IDIAP COMMUNICATION REPORT

FINDING WITHOUT SEARCHING

Andrei Popescu-Belis

Idiap-Com-01-2010

JANUARY 2010

Centre du Parc, Rue Marconi 19, P.O. Box 592, CH - 1920 Martigny
T +41 27 721 77 11 F +41 27 721 77 12 info@idiap.ch www.idiap.ch

Finding without searching

Receiving the information and services you need, exactly when you need them

Andrei Popescu-Belis
andrei.popescu-belis@idiap.ch
January 19, 2010

Contents

1. Description of the problem

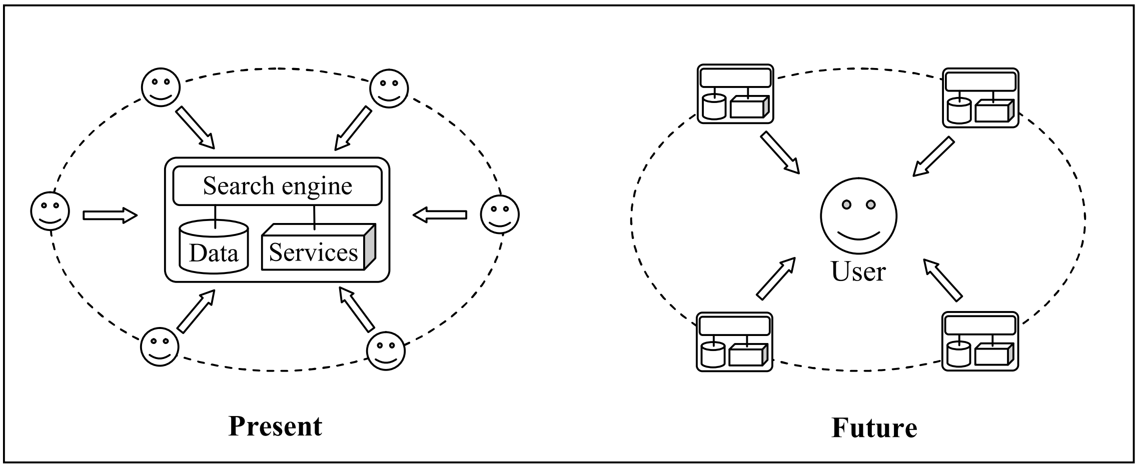
Advances in information technology have tremendously increased the amount of information and services that are available online. To classify and access this digital universe, portals and search engines constitute entry points for people looking for specific information or services. These tools are optimized to provide adequate results in response to a query submitted by a human user, and in certain cases they offer some help to formulate the query itself. This model of access to Web-based information and services is clearly driven by explicit demands from the users, who must invest some effort into searching, without always knowing if the results will be worth the effort. Therefore, in many situations, relevant information or services remain undiscovered or unused because the tradeoff between the cost and the benefits of a search do not encourage users to start this process.

Despite the emergence of the World Wide Web, the search-and-retrieval model of information access does not represent a radical change with respect to more traditional ways of disseminating knowledge or providing services. For instance, when compared to traditional paper-based libraries, Web-based access to knowledge is significantly quicker, and yields potentially more up-to-date results, but remains essentially within the same search and retrieval paradigm, similar to lookup in paper-based or computerized directories.

We believe that one of the major challenges in computer science is to reverse the traditional model of information access by removing the need for formulating explicit searches for information and services. Instead, context-aware software should bring to people exactly the information and services they need, at the right moment, with no effort required from users apart from deciding whether or not to consider and to use the results. Enabling users to find information without searching is one of the radical challenges that must be solved in order to benefit fully from the wealth of information and services available over the Web.

