Conversational Framework for Web Search and Recommendations

Saurav Sahay and Ashwin Ram

ssahay@cc.gatech.edu, ashwin@cc.gatech.edu
College of Computing
Georgia Institute of Technology
Atlanta, GA

Abstract. In this paper, we describe a Conversational Interaction framework as an innovative and natural approach to facilitate easier information access by combining web search and recommendations. This framework includes an intelligent information agent (Cobot) in the conversation that provides contextually relevant social and web search recommendations. This setup leverages the information discovery process by integrating web information retrieval along with proactive connections to relevant users who can participate in real time conversations. We describe the conversational framework and report some preliminary experiments in the system.

1 Introduction

The medium of online conversation allows for sharing ideas, asking questions or discussing issues and solutions interactively along with others. It is an age-old communications practice that helps cultivate creativity, exploratory ideas, perspectives and experiences to take better decisions individually or collectively in the process. Several problems persist with using existing search tools as a means of learning, investigating or exploring about some complex and open-ended information topic. Collaborative social search involves different ways for active involvement in search related activities such as use of social network for search, use of expertise networks, involving social data mining or crowdsourcing to improve the search process.

Social psychologists have experimentally validated that the act of social discussions have facilitated cognitive performance[16]. People in social groups can provide solutions (answers to questions), pointers to databases or other people [1][3], validation of ideas[2], can serve as memory aids[5] and help with problem reformulation.

The goal, we envision, is to move search from being a solitary activity to being a more participatory activity for the user using natural dialogue conversations mixing social search with traditional web search techniques. The search agents perform multiple tasks of finding relevant information and connecting the users together; participants provide feedback to the system during the conversations that allows the agent to provide better recommendation temporally in the conversation. This framework is different from classical IR or Question Answering (QA). The focus of classic IR systems is on retrieving relevant documents from a large document collection in response to a query.

While QA deals with more complex understanding of natural language queries, it does not involve a back and forth interaction to continuously monitor, adapt and explore in continuum about some information or questions. This Conversational approach helps users search, explore and ask questions in natural language, leaving the task of user intent comprehension on the system, while the conversational search agents bring together people and different artifacts like documents, facts and opinions together in the conversation to provide a knowledge-rich participatory atmosphere. Cobot uses technology for operationalizing a user's intent into computational form, dispatching to multiple, heterogeneous services, gathering and integrating results, finding people in the community who best match the ability to respond to user's request and presenting them to the user as a set of solutions to their request. This conversational framework process involves a series of dialogue interactions, agent recommendations and feedback activities.(Figure 1)

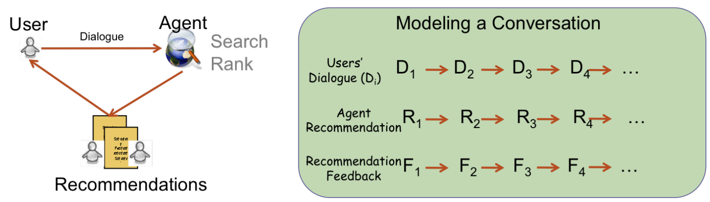

Fig. 1. Modeling Conversational Search

2 Framework

Figure 2 gives a high level architecture of the Conversational Interaction framework. The framework is built around constructs to support memory update and access, categorization and learning in the system. The framework allows for the ability to start conversations, get connected to people and get relevant information for the information need in context.

While developing the Conversational Interaction framework, we are adhering to some guiding principles which are as follows:

- Cobot is an Information Agent with Memory, Categorization and Learning modules to remember, understand and improve recommendations over time for the user.
- Different conversation facets (topic, message, asker, presence, time of asking) should have different metrics for comparison to provide for search criteria beyond query-relevance
- Ability to reformulate relevant queries from conversational sentences and paragraphs
- Ability to understand the progression of conversation context to determine suitable interference points.
- Critique based feedback in search results (eg. ability to like different facets) to support personalization of results

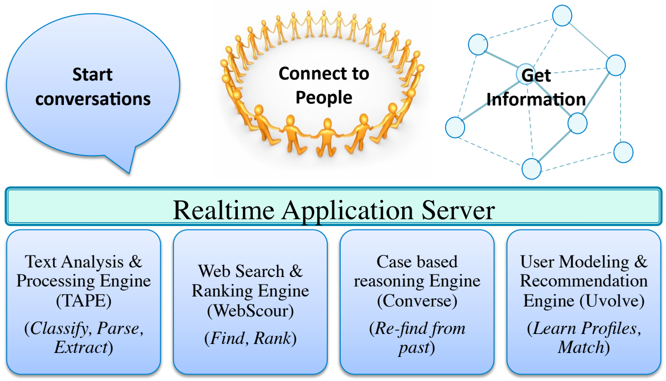

Fig. 2. System Architecture

– Support for quick access to past conversations (Ability to re-find information)

Some differences between searching conversations and traditional web search can be attributed to factors like chronological ordering of conversations, lots of coreferences and informal nature of the language. Traditional text ranking algorithms like BM25[9] might not work due to factors like short length of these conversations.

