contain a list (detected by splitting on “and”, “or”, or commas, and counting the number of resultant items). *HasNLElements* is true if the query contains connective stopwords, e.g., “by”, “with”, “for”, etc. *HasObservations* is true if the query contains terms such as “feeling”, “pain”, and “hurt” that may indicate an observed medical symptom. We use “Portion” as a prefix of features to denote the fraction of all actions in the session where a feature was observed.

We trained a logistic regression classifier for each class (*evidence-only* or *hypothesis-only*) as well as for *both*. We repeated the following process for each experimental run: (i) randomly permute the 1,238 labeled sessions; (ii) split the 1,238 sessions into five equally-sized sets, and; (iii) perform five-fold cross-validation, using four folds for training and the remaining fold for testing. Since logistic regression provides probabilities that each test case is an instance of the positive class, we set a threshold value ( $\alpha$ ). If the probability exceeds that threshold, we consider the instance to be in the positive class, and negative class otherwise. For any given threshold, performance is averaged over the five runs.

**Table 2. Feature names and weights for the three classifiers. Sorted in descending order by the absolute values for *Both*.**

| Feature Name         | Evidence | Hypothesis | Both   |
|----------------------|----------|------------|--------|
| HasMedicalSites      | +1.396   | −0.803     | +1.682 |
| HasObservations      | +0.992   | −0.444     | +1.135 |
| HasCausePatterns     | +0.282   | +0.901     | +0.973 |
| HasListElements      | +0.431   | +0.330     | +0.700 |
| HasBodyParts         | +0.772   | +0.129     | +0.696 |
| PortionQueries       | +0.396   | −1.384     | +0.543 |
| NumTreatPatternsURL  | −0.067   | +0.107     | +0.474 |
| HasNLElements        | +0.525   | −0.589     | +0.458 |
| PortionRepeatQueries | +0.230   | −0.128     | +0.339 |
| HasSymptomPatterns   | +0.427   | −0.523     | +0.287 |
| PortionImages        | −0.095   | −0.182     | −0.156 |
| HasMedications       | +0.200   | +0.025     | +0.129 |
| HasTreatmentPatterns | −0.060   | +1.770     | +0.123 |
| NumYoutubeViews      | −0.086   | −0.083     | −0.116 |
| HasSocialSites       | +0.068   | −0.340     | +0.114 |
| NumSearches          | −0.108   | +0.052     | −0.103 |
| NumRepeatedQueries   | +0.084   | +0.008     | +0.074 |
| PortionHealthRelated | +0.051   | −0.099     | +0.072 |
| HasAdultContent      | −0.049   | +0.196     | −0.042 |
| NumActions           | −0.002   | −0.040     | −0.003 |
| AvgPageDwellTime     | +0.002   | −0.007     | +0.002 |



**Figure 1. ROC curves (*evidence*, *hypothesis*, and *both*).**

Figure 1 shows the receiver-operator characteristic (ROC) curves as we reduce  $\alpha$  for *evidence-only*, *hypothesis-only*, and *either* classes. A threshold of 0.0 corresponds to the upper right corner of the figure, where all instances are classified as true. Conversely, a threshold of 1.0 corresponds to the lower left corner, where all instances are classified as false. Baselines are determined using the maximum likelihood estimator (MLE) of the positive class from the training data, and then using the MLE as the threshold value of a random roll for each instance. Therefore the baseline estimator uses only frequency information in the training data, and no instance-specific features. The false-positive and true-positive rates of the baseline estimator are averaged over five runs, which

results in a single point on the curve: a single (false positive rate, true positive rate) coordinate, shown for each method in Figure 1.

Table 2 shows the weights assigned by the classifier to each of features used in the classifying whether a session is evidence-directed, hypothesis-directed, or both per the definitions of these categories for the study. We focus on the sessions identified as EHS-related by the *Both* classifier for the remainder of the paper as it had the highest accuracy and allowed us to extract and study examples of hypothesis- and evidence-directed search behaviors.

The presence of medical Websites (*HasMedicalSites*) such as WebMD and PubMed is the most predictive feature of EHS sessions, followed by queries associated with pain or feelings (*HasObservations*). Other features such as *NumYouTubeViews* and the fraction of the session that are image searches (*PctImages*) are negatively correlated with the search session being medically-oriented. Features such as the average amount of time users spend on a page (*AvgPageDwellTime*) and the number of actions in a session (*NumActions*) had little predictive value for this task.

Using the *Both* classifier, we automatically labeled the original full data set described in Table 1, setting  $\alpha$  to 0.95. We did this to increase classification precision, and to help ensure that the recovered set contains few false positives at the cost of missing a portion of the EHS sessions. Following automatic labeling, we obtained 249,833 sessions (from 94,498 users) that we considered as EHS-related. Hereafter we refer to these EHS sessions as *M*, and we use them for the analysis performed in the rest of the paper.