This challenge can be metaphorically compared to the development of tap water technology. Water had first to be drawn from wells and transported with significant efforts, but the introduction of domestic water systems, with pumps, pipes and faucets, enabled the delivery of water directly to the houses. Similarly, the Web now requires "information faucets", which can be open or closed at will, and which can learn to deliver exactly the information that the users need, whenever and wherever they need it. In fact, broadcast media such as radio or television already constitute primitive types of "information faucets", because they do not require from users any explicit searches for information other than opening or closing a communication channel. However, the information they provide is almost totally unrelated to the needs of a specific user at a given moment, place and activity.

The challenge proposed here calls for a Copernican revolution in the universe of information providers and information users, which would *put the users in a central position*, rather than the information sources or the search engines. Unlike the current conception of this universe, in which information and service providers occupy the focal points to which people must refer, we believe that people will benefit fully from the information universe only when they are at its center, surrounded by providers that tend to their needs, making information and services accessible in a natural, adaptive and effortless manner.



2. Usage, challenges and expected benefits

Information needs are pervasive in human activities, and the situations in which a personal information assistant would be helpful are innumerable. We envision here three scenarios of use, from the most challenging to the most feasible one: a mobile universal guide, a writing assistant for knowledge workers, and a meeting assistant for people involved in business discussions.

1. The *mobile universal guide* would be a small-size device that permanently displays context- and activity-aware information, personalized for its owner. Depending on his/her information needs, the owner could refer to the mobile universal guide – by looking at the

screen or possibly activating an audio output – and find information or services that would precisely answer his/her needs, without the effort that is at present required to start searching for such resources[1]. Displaying accurate and rich information that is exactly relevant to the ongoing context and activity, and that is accessible at a cost no higher than consulting the screen of a portable device, could considerably simplify and improve our way of life. Examples of applicable contexts/activities are: travelling to an unknown place, fact-checking in human conversations and interactions, shopping, sightseeing, etc., while examples of relevant information are: time and platform of next train; name of best-known Durrenmatt's play; location of a convenient hotel in Zermatt; product reviews and pricing information; and so on.

2. A *writing assistant* for spontaneous retrieval would also prove helpful when writing a document, such as an academic paper or a technical report. The assistant would monitor the user's output in a text processing system and the his/her other actions, and would provide suggestions about useful resources such as previous notes or drafts, similar work, word proofing and thesaurus information, and so on. This assistant (as the previous one) could adapt its suggestions over time, based on implicit relevance feedback from the user, i.e. based on feedback that is inferred by the assistant from the user's actions following its suggestions.

3. A *meeting assistant* would be helpful to participants in technical or business meetings, who often need to refer to specific information about the topic of their discussion, such as budget figures, previous decisions, or data about their competitors. Although this information can be found within the organization's information systems or on the Web, the discussion flow during a meeting makes it impossible for participants to start searching for relevant documents and locating precise information within them. A digital assistant that would be able to track the contents of an ongoing meeting discussion, and to retrieve documents and data based on these observations and on the meeting's context, would be of considerable use to the progression of a meeting. Such an assistant would monitor the discussion on a permanent basis, but participants could turn to it (open the "information faucet") only when their information needs and attention resources prompt them to do so. References to a preliminary version of such an assistant are given in Section 4.

3. Related research domains

Solving the challenge of "finding without searching" requires mainly the capacity to detect the information needs of a user based on fine-grained modeling of their activities and context, to match these needs with information and services that have been previously indexed with similar precision, and to present the results in a user-friendly manner. The paradigm shift required to solve the challenge relies on the convergence of several research directions in computer science, artificial intelligence, signal processing, as well as cognitive science and social science. The steps towards a solution outlined in Section 5 are related to these domains.

---

[1] The most widespread antecedent of such a device, albeit a very primitive one, is the watch, which offers its owner temporal information with minimal effort.

Our main proposal is to design software that infers a user's informational context from his/her language production, using an ontology-based approach to ensure precision, and to estimate the user's information needs from this representation. The information and services on offer should also be indexed using an ontology, replacing traditional keywords with concepts. Finally, an appropriate matching procedure between the information needs and the resources should take into account personalization, implicit relevance feedback, and the overall context of use, while keeping in mind privacy concerns and avoiding unwanted information.