Text Analysis and Processing Engine(TAPE) processes conversations, pushing it through the various steps of analysis, processing and storage within the system. The current system is being designed and developed for health domain and engages in it the use of medical ontologies coupled with natural language processing components. TAPE (Figure 3) produces and maintains the knowledge representation by processing information from agent's working memory of conversation, user models and knowledge-bases. The agent's task is to use the sub-modules for extracting meaningful queries from conversations, classifying messages into relevant categories, and calling the right combination of algorithms for retrieving candidate recommendations.

2.1 Memory

Language and interaction (percepts) creates usable memories, useful for making decisions about what actions to take and what information to retain. Cobot framework (we interchangeably use the terms Cobot and the Conversational Interaction framework) leverages these interactions to maintain users' episodic and long term semantic models, agent's per conversation working memory of topics, users and messages (Figure 4). The agent analyzes these memory structures to bring in external recommendations into the system by matching with the contextual information need(Categorization). The social feedback on the recommendations are registered in the indices for the algorithms to generate their contextual relevance. Paper [13] also describes the architecture of Cobot System in more detail.

The purpose of Episodic Memory is to capture the user's short-term interactions and interests. Based on user's frequency of interactions and diversity in topics, this memory empirically varies in the range of a few days for different users. The Semantic

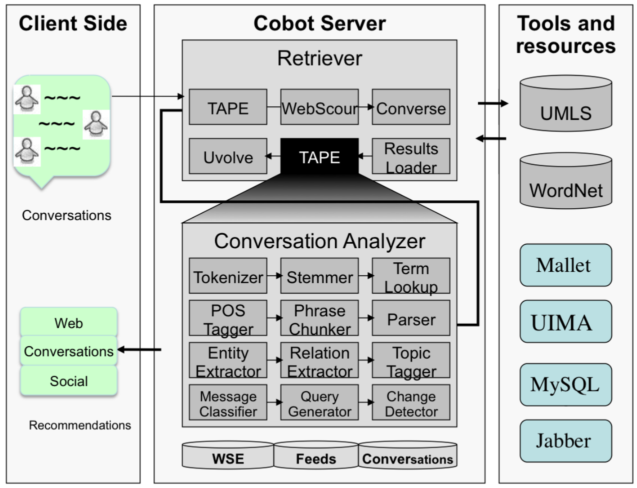

Fig. 3. TAPE Engine

Memory captures the user's long-term profile. These are the topics that interest the user in general and for a prolonged time. These interests change less frequently and represent general criteria of recommendation to the user. Many times, users might be interested in some temporary information need. Such information need not be incorporated in the long term user memory. The episodic memory captures such short-term interests. The episodic memory forms a sort of staging area and the concepts from this memory are selectively and periodically moved to the semantic memory in a crossover process.

The nodes of the semantic memory are concepts extracted from user's interactions. The concepts are connected with associations which develop when concepts co-occur frequently. Over a period of time when the user participates in more interactions, new concepts are added to the semantic memory. Episodic Memory is represented as a Case based Reasoning like kNN system. Short term interaction episodes containing frequent concepts from conversations with interaction feedback are stored. We also call this episodic store as our Level 1 (L1) memory. This memory is searched first during the recommendation stage to prune the search space to a smaller size. Semantic Memories for this smaller search space (Level 2 or L2 Memory) are searched next to refine the ordering of recommendations and find the best matches.