## 4. DYNAMICS OF ATTENTION

We have focused on the automated classification of EHS sessions. This is necessary to create a subset of all sessions that we can use to study the dynamics of the focus of attention in exploratory health search. We now turn to extracting foci from those sessions and characterize focus of attention during health searches.

### 4.1 Extracting Intentional Foci

Beyond focusing on evidence- and hypothesis-directed phases of EHS, we magnified our analytic resolution to the sub-session level, to examine user behavior at each step in sessions. To better understand user activity within a single session, we define several concepts and provide a formulation for investigating the flow of intentions during exploratory health search. As mentioned above, we consider actions to only be queries submitted to a search engine. We consider sessions to be temporally ordered sets of actions. For any action, we attempt to extract a *focus* associated with that action. We shall focus here only on actions that are search queries as: (i) extracting information from the URL string itself provides only partial coverage of the remaining actions and we have no gauge of how well these extractions correlate to current intent, and; (ii) content analysis of Web pages may introduce significant noise into the process of extracting foci. Web pages containing medical content may cover many topics on a single page, raising the likelihood of making erroneous extractions of foci.

#### 4.1.1 Defining Medical Foci

During our analysis of the 1,238 labeled sessions, we identify three types of foci. These include the pursuit of information on symptoms, causes, and remedies, corresponding to evidence-directed, hypothesis-directed (diagnostic intent), and hypothesis-directed (informational intent) behaviors respectively. For the focus of attention study, we take as *causes* conditions that explain one or more symptoms. A *symptom* is a physical manifestation that is seen as a departure from normal function, even if the un-

derlying cause may not be a malady. A *remedy* is a step taken to treat a cause. Symptoms expressed in queries likely indicate evidence-directed phases, and causes input as queries likely indicate hypothesis-directed phases. Remedies in queries are indications of users seeking treatment information and are likely indications of users with non-diagnostic, hypothesis directed goals.

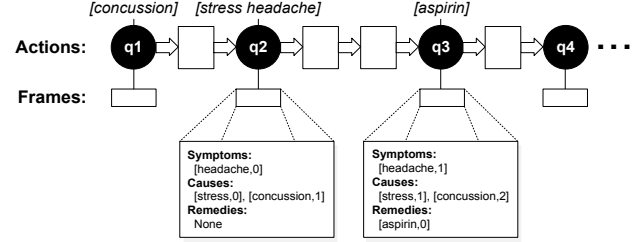


Figure 2. Generating references and frames for foci.

When we extract a focus from an action, we consider that instance to be a *reference* to that focus. We may extract any number of foci from a given action, for example a query for [*exercises for chronic back pain*] is identified as the symptom “back pain” and remedy “exercise.” We define a *frame* to be the representation of the user’s focus of attention over a single action. Each frame consists of three ordered lists, one each for symptoms, causes, and remedies, respectively. Each item in a list maintains the number of times that particular cause or symptom has been referenced thus far in the session and the number of steps since the last reference (which is zero if the item is first seen in the current frame).

Figure 2 shows an example of a medically focused search session, with queries (black circles) and page views (rectangles) as the actions. The figure shows how we extract frames from the queries of these sessions and how the frames evolve over the course of the search session as more queries are issued by the searcher.

#### 4.1.2 Identifying Medical Foci in Sessions

Identifying diagnostic foci from search sessions is challenging. We see large variations in users’ search behaviors, including how they specify their information needs and the strategies they employ in exploratory health search. To identify symptoms, causes, and remedies in the context of a medically relevant query, we use methods similar to those described earlier for labeling EHS sessions. For each focus type, we employ a set of rules, partly based on compiled dictionaries of potentially salient terms and phrases identified via a manual review of query logs. We formulate a set of regular expressions based on common patterns for each focus type that appeared in the log data. Some examples are as follows:

- Symptoms (evidence directed):
  - rules such as <body part> pain, where body part is derived from lists of major appendages, organs, etc.;
  - terms/phrases such as “ache” and “dizziness”, and;
  - expressions (e.g., query starts “pain in” or “causes of”).
- Causes (hypothesis-directed – *diagnostic* intent):
  - rules such as <body part> failure;
  - terms/phrases such as “acid reflux” and “sinusitis”, and;
  - expressions (e.g., “symptoms of” or “diagnosis of”).
- Remedies (hypothesis-directed – *informational* intent):
  - terms such as “treatment”, “clinic”, and “doctor”, and;
  - expressions (e.g., “cure for” or “treatment for”).

These rules and resources are used to extract foci automatically from all queries in all sessions in *M*. Although the generalizability of this methodology is limited, we believe that the labels can pro-

vide the basis for generating initial insights about patterns of activity in exploratory health search. We believe that there is promise in the longer term with refining the labeling methodology using of machine learning.