The main cornerstone of the present challenge is to improve the computer's capabilities to manipulate ontologies – i.e. the formal representation of concepts and their relations. A well-known topic in artificial intelligence, knowledge management and natural language processing[2], ontologies received recent interest in relation to the conceptual annotation of large quantities of text, for instance in the ongoing DARPA OntoBank / OntoNotes project[3], and several large ontologies are now available[4]. In parallel, content-representation languages based on XML are being developed for data (RDF, OWL) as well as for Web services (WSDL) or for multimedia content (MOWL)[5]. However, ontology-based indexing is still an open problem, and the use of ontologies for matching the information context of a user inferred from speech with the information and services generally available is a novel approach in the language domain, related to the problem of automatic term extraction from texts[6].

From an applicative point of view, a number of systems for "query-free search" or "just-in-time retrieval" have been proposed, but none of them has taken advantage of the conceptual approach proposed here, and therefore their performance and influence have remained quite limited. For instance, some studies have explored the possibility of providing support information to users based on their current activities with a computer[7], while others have enriched TV broadcast

---

[2] Vossen P. (2003) "Ontologies", in Mitkov R., ed., *The Oxford Handbook of Computational Linguistics*, Oxford, UK: Oxford University Press, 2003, pp. 464-482.

[3] Hovy E.H., Marcus M., Palmer M., Pradhan S.S., Ramshaw L., and Weischedel R. (2006), "OntoNotes: The 90% Solution", *Proc. of HLT-NAACL 2006 (Human Language Technology / North American Association of Computational Linguistics)*, New York, NY; Pradhan S.S., Hovy E.H., Marcus M., Palmer M., Ramshaw L., and Weischedel R. (2007), "OntoNotes: A Unified Relational Semantic Representation", *Proc. of ICSC 2007 (IEEE International Conference on Semantic Computing)*, Irvine, CA.

[4] For instance Omega – Philpot A.G., Hovy E.H., and Pantel P., "The Omega Ontology" (2005), *Proc. of ONTOLEX Workshop at IJCNLP 2005 (International Joint Conference on Natural Language Processing)*, Jeju Island, Korea – or OpenCyc (http://www.opencyc.org) or the lexically-based Wordnet – Fellbaum C. (1998), *WordNet: An Electronic Lexical Database*, Cambridge, MA: The MIT Press.

[5] For instance, RSS, the Rich Site Summary or RDF Site Summary is a representation language giving access to news sites or other rapidly changing content, signaled by a sign. RSS is mostly used for the keyword-based representation of news feeds to which a user subscribes, although conceptual representations are potentially supported as well. The feeds are updated automatically by RSS-compliant readers, thus simulating a query-free behavior, but without elaborated filtering of the feed's contents.

[6] Jacquemin C. (2001), *Spotting and Discovering Terms through Natural Language Processing*, Cambridge, MA: The MIT Press.

[7] Hart P.E. and Graham J. (1997), "Query-free information retrieval", *IEEE Expert: Intelligent Systems and Their Applications*, 12(5):32–37; Rhodes B.J. and Maes P. (2000), "Just-in-time information retrieval agents", *IBM Systems Journal*, 39(3–4):685–704; Budzik J. and Hammond K.J. (2000), "User interactions with everyday applications as context for just-in-time information access", *Proc. of IUI 2000 (5th International Conference on Intelligent User Interfaces)*, New Orleans, LA.

news with pointers to more detailed articles related to the current topic[8]. On another plan, voice-directed search was also explored[9].

Modeling the user's context and activities should benefit from more inspiration from social sciences, to make context-modeling techniques more robust and more realistic[10]. One component of context-awareness, geo-location, has already an influence on the behavior of popular search engines, and plays a major role in location-based services for mobile devices (e.g. to look for the closest restaurant). Human activity modeling from large datasets is only emerging as a computational problem[11], as is social signal processing[12], two fields that should provide essential contributions to context and activity inference from physical sensors.