2.2 Categorization

The next important step in the development of an information agent is to enable it with constructs to identify important signals from the conversations, classify them in the right schemas and group them together to further aid in generating good recommendations. In order to test some of Cobot's algorithms, we crawled WebMD forum that consists of posts and responses on different health topics. The crawler extracted all posts dating back up to one year or 20 pages of posts for each subforum. The data so extracted

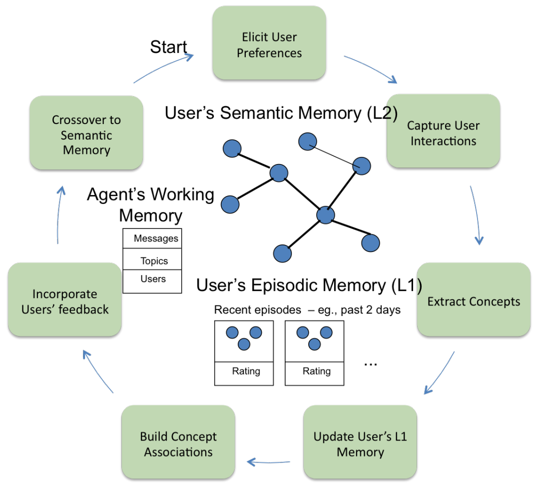

Fig. 4. Memory Structures

includes forums, subforums, conversations, users and their ratings. We extracted more that 64000 conversations from WebMD forums.

Here's an annotated sample post and one of its responses that are typical of the dataset. Bold face maps medical concepts and extracted relationships (highlighted in bold italicized). The method for extracting terms and relationships is described in detail in this paper[12].

Post (**AskQuestion** category): *Has anyone experienced* **cystic acne** *appearing once you started taking* **Adderall***? I have found that when I take my* **daily dose***, by the end of the day a cystic-like* **pimple** *has appeared on my face. If I skip a* **daily dose** *or two of* **medication***, I don't have any real* **acne** *issues. I am 42 years old and have had* **acne** *before taking* **Adderall***. But I have never had these large painful bumps. Can anyone help me???*

Response (**SuggestSolution** category): *I don't think it's the* **medication***. I've had* **cystic acne** *for a long time - including years before I started taking* **ADD Meds***. It's linked to two things. My time of the month and* **STRESS***.* **AD/HD stimulants** ***can increase*** **stress***. Instead of an* **antidepressant***, like some people have, get a* **beta blocker***. You don't get* **sleepy***. I also don't think it's* **depression** *that people get with the* **meds***, it's the* ***anxiety*** **which can cause** ***depression*** *like* **symptoms***.*

Conversational interactions are classified into one of the following categories in Cobot to strategize for query reformulation stage and to help make the decision if the agent should insert some type of recommendation into the conversation:

– *ASK QUESTION*: Asking a question, e.g. somebody posts a problem. This is usually, but not always, the first post of a thread.

– *DITTO*: Repeating a question, e.g. “Yes, I also have the same (or a very similar) problem”.
– *ASK CLARIFICATION*: Asking for more details about the problem, e.g. “Can you please provide more details?”
– *FURTHER DETAILS*: The person who is facing a problem provides more detailed information about it, possibly after somebody asks for more details.
– *SUGGEST SOLUTION*: Suggesting a solution
– *EXPRESSIVE* (Thanks for suggestion/solution, complaints about suggestion/solution, reject/accept solution)
– *OTHER* (Not fitting the above categories)

Choice of Features

The choice of features to predict the type of message labels is extremely important to get good results for this problem. In most text classification problems, a simple ‘bag of words’ approach is taken to populate the vector space of features. These features are statistically extracted using techniques like ‘term frequency - inverse document frequency’ (TFIDF) or z-score method. These statistical features make the space of possible feature set extremely large thus requiring huge training data to come up with good decision boundaries for classification of data into the right categories. In contrast, we have used a mix of syntactic and semantic features for our data exploiting medical ontologies like UMLS (Unified Medical Language System) and WordNet. We extracted the following features for the Message Classification problem:

– Position of the message thread
– Length of message
– Number of responses of the user for that forum
– Emotive Features (vector of words, testing for binary presence)
– Question words (vector of words, testing for binary presence)
– Previously responded in the forum or not
– Number of previous responses
– response time windows
– words in the thread (high information gain 5950 words vector from the corpus)

In order to develop a message classifier that could categorize the messages into one of the above categories, we manually tagged 412 different conversation threads with different message categories. We used this labeled data from different WebMD forums to evaluate the classifiers using 80% of data for training and the rest for testing the models.

We used three standard algorithms to compare the accuracies of message classification system using rich feature extraction to aid in classification. In the first two approaches involving Bayes’ Classification and Support Vector Machines, this problem is a standard multi class text classification problem. Third approach using CRFs formulate the problem as a Sequence Labeling problem. Conditional Random Fields (CRFs)[6] are discriminative conditional probability distribution models that allow to take advantage of the sequential nature of conversations better. From the experiments, we see that CRF was able to pick up the right categories from the messages and was able to do better (Table 1) that the other standard methods.