### 4.1.3 Extracting Medical Foci

The set of sessions in  $M$  comprised 5,292,871 actions. We labeled 370,498 (7%) of these actions as “non-medical,” meaning either the query or URL contained strong indicators of non-medical intent in the action. We used two types of indicators: (i) we maintained a list of popular socially-oriented websites that occurred often in our log data (e.g., Facebook, MySpace), and excluded any actions with a URL from any of these domains, and; (ii) we also maintained a keyword list containing words that indicated a completely different intent than medical search. For example the word “lyrics” was in this list, meaning a query such as “headache lyrics” would be excluded. An initial list was generated and this list was updated during the analysis as false positives emerged. An examination of a random sampling of excluded actions revealed that no actions were falsely excluded.

Of the remaining actions, 2,329,231 (44%) contain queries that we consider as serving as strong signals of search intent from the user, from we may be able to directly extract medical foci. Of this number, we are able to extract at least one cause focus from 652,180 (28%), at least one symptom focus from 1,213,529 (52%), and at least one remedy focus from 92,169 (4%). 141,660 (49%) of all EHS search sessions in  $M$  contain at least one non-medical query. Non-medical queries are ignored for our analysis. The remaining 2,593,142 (49%) of actions in  $M$  are URL visits that may offer evidence of medical intent. However, as stated earlier, we only use queries in this work.

## 4.2 Characterizing Foci Dynamics

We now present an analysis of the characteristics of the foci extracted from  $M$ , specifically examining the three focus categories and the transitions among those categories during sessions. We begin by looking at how health search sessions start, how they end, and transitions that occur within them.

### 4.2.1 Starts, Ends, and Transitions

Table 3 displays how often each of the medical foci types appears at the outset and conclusion of sessions. The table also shows the percentages for each of the start and end points.

**Table 3. Start and end states of health search sessions.**

| Session starts |              |               | Session ends   |              |               |
|----------------|--------------|---------------|----------------|--------------|---------------|
| <i>Symptom</i> | <i>Cause</i> | <i>Remedy</i> | <i>Symptom</i> | <i>Cause</i> | <i>Remedy</i> |
| 157,433        | 62,785       | 7,528         | 164,218        | 68,312       | 9,346         |
| 69.1%          | 27.6%        | 3.3%          | 66.5%          | 29.7%        | 3.8%          |

The EHS sessions identified primarily begin and end with symptoms, highlighting the prominence of evidence-directed reasoning in the data. Hypothesis-directed foci (causes and remedies) tend to occur toward the ends of EHS sessions, suggesting a shift between evidence-directed and hypothesis-directed phases over the course of sessions. We shall examine such transitions later in this section. We note that the total number of references within and across foci types increases from session start to session end. This finding suggests that the average density of types (the number of different foci in one session start/end) increases. There could be many reasons for this including exposure to Web content stimulating users to become more medically focused as the session proceeds [28].

We were interested in the number of focus types that occur in a given session. In most sessions in  $M$  (68%), searchers stay with the same focus type for the duration of the session, even though the specific symptom or condition being pursued might change. The distribution of focus categories is shown in Table 4. Focus categories can include multiple foci, e.g., *SymptomCause* means that the focus is on symptoms and causes together for the duration of the session (e.g., query for [*chest pain heart attack*], [*chest pain indigestion*], etc.). *Other* refers to queries that were identified as being non-medical.

**Table 4. Number of sessions with single focus of attention.**

| Focus type         | Count   | % sessions |
|--------------------|---------|------------|
| Symptom            | 125,232 | 63.8%      |
| SymptomCause       | 44,365  | 22.6%      |
| Cause              | 10,433  | 5.3%       |
| Other              | 7,564   | 3.9%       |
| SymptomRemedy      | 3,879   | 2.0%       |
| Remedy             | 2,241   | 1.1%       |
| SymptomCauseRemedy | 1,641   | 0.8%       |
| CauseRemedy        | 829     | 0.4%       |

The findings of our analysis shows that sticking with symptoms is particularly common. Non-medical sessions (labeled as *Other*) were initially labeled as EHS sessions by the classifier described in Section 3, but were not found to be associated with medical intent per the rules used in the focus extraction. One explanation for the difference in label is that the classifier was based on all actions in the session (including URLs visited, e.g., the *HasMedicalSites* feature was important in EHS session labeling), whereas the focus extraction phase was based only on query actions.

92,934 (32%) health search sessions contained a transition among different focus types. A transition involves the expansion of an existing type to include another (e.g., symptom  $\rightarrow$  symptom + cause), or a complete change in the type (e.g., symptom  $\rightarrow$  cause). In Table 5, we summarize the frequency of each transition type.