4. Steps towards a solution

Our research interests lie broadly within the domain of natural language engineering, multimodal processing and human-computer dialogue. In particular, the problem of reference resolution appeared to us as a crucial one for natural language understanding, with both theoretical and applicative potential. We have proposed methods for resolving references to entities in narrative texts[13], i.e. for building representations of these entities based on the referring expressions that mention them, and for resolving references to documents in human-human dialogues within meetings[14]. These studies are steps towards determining precisely the contents of a monologue or dialogue, offering a starting point for information searches based on entities rather than words.

Approaches to dialogue analysis in the meeting context[15] and to information access within multimedia meeting archives[16] were also proposed, with special emphasis on the precise

[8] Henziker M., Chang B.-W., Milch B., and Brin S. (2005), "Query-free news search", *World Wide Web: Internet and Web Information Systems*, 8:101–126.

[9] Franz A. and Milch B. (2002), "Searching the web by voice", *Proc. of Coling 2002 (19th International Conference on Computational Linguistics)*, vol. 2, p. 11–15, Taipei, Taiwan.

[10] Bolchini C., Curino C.A., Quintarelli E., Schreiber F.A., and Tanca L. (2007), "A data-oriented survey of context models", *ACM SIGMOD Records*, 36(4):19–26.

[11] Farrahi K. and Gatica-Perez D. (2008), "Daily Routine Classification from Mobile Phone Data", *Proc. of MLMI 2008 (5th Workshop on Machine Learning for Multimodal Interaction)*, Utrecht, p.174-185.

[12] Vinciarelli A., Pantic M., Bourlard H., and Pentland A. (2008), "Social Signal Processing: A Survey on Nonverbal Behaviour Analysis in Social Interactions", *Proc. of ACM International Conference on Multimedia*, "Brave New Topic" session, Vancouver, Canada.

[13] Popescu-Belis A. (2003), "Evaluation-Driven Design of a Robust Reference Resolution System", *Natural Language Engineering*, vol. 9, n. 3, p.281-306.

[14] Popescu-Belis A. and Lalanne D. (2004), "Reference Resolution over a Restricted Domain: References to Documents", *Proc. of ACL 2004 Workshop on Reference Resolution and its Applications*, Barcelona, Spain, p.71-78; Popescu-Belis A. and Lalanne D. (2006), "Detection and Resolution of References to Meeting Documents", in Renals S. and Bengio S., eds., *Machine Learning for Multimodal Interaction II*, LNCS 3869, Springer-Verlag, Berlin, p.64-75.

[15] Popescu-Belis A. (2008), "Dimensionality of Dialogue Act Tagsets: An Empirical Analysis of Large Corpora", *Language Resources and Evaluation*, vol. 42, n. 1, p.99-107.

[16] Popescu-Belis A., Baudrion P., Flynn M. and Wellner P. (2008), "Towards an Objective Test for Meeting Browsers: the BET4TQB Pilot Experiment", in Popescu-Belis A., Bourlard H. and Renals S., eds., *Machine Learning for Multimodal Interaction IV*, LNCS 4892, Springer-Verlag, Berlin, p.108-119.

definition and evaluation of the problems to be solved, and on the construction of training and test data, within the Swiss IM2 NCCR[17].

We are currently coordinating an ongoing effort to build an "automatic content linking" prototype within the AMIDA EU project[18], which implements the idea of effortless retrieval of documents based on the perceived conversational content in a meeting[19]. This demonstrator uses only word matching, therefore its precision and versatility are still limited to specific meetings and document databases, within a fixed project. However, the demonstrator provides an existing experimental platform, which awaits generalization of the information sources, addition of services, development of an ontology-based matching procedure, and increasing mobility and context-awareness.