Table 1. Message Classification

	Accuracy	Time(sec)
Bayes' Classifier	53.5	0.1
SVM Classification	56.1	1.8
Linear Chain CRF	67.9	9.8

2.3 Recommendations

Cobot provides three types of recommendations. It recommends and notifies relevant people who may be interested in joining conversations. It provides topic specific web recommendations and it also suggests past similar conversations from the system.

People Recommendation: While designing a recommender system, it is important to take into account the domain implications and fine-tune the algorithms accordingly. To provide social recommendations with a high degree of conversion rate, the system needs to identify people who can provide answers to asked questions, share similar health experiences and provide topic specific opinions and advice. Our system is built around health information domain therefore users are generally not concerned with building their social ties, instead, the goal is to serve the user's contextual information need. One important aspect in this domain is reputation of the recommended users, since there is no prior information and relationship of these users with the person who starts a conversation. We are building the reputation system by allowing users with the ability to rate conversations. The system takes into account factors for weighting the users differently (based on types (asker, responder, viewer), length of conversation, topics, etc.)

Our system currently tries to find a recently active user first who participated in similar conversations. Different conversational facets are matched with episodic memories and a spreading activation search on the semantic net is performed for recommending the best 3-4 users for the conversation. The activation is spread to the neighboring nodes proportional to the weight of each connecting association in the semantic net. There are several parameters in the system that can be learnt based on activity of users. Parameters for episodic memory window size, semantic memory learning and unlearning rates, concept co-occurences and feedback strengths for associations are initially set heuristically and can be fine-tuned to suit individual users.

Knowledge Recommendation: For web search and conversation recommendations, we reformulate queries from the conversation snippets based on occurrence of concepts and relationships and types of messages. For a given target query Q_t, past conversations are ranked so that the results which are most likely related to the learned preferences of the community are promoted[14][8][7]. This kind of personalization is based on the reuse of previous search episodes: the promotions for Q_t are those results that have been previously selected by community members for queries that are similar to Q_t.

Cases are represented as tuples made up of the query component (a set of query terms, Q_i used during some previous search session) along with web recommendations and past conversations with their community hit counts. Our formulation is based on similar work reported in Paper [14]. Each case is a summary of the communitys search experience relative to a given query.

Each new target problem (corresponding to a new query Q_t) is used to identify a set of similar cases in the case base by using a term-overlap similarity metric to select the n most similar search cases for Q_t .

These search cases contain a range of different result pages and their selection frequencies. Bearing in mind that some results may recur in multiple cases, the next step is to rank order these results according to their relevance for Q_t. Each result R_j can be scored by its relevance with respect to its corresponding search case, C_i by computing the proportion of times that R_j was selected for this cases query Q_i.

During the development of retrieval stage of the CBR system for Cobot, it was often observed that number of results retrieved were very large since the retrieval stage entailed a meta-search which queried many search engines which returned large number of results. We wanted to show only the top 2 to 3 results /conversations from the retrieved case base. Consequently sorting and ranking results according to their relevance to the ongoing conversation was necessary.

Relevance of a result with respect to the current target query Q_t) is calculated by computing the weighted sum of the individual case relevance scores, weighting each by the similarity between Q_t and each Q_i. In this way, results which come from retrieved cases (C_1, ..., C_n) whose query is very similar to the target query are given more weight than those who come from less similar queries. The relevance of a Result R_j to a target query Q_t and the case library comprising of cases from C_1 to C_n cases is expressed as:

$$WRel(R_j, Q_t, C_1...C_n) = \frac{\sum_i Relevance(R_j, C_i) * Similarity(Q_t, C_i)}{\sum_i Exists(R_j, C_i) * Similarity(Q_t, C_i)}$$

Similarity between the query and case is computed by finding the similarity between the query and case queries. We are using Jaccard Similarity as the similarity metric in our design. In this way, for given user, with query Q_t we produce a ranked list of results R_j that come from the communitys case base and that, as such, reflects the past selection patterns of this community. If the case base doesnt retrieve cases or the similarity confidence of the retrieved results is less than a user specified threshold t then, Q_t is used by the meta-search module to retrieve a set of web search results.

The top 3 results from the ranked results obtained either from the case base or the meta search engines are shown to the user. In this way, results that have been previously preferred by community members are either promoted or marked as relevant to provide community members with more immediate access to results that are likely to be relevant to their particular needs. This framework promotes community preferred results and conversations to the user.