**Table 5. Frequency of occurrence of types of transitions among categories.**

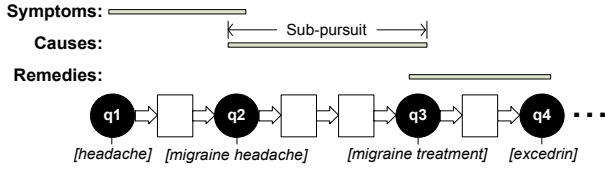
| Transition type              | Count  | % transitions |
|------------------------------|--------|---------------|
| Symptom $\rightarrow$ Cause  | 40,618 | 43.4%         |
| Cause $\rightarrow$ Symptom  | 35,779 | 38.6%         |
| Symptom $\rightarrow$ Remedy | 6,628  | 7.2%          |
| Remedy $\rightarrow$ Symptom | 4,240  | 4.7%          |
| Cause $\rightarrow$ Remedy   | 3,559  | 3.7%          |
| Remedy $\rightarrow$ Cause   | 2,110  | 2.3%          |

The statistics on transitions show that transitions between symptoms and causes are most common. Searchers’ exposure to Web content may lead them to link their perceived symptoms with causes. We found that transitions from causes to symptoms occurred almost as frequently as transitions from symptoms to causes. This implies the existence of oscillations or loops among these categories, perhaps in the spirit of the hypothetico-deductive cycle described earlier. We examine such cycling later in the paper.

### 4.2.2 Pursuits

We now move to characterizing sequences of foci, and define a *pursuit* as a series of consecutive queries within a search session. There are two types of pursuit: (i) *full pursuits* over all session actions, and (ii) *sub-pursuits* on particular foci types (e.g., symp-

toms) over a subset of actions in the session. A typical set of pursuits for a search session is illustrated in Figure 3. The duration of each sub-pursuit is marked above its session actions. For completeness, the figure contains both queries and page views, but our analysis was performed using only queries.



**Figure 3. Pursuits and sub-pursuits within a search session.**

We sought to characterize the differences in the length of pursuits by focus type. A better understanding of the length of time that searchers pursue on each condition can be useful for understanding differences in search strategies for different foci or for predicting when transitions among pursuits are going to occur. Table 6 shows the mean average number of queries associated with pursuits for each type. For example, the average length of a *SymptomCause* pursuit (symptoms and causes) was 8.33 queries.

**Table 6. Average pursuit length for each focus type.**

| Focus type         | Pursuit length |
|--------------------|----------------|
| SymptomCauseRemedy | 13.46          |
| CauseRemedy        | 11.98          |
| Symptom            | 11.69          |
| SymptomRemedy      | 10.43          |
| Remedy             | 10.27          |
| SymptomCause       | 8.33           |
| Other              | 8.20           |
| Cause              | 7.91           |

As shown in Table 6, the average lengths depending on the focus type. We find that the pursuit lengthens with increases in the number of focus types included in the pursuit. We note that when remedy foci are involved, the pursuit length tends to be longer than in pursuits without remedy foci. For example, symptom/cause pursuits average a little over eight steps in length. If we add remedies to the pursuits, the number jumps to over 13. Indeed, if we focus only on the length of the sub-pursuits for each of the three main focus types (for sessions with at least one transition so we can focus on sub-pursuits), we see that remedy-related pursuits contain more user actions on average (Table 7).

**Table 7. Average sub-pursuit length for each focus type.**

| Focus type | Sub-pursuit length |
|------------|--------------------|
| Symptom    | 4.49               |
| Cause      | 5.08               |
| Remedy     | 5.94               |

The pursuit of remedies may indicate a desire to address a known condition. We speculate that searchers appear to be more engaged in that task than in symptom or cause searching, or may experience more difficulty in finding relevant search results and hence need to revise their query statements more frequently.

#### 4.2.3 Oscillations

As mentioned earlier an oscillation instance involves the transition from one focus type to another and then back to the original type (e.g., *Symptom→Cause→Symptom* describes an oscillation instance where the focus is symptom, then cause, and then returning

to symptom). We mined the foci extracted from our EHS sessions, extracted oscillations, and obtained the distribution among oscillation types (summarized in Table 8). Such oscillations were observed in the EHS sessions of 16.2% of the users in *M*.

**Table 8. Frequency of occurrence for oscillation types.**

| Oscillation type       | Count  | % oscillations |
|------------------------|--------|----------------|
| Symptom→Cause→Symptom  | 10,034 | 51.4%          |
| Cause→Symptom→Cause    | 7,490  | 38.4%          |
| Symptom→Remedy→Symptom | 989    | 5.1%           |
| Remedy→Symptom→Remedy  | 526    | 2.7%           |
| Cause→Remedy→Cause     | 301    | 1.5%           |
| Remedy→Cause→Remedy    | 171    | 0.9%           |

As noted in our earlier analysis, transitions between causes and symptoms occur often, and more frequently than any other type of transition. This evidence supports the hypothesis that searchers engage in forms of hypothetico-deductive searching as part of diagnostic search. We now explore this hypothesis in more detail.