We believe that the solution to the challenge put forward here requires a number of distinct, but related capacities: (1) detect a person's information needs from their activities, (2) represent these needs in a formal way, (3) represent available information and services in a similar format, (4) match the two representations, and (5) display the results. The solution to "finding without searching" relies therefore on progress over each of these steps, and on their successful integration into an application with a demonstrable utility.

Depending on the point of view, one can either emphasize the value of the scientific contributions to each step, which can be achieved for instance through a number of PhD theses, or the practical value of the application and its commercial potential, which can be exploited for instance by Swiss CTI projects. The long-term vision proposed here as the first application in Section 2 lies probably many years ahead, but the more tractable aspects of the challenge – the third application in Section 2 – constitute first steps towards a solution.

The most significant scientific and technological achievements that are needed are the following. In order to infer the needs of a user at a given moment in terms of information and services, it is necessary to acquire knowledge about: (1) the specific activity or task that the user is carrying out; (2) the overall context of this activity; (3) the long-term preferences of the user. For the first point, the analysis of the linguistic output of the user (spoken or written) should be key to an initial solution, as this data provides the most informative cues about their activity[20].

---

[17] Interactive Multimodal Information Management (http://www.im2.ch), a Swiss National Center of Competence in Research, started in 2002, and headed by the Idiap Research Institute and EPFL. The author was the head of one of its seven modules in the second phase (2006-2009), IM2.DMA, on Data Management and Access.

[18] Augmented Multiparty Interaction with Distance Access (http://www.amiproject.org), a European Integrated Project started in 2007 and headed by the Idiap Research Institute and the University of Edinburgh. The author was the head of AMIDA.WP6, the work package on Applications, Integration and Evaluation (2008-2009).

[19] Popescu-Belis A., Boertjes E., Kilgour J., Poller P., Castronovo S., Wilson T., Jaimes A. & Carletta J. (2008), "The AMIDA Automatic Content Linking Device: Just-in-Time Document Retrieval in Meetings", in *Proc. of MLMI 2008 (5th Workshop on Machine Learning for Multimodal Interaction)*, Utrecht, p. 272-283; Popescu-Belis A., Poller P., Kilgour J., Boertjes E., Carletta J., Castronovo S., Fapso M., Flynn M., Nanchen A., Wilson T., de Wit J., & Yazdani M. (2009) - A Multimedia Retrieval System Using Speech Input. *Proceedings of ICMI-MLMI 2009 (11th International Conference on Multimodal Interfaces and 6th Workshop on Machine Learning for Multimodal Interaction)*, Cambridge, MA; see also AMIDA Project Deliverable D6.7, April 2008.

[20] This means that the initial application is less suitable to situations where the users remain entirely silent. It should however be possible to talk directly to the application and prompt it to return results.

In some contexts, these cues can be supplemented with information from the user's agenda, location, or computer/PDA usage, which provide more elements for contextualization. Preferences can be learned using a relevance-feedback mechanism on the output of the entire process, when relevance is inferred from the observed use of the information on the physical device.

The main theoretical advance needed to solve the problem of query-free information access with sufficient precision is, in our view, a robust and precise representation formalism for the description of information needs. The main inspiration source should be the work on ontologies cited above. A mechanism should extract the main concepts from a speaker's linguistic output, starting from the speech recognition results and performing robust linguistic analysis and disambiguation at the syntactic, semantic and discourse levels. Similarly, available information and services should also be represented conceptually using elements from a large-scale ontology – this can be done upon creation of the resources, or by automatic *a posteriori* semantic indexing, or using information from social bookmarking. The most relevant information and services for a given user can be found by matching the user's information needs with the available sources, i.e. matching two conceptual representations expressed using the same formalism.

Finally, to provide results in an unobtrusive and easily accessible form, user-friendly interfaces must be designed and implemented on suitable devices. These could start from specific use cases, such as those identified in Section 2, starting from more tractable contexts such as business meetings or technical writing, and aiming later at the more ambitious mobile universal guide that enables "finding without searching" in any context and activity.