Fig. 5. System Prototype

3 Discussion

From a brief usability study of the system prototype (Figure 5)[11], we learnt that socially powered search feature and the ability to collaboratively search together and discuss issues with real people instead of solitary search engine is very useful. Websites like Vark.com[4] are doing social search for generic question answering effectively using IM based messaging bots and other channels.

There are many technical challenges in community based information and recommendation systems. Cobot is being developed around the principle of *Suit the user, make it easy, make it good*. Cobot's approach and solution to next generation of socially enabled search is uniquely driven by new trends on the web, requiring new technologies for an integrated socio-semantic search experience. Instead of relying on search engines that inundate the user with a multitude of information, Cobot models the information finding task as an interactive and collaborative recommendation process within a social community. The user describes his need in natural language to a trusted community which is modeled via text conversations familiar to most users. Our agent based conversational framework for web search and recommendations uses a 'wisdom of crowds' approach to compensate for the limitations of traditional search engines and uses the experience of real users by proactively bringing them to participate in the conversations.

4 Acknowledgement

We acknowledge and thank our past project members, Alejandro Dominguez for writing the WebMD forum crawler, Bharat Ravisekar for working on a Personalized Feed Recommender based on Semantic Nets and Hrishikesh Pathak for implementing the case based reasoning module in Cobot. We also thank the contributions of Anushree Venkatesh and Stephanie Ahn as members of the Cobot project.

References

1. R. Cross, R. E. Rice, and A. Parker. Information seeking in social context: structural influences and receipt of information benefits. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 31(4):438–448, 2001.
2. B. M. Evans and E. H. Chi. Towards a model of understanding social search. In *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 485–494, New York, NY, USA, 2008. ACM.
3. E. A. Fox, D. Hix, L. T. Nowell, D. J. Brueni, D. Rao, W. C. Wake, and L. S. Heath. Users, user interfaces, and objects: Envision, a digital library. *J. Am. Soc. Inf. Sci.*, 44(8):480–491, 1993.
4. D. Horowitz and S. D. Kamvar. The anatomy of a large-scale social search engine. In *WWW*, 2010.
5. I. Karasavvidis. Distributed Cognition and Educational Practice. *Journal of Interactive Learning Research*, pages 11–29, 2002.
6. J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, pages 282–289, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.
7. K. McCarthy, L. McGinty, B. Smyth, and M. Salamó. The needs of the many: A case-based group recommender system. *Advances in Case-Based Reasoning*, 4106:196–210, 2006.
8. M. J. Pazzani and D. Billsus. Content-based recommendation systems. pages 325–341, 2007.
9. S. E. Robertson and S. Walker. Some simple effective approximations to the 2-poisson model for probabilistic weighted retrieval. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '94, pages 232–241, New York, NY, USA, 1994. Springer-Verlag New York, Inc.
10. B. Rogoff. *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press New York, 1990.
11. S. Sahay, S. Ahn, S.-C. Lu, B. Sherwell, A. Venkatesh, and A. Ram. Healthbuzz: Contextual social search and conversations. In *The Third Annual Workshop on Search in Social Media (SSM 2010)*, February 2010.
12. S. Sahay, S. Mukherjea, E. Agichtein, E. V. Garcia, S. B. Navathe, and A. Ram. Discovering semantic biomedical relations utilizing the web. *ACM Trans. Knowl. Discov. Data*, 2(1):1–15, 2008.
13. S. Sahay, A. Venkatesh, and A. Ram. Cobot: Real time multi user conversational search and recommendations. In *Recommender Systems and the Social Web*, volume 532. CEUR Workshop Proceedings, 2009.
14. B. Smyth, P. Briggs, M. Coyle, and M. P. O'Mahony. A case-based perspective on social web search. In *Proceedings of the 8th International Conference on Case-Based Reasoning: Case-Based Reasoning Research and Development*, pages 494–508, Berlin, Heidelberg, 2009. Springer-Verlag.
15. B. Wilson and H. Meij. Constructivist learning environments: Case studies in instructional design. *IEEE Transactions on Professional Communication*, pages 0361–1434, 1997.
16. O. Ybarra, E. Burnstein, P. Winkielman, M. C. Keller, M. Manis, E. Chan, and J. Rodriguez. Mental Exercising Through Simple Socializing: Social Interaction Promotes General Cognitive Functioning. *Pers Soc Psychol Bull*, 34(2):248–259, 2008.