#### 4.2.4 Diagnostic Cycles

As discussed earlier, the hypothetico-deductive method [25] has been proposed as a description of an iterative process of medical diagnosis involving a cycling between evidence-directed and hypothesis-directed phases, with an ongoing revision of the diagnoses under consideration and information being pursued. We sought to better understand potential cycles of revision in diagnostic search. We were particularly interested in *cycles* where users were observed to repeat a search for symptoms but pursue different causes or remedies with the progression of a session. Although the average number of new foci per session is around one (suggesting that if users do oscillate, they do not do so repeatedly), we find evidence that diagnostic cycles occur. Table 9 presents the frequencies with which we observed cycles with different numbers of iterations. A cycle with two iterations means that the user searched for a symptom, a cause, then returned to the original symptom, and proceeded to search for a different cause.

**Table 9. Frequency of cycles with different iteration counts.**

| Num. iterations | Count | % iterations |
|-----------------|-------|--------------|
| 2               | 2,122 | 88.4%        |
| 3               | 229   | 9.5%         |
| 4               | 43    | 1.8%         |
| 5               | 4     | 0.2%         |
| 6               | 3     | 0.1%         |

Cycles were observed in the EHS sessions of 3.1% of the users in *M*. An example of such a cycle taken from the logs is a user searching for [stomach pain], who first considered [irritable bowel syndrome], then [appendicitis], before settling on [gallstones] and proceeding to explore treatment options ([gallstone surgery], [cholecystectomy], etc.). Although our rule-based detection mechanism is limited (e.g., we identify users returning to the exact same symptom (and associated synonyms) following each iteration), we find that diagnostic cycles appear in Web search, and are likely to occur even more frequently than we found with our limited methodology for identifying actions.

## 5. CLUSTERING AND PREDICTING

We now describe two applications that demonstrate the potential value of harnessing representations of searcher focus of attention to understand and support exploratory health search behaviors.

The two applications are (i) clustering EHS sessions to better understand the health-related search strategies that users employ, and (ii) predicting future foci given a recent history of exploratory health searching. These applications seek to identify common patterns of exploratory health search sessions.

### 5.1 Generalizing Sessions for Analysis

Extracting the foci from various sessions provides a microscopic view of EHS search behavior. However, to represent sessions in a way that allows comparability among the sessions, we seek to create a more macroscopic view of user behavior. To do this, we represent an EHS session in  $M$  as a character sequence and depict each action in the session with a character. Such a mapping affords easy manipulation and analysis, and also removes identifying information, further preserving user privacy without destroying the salient aspects of search behavior that are necessary for sequence clustering and prediction. Downey *et al.* [9] introduced formal models and languages that encode search behavior as character sequences, with a view to comparing search behavior in different scenarios. In a similar manner, we formulated an alphabet designed to represent the focus of attention at each query in the session. The alphabet appears in Table 10. Each symbol represents a set of foci, e.g., ‘A’ is used to represent a query that has both symptom and cause elements, such as [headache migraine].

Table 10. Symbols assigned to queries based on medical foci.

| Symbol | Focus type         |
|--------|--------------------|
| S      | Symptom            |
| C      | Cause              |
| R      | Remedy             |
| A      | SymptomCause       |
| B      | SymptomRemedy      |
| D      | CauseRemedy        |
| E      | SymptomCauseRemedy |
| N      | Non-medical        |

If a user searched for symptoms across multiple queries, pursued a cause, returned to the symptom, and ended up querying for a remedy, the session would be represented as “SSSCSR.” We created a sequence representation for each of the sessions in  $M$ , hereafter referred to as  $M_S$ . We use these sequences for the analysis presented in the remainder of this section. For the purposes of this analysis, we exclude the non-medical actions (‘N’ in Table 10) as they were not pertinent to either of our application scenarios. We now describe the first application: session clustering.

### 5.2 Clustering Exploratory Health Sessions

Clustering exploratory health search sessions based on the sequences described in the previous section promises to be useful for understanding foci dynamics. Clusters emerging from this analysis could help inform the design of tools to support medical search. We clustered the sessions using the approach described by Cadez *et al.* [7], which was used to cluster users’ browsing categories of a popular news website (msnbc.com) by learning a mixture of first-order Markov models using the Expectation-Maximization procedure. We ran the clustering algorithm (which has a built-in model-selection procedure) across all character sequences in  $M_S$ . This yielded a set of 12 clusters that represented different aspects of diagnostic search strategies. Since the clusters were based on sets of symbols rather than the detailed search strategies themselves, there was some redundancy across the set of clusters generated. For example, although one cluster com-

prised only sessions with all symptoms, and another cluster comprised sessions with all causes, both had consistent foci *types* that were unchanged over the course of the session. To remove redundancy and to simplify our analysis of the clusters, we manually grouped the clusters into five distinct cluster categories. The goal of the re-grouping was only to remove redundancy, and no other manipulation of the clusters was performed during this procedure.

We assigned names to each of the categories and list them below. For each category, we indicate whether the pursuit is evidence-driven, hypothesis-driven, or some combination of the two. We also show the percentage of all sequences in  $M_S$  that each category comprises, and provide a canonical example to clarify the nature of the search strategies observed in each category:

1. **Constant (mostly evidence-directed) (75%):** Foci (cause, symptom, or some combination, sometimes remedy) remain the same for the entire session. For example, a user searching for symptoms of headache persists with headache-related symptom searches over all queries in the duration of their session.
2. **Incremental (evidence-directed) (10%):** Assuming a natural progression from symptom, to cause, to remedy, searches transition among each of these foci types in that order. For example, a searcher first searches on headache symptoms, then headache causes, and then headache medication.
3. **Incremental (hypothesis-directed) (7%):** Searches do not conform to the natural progression described above, and instead progress in opposite direction (i.e., from remedies/causes to symptoms). For example, searches may start with a possible explanation for experienced symptoms and end with the searcher seeking more information about these symptoms.
4. **Oscillating (both evidence- and hypothesis- directed) (5%):** Searches alternate among categories. For example, a searcher may seek causes for a symptom and then search for a different symptom.
5. **Cyclic (evidence- and hypothesis- directed) (3%):** Searches anchor on symptom and explore related causes and remedies. For example, search for a symptom, identify a cause, return to the same symptom, then search for an alternative cause.

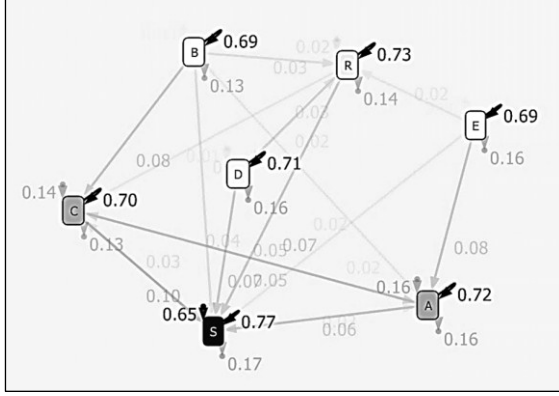
Through sequence clustering, we could summarize and distill exploratory health search strategies into five key elements. Such analyses can provide search engines with insight about the evolving informational goals and intentions of searchers pursuing health information, and to ultimately inform the design of more effective search engines.

### 5.3 Predicting Foci in EHS Sessions

Rather than considering the character sequence holistically across a session, we can consider transitions among foci. We can seek to predict whether a searcher will stick with a current set of foci or to transition to new foci. Being able to predict such transitions in real-time based on interaction history could enable search providers to offer proactive support or intervene if transitions to causes are anticipated (especially if those causes could be identified as inappropriately escalatory [30]).

We generated a state transition diagram, displayed in Figure 4, using all sequences in  $M_S$ , given the frequency with which foci-foci transitions are observed. The symbols in each of the nodes correspond to Table 10. The darker the node, the more occurrences there are of that state (foci) in our data. Also shown are the empirically based probabilities of remaining in each state, and the probability of starting or ending an EHS session in each state.





**Figure 4.** State transition diagram across all sequences in  $M_S$ .

The transition diagram illustrates that once users get into a particular state they do not usually leave. The diagram also shows that although all states have a similar chance of being the last state in the session, EHS sessions primarily start with symptoms (65% of the time), causes (14%), or a combination of symptoms and causes (16%). It is less common for medical sessions to begin with a remedy-oriented query (also supported by the data in Table 3). From these statistics, it also seems that evidence-driven pursuits occur more frequently than hypothesis-driven search.

We applied the sequence prediction algorithm from [7] to the character sequences in  $M_S$ . The task was to predict the next symbol in sequence given the previous sequence. We used five-fold cross validation and varied the size of the window  $w$  containing the sequence of symbols preceding the predicted symbol (foci), and used this window in both training and testing. We measured the accuracy with which we were able to predict the next state, given the previous  $w$  symbols in the sequence, averaged across 100 randomized runs and within each run across the five folds. The findings show that we are able to predict the next state with accuracies ranging from 86% ( $w=2$ ) to 92% (when  $w=8$ ). A default strategy of always predicting that users remain in the same state as the last action in  $w$  has an accuracy of 74%. Performance degraded when  $w$  exceeded eight steps, perhaps because of the diminished amount of training data.

## 6. DISCUSSION AND IMPLICATIONS

We presented log-based analyses of the dynamics of focus of attention in exploratory health search. The analyses provide views into the behaviors of people searching for health information on the Web. We first presented a schema for categorizing the informational goals behind exploratory health search, touching on informational and diagnostic intentions, and on evidence- and hypothesis-directed strategies. We then reviewed methods for classifying sessions as being medically focused, studied the dynamics of foci (symptoms, causes, remedies) in those sessions, and highlighted patterns of activity demonstrated by searchers, such as the prevalence of symptom-oriented search, and the tight interplay between searchers for symptoms and causes when transitions are observed. We identified the existence of cycles and oscillations among foci that indicate that users may employ patterns of medical search that can be viewed as a rough approximation to hypothetico-deductive reasoning, using the Web as a source of information, in lieu of probabilistic inference.

Our analyses also reveal that users tend to remain in a single state during exploratory health search. Further study is needed to de-

termine whether such search behavior is associated with EHS or is simply an indication of the persistence of a goal in the absence of satisfaction. The log-based nature of the study means that we cannot validate searchers' true intentions. Beyond consumer health search, sessions under consideration in our study likely also included professional medical searches including those by practitioners and students among numerous user cohorts searching for different purposes. Another concern is that since users use various search engines, sites, and other non-Web resources as they perform exploratory health searches, our approach may be unable to detect all of the nuances of such searching amidst the larger information context and motivations of searchers. We are interested in further investigating the underlying causes and intentions behind the observed behaviors via user studies.

The work presented here can be extended in several ways. First and foremost, we note that the current implementation uses a restricted set of ad hoc rules to label medical sessions, as well as to extract foci. We believe that the rules and terms being considered can be expanded and that potential biases in the existing set need to be explored and characterized. Machine-learning methods could be used to learn features of foci and to more robustly label sessions and identify foci in EHS sessions. Our analyses were also limited to within-session analysis. Although this scope provided a significant cache of information on searchers' activities, we believe that extending the analysis beyond the boundaries of a single session will provide even richer information concerning the user's intentions, enabling us to form a more accurate model of their reasoning process. For example, our earlier studies in this area revealed trends in how users re-access medical information over time [28], and significant benefit from multi-session features in predicting real world healthcare utilization [29]. In an attempt to reduce noise, we solely considered queries in search sessions. However, we know that a significant portion of the actions in each session were post-query browsing of web pages. These actions include navigating to pages on social networking sites that support groups for specific diseases. Future work will extend our analysis to include these and other post-query browsing actions.

The clustering and prediction tasks suggest directions for potential uses of information on the dynamics of focus of attention. The clustering provides a way to group individual searches so as to identify predominant exploratory health search strategies. Search systems or medical websites could make use of this kind of information to better understand how people perform health search. For example, a search system may expand queries using symptoms that a user has previously searched on so as to focus the search (potentially while sharing this strategy with the user) on the larger set of symptoms. Likewise, our ability to predict user actions (at a coarse level, in the form of foci) suggests that a search service might accurately anticipate shifts and more complex patterns of the users focus of attention during the course of a session. Such competencies demonstrate the potential utility of extracting foci from raw search logs, and have several implications in the design of search systems for medically-focused search. The use of representations of foci can provide useful handles on the structure and flow of a user's search session; search engines could provide an advanced search interface that exposes additional controls for users. Prediction could be used to anticipate forthcoming states of a user and search, such as an understanding of when a searcher might soon become inappropriately anxious based on the content of search results [28], enabling a system to act to provide guidance to users. Search systems could support medical search-

ers by pre-fetching results of possible interest or suggesting useful search strategies. In other approaches, search engines could be linked to more formal diagnostic decision support based on probabilistic and decision-theoretic inference [17].

## 7. CONCLUSIONS

We explored patterns of behavior of users searching for health information online. We sought insights on varying intentions and on the dynamics of focus of attention in search. As part of this work, we demonstrated that it is possible to extract search sessions that contain medically oriented searches. We demonstrated a simple mechanism to mine foci from a set of sessions that were extracted automatically. These sessions were used in analyses of user behavior during exploratory health search. The analyses included the generation of encodings of foci sequences that were used to discover common search strategies that people employ during health search. We showed that we were able to predict transitions among different focus types based on a recent sequence of user actions. There is much to be done to refine and extend the methods that we have presented. However, we are heartened by the initial results and believe that continuing this line of research will ultimately lead to new methods and tools that assist people with finding answers to their questions about illness and health.

## REFERENCES

- [1] S. Ayers and J. Kronenfeld. (2007). Chronic illness and health-seeking information on the Internet. *Health*, 11(3): 327–347.
- [2] L. Baker, T.H. Wagner, S. Singer, and M.K. Bundorf. (2003). Use of internet and e-mail for health care information. *J. Amer. Med. Assoc.*, 289(18): 2401–2406.
- [3] M.J. Bates. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13 (5): 407–424.
- [4] D. Beeferman and A. Berger. (2000). Agglomerative clustering of a search engine query log. *KDD*, 407–416.
- [5] S.K. Bhavnani. (2002). Domain-specific search strategies for the effective retrieval of healthcare and shopping information. *SIGCHI*, 610–611.
- [6] S.K. Bhavnani, R.T. Jacob, J. Nardine, and F.A. Peck. (2003). Exploring the distribution of online healthcare information. *SIGCHI*, 816–817.
- [7] I.V. Cadez, D. Heckerman, C. Meek, P. Smyth, and S. White. (2001). Model-based clustering and visualization of navigation patterns on a web site. *MSR-TR-00-18*.
- [8] R.J. Cline and K.M. Haynes. (2001). Consumer health information seeking on the Internet; the state of the art. *Health Education Research*, 16(6): 671–692.
- [9] D. Downey, S.T. Dumais, and E. Horvitz. (2007). Models of searching and browsing: Languages, studies and applications. *IJCAI*, 2740–2747.
- [10] M.S. Eastin and N.M. Guinsler. (2006). Worried and wired: effects of health anxiety on information-seeking and health care utilization behaviors. *Cyber. and Behav.*, 9(4): 494–498.
- [11] A.S. Elstein, L.S. Shulman, and S.A. Sprafka. (1978). *Medical Problem Solving: An Analysis of Clinical Decision Making*. Harvard University Press.
- [12] G. Eysenbach and C. Kohler. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative studies using focus groups, usability test, and in-depth interviews. *Brit. Med. J.*, 324: 573–577.
- [13] G.A. Gorry and G.O. Barnett. (1968). Sequential diagnosis by computer. *J. Amer. Med. Assoc.*, 12(205): 849–854.
- [14] Health on the Net Foundation. [www.hon.ch/HONcode/Pro](http://www.hon.ch/HONcode/Pro).
- [15] W.R. Hersh and D.H. Hickam. (1998). How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review. *J. Amer. Med. Assoc.*, 280: 1347.
- [16] W.R. Hersh et al. (2002). Factors associated with success in searching MEDLINE and applying evidence to answer clinical questions. *J. Amer. Med. Inform. Assoc.*, 9: 283–293.
- [17] E.J. Horvitz, D.E. Heckerman, B.N. Nathwani, and L.M. Fargan. (1984). Diagnostic strategies in the hypothesis-directed pathfinder system. *First Conference on Artificial Intelligence Applications*, 630–636.
- [18] T. Lewis. (2006). Seeking health information on the Internet: lifestyle choice of bad attack of cyberchondria? *Media, Culture, and Society*, 28(4): 521–539.
- [19] Marchionini, G. Exploratory search: From finding to understanding. *Communications of the ACM* 49, 4, (2006), 41–46.
- [20] K. Purcell, L. Rainie, A. Mitchell, T. Rosenstiel, and K. Olmstead. (2010). Understanding the participatory news consumer. *Pew Internet and American Life Project*, March 1.
- [21] J. Schwarz and M.R. Morris. (2011). Augmenting web pages and search results to support credibility assessment. *SIGCHI*, 1245–1254.
- [22] E. Sillence, P. Briggs, L. Fishwick, and P. Harris. (2004). Trust and mistrust of online health sites. *SIGCHI*, 663–670.
- [23] A. Spink, Y. Yang, J. Jansen, P. Nykanen, D.P. Lorence, S. Ozmutlu, and H.C. Ozmutlu. (2004). A study of medical and health queries to Web search engines. *Health Information and Libraries Journal*, 21: 44–51.
- [24] S.M. Weiss, C.A. Kulikowski, S. Amarel, and A. Safir. (1978). A model-based method for computer-aided medical decision-making *Artificial Intelligence*, 11(1-2): 145–172.
- [25] W. Whewell. (1837). *History of the Inductive Sciences*. Parker and Deighton. Cambridge, UK.
- [26] R.W. White and S.M. Drucker. (2007). Investigating behavioral variability in web search. *WWW*, 21–30.
- [27] R.W. White and E. Horvitz. (2009). Experiences with web search on medical concerns and self-diagnosis. *AMIA*, 696–700.
- [28] R.W. White and E. Horvitz. (2009). Cyberchondria: studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems*, 27(4), Article: 23.
- [29] R.W. White and E. Horvitz. (2010). Web to World: Predicting transitions from self-diagnosis to the pursuit of local medical assistance in web search. *AMIA*, 882–886.
- [30] R.W. White and E. Horvitz. (2010). Predicting escalations of medical queries based on web page structure and content. *SIGIR*, 769–770.
- [31] B.M. Wildemuth. (2004). The effects of domain knowledge on search tactic information. *JASIST*, 55(3): 246–258.