Knowledge Representation in the Agent-based Travel Support System

Marcin Paprzycki, Austin Gilbert, and Minor Gordon

Computer Science Department, Oklahoma State University, Tulsa, OK 74106, USA
{marcin,minorg,austirg}@a.cs.okstate.edu

Abstract. One of the interesting problems in implementing an e-commerce system is the question of what knowledge is required and how it should be implemented. In this note we attempt to address this problem in the guise of an agent-based e-commerce system for providing personalized travel services.

1 Introduction

As the amount of information contained within the digital landscape increases exponentially, we are forced to reconcile our business models with the new information paradigm. The crux of this reconciliation ultimately lies within the ever-increasing "intelligence" of computers. Human-centered business processes must be migrated into the new world of computer-driven commerce. The essence and core of this migration is knowledge, the common bridge between humans and computers. In the business world, the most important knowledge is that directly related to markets, products, strategies and above all, human behavior [4, 14]. Knowledge processing capabilities in humans are transparent; however, in order for computers to assume these capabilities, we must decompose the human processes into their component parts. This is the focus of the newly emerging field of knowledge management [8]. In this context knowledge consist of information, which has been aggregated from raw data sources, and refined to reveal trend patterns, from which knowledge of situation and motivation may be created.

Knowledge is even more crucial to e-commerce than it is to traditional business, because e-commerce lacks the intuitive/assumptive component of human-to-human transactions. Missing is the human touch that customers associate with successful brick-and-mortar establishments. This deficiency must be compensated through the use of knowledge management in the e-commerce environment [5]; knowledge management is the only way to make up for the helpful suggestions of a salesperson in the local store, often the most crucial aspect of the customer experience. With knowledge management, e-commerce systems can close the gap between the customer and the product in a way that is analogous, though not the same, as a human seller.

In order to successfully fulfill these roles, knowledge within the e-commerce system must reflect (1) the information environment from which the knowledge is derived, and (2) the functions of the system transforming that information into knowledge. In addition, it must be possible to represent knowledge in an electronic form. The system should be accurate, relevant, robust, and extensible enough to fully utilize this knowledge, and adapt to the continuously changing market of buyers and sellers.

1.1 Information Environment

The Internet is the natural environment for a modern e-commerce system, and consequently, nearly all available input into a system is acquired from Internet sources. The most important of these sources are the end-users: their choices are the basis for all directives and actions of the system. End-user feedback is especially crucial to service-oriented e-commerce systems, which do not sell physical products, but provide information services. Examples of such systems include popular comparison-shopping sites, like mySimon.com, DealTime.com, and PriceGabber.com, which aggregate product information from across the Internet and offer it to consumers as a value-added service. By necessity, the information domains searched by these systems and the content they deliver is determined by the nature of the end user's request and motivation for making it.

1.2 The Travel Support System

One such information domain is travel. Sites like travelocity.com, expedia.com, orbitz.com, et al. have popularized comparative shopping for inexpensive travel arrangements. While the knowledge management processes of such systems are typically proprietary, there are elements common across the industry as well as similarities with other e-commerce systems. These processes are the focus of our investigation, which has led to the development of a prototypical knowledge-centric e-commerce system within the travel domain. The goal of the system is to replicate the role of a human travel agent in planning itineraries, and it was initially proposed in [11, 12]. An agent-based based framework and a high-level overview of the system were presented in [1]. Some technical issues were discussed in [13].

The knowledge in the system exists to provide the most personally relevant information to the customer, just as a human travel agent would. There are two general classes of knowledge in the travel support system: travel-related and customer-related (corresponding to the content management and content delivery spheres of a general e-commerce system [3]). Though our focus is on the knowledge within the system, for the purposes of this note we will omit most questions related to knowledge management directly related to the customer, which have been discussed in [14] and concentrate our attention on the travel-related knowledge domain.

Due to the lack of space we will not be able to go into the details and we will proceed as follows. First, we will summarize the general functionality of the agent-based Internet Travel Agency, augmented by the specification of the types of travel-related knowledge that is stored in the system. We will then analyze a sample case of an imaginary traveler, attempting to golf during a business trip, to point out the travel-related knowledge required to support his needs. In the next section we will provide some technical details of how this knowledge is to be handled in the system.

2 Internet Travel Agency – Summary

On the functional level, the proposed Internet Travel Agency works as follows: travel-related information is primarily obtained from Verified Content Providers (*VCPs*: trusted sources on the Internet [12]). Metadata (only) obtained in this way is cross-referenced into multiple classification schemes: location, geospatial extent, travel-related ontology, popularity and other categorizations. The unstructured data available on the Internet (outside the *VCPs*) may be searched for relevant information as well [1]. Information garnered from the Verified Content Providers and the Internet is manipulated in a source-neutral format (see below); the original sources may be databases, web sites or other infomediaries [16], related to the travel support system.

Customers connect to the system using Internet-enabled devices – web browsers, WAP phones and others – and request travel-related information [13]. The initial response to a request is prepared from the *VCP*-derived information *indexed* in the system (via acquisition of data from the *VCPs*), and then filtered to match the user's personal preferences [12, 1]. The system also attempts to sell additional services to the user through targeted advertisements [4]. For responses that could benefit from more in-depth information than that already prepared, search agents can be sent to the Internet, and information obtained from them integrated into the response (as subsequently indexed in the system). Similarly, in the process of interacting with the system, the customer's queries may be refined as more specific content is presented, which will be combined with accompanying refinements in directed advertising. All of the details of customer-system interactions are then stored in the user behavior database for later data mining. The complete system framework is designed in terms of agents and therefore separate agents perform all of the above-described functionalities [7]. This includes the main repositories of travel knowledge: the travel expert and the advertising expert.

3 Aspects of Knowledge Within the System

In order to build the framework of a knowledge management system, we must divide the knowledge according to its function within the system – i.e. it must be divided by the questions it answers. For this we create a viable "story" of a transaction, and look into the aspects of knowledge necessary to its successful completion, e.g. the questions that *must* be answered. The sample case is:

Our customer, Robert, goes on a business trip to Palm Springs, Florida. The trip isn't all business, however, and Robert intends to play golf while he is there.

He has never been golfing in Florida before, though he has played on other trips. Robert would like to know the courses that are close to his destination, and of these, which have tee times available during his stay, and what are these times.

The questions necessary for the system to understand Robert's request are those typically one asks when trying to understand a given situation: "Who?", "What?", "When?", "Where?", and "Why?". Though these questions may seem simplistic, their answers provide a nice, significantly divided picture of the knowledge involved in the interaction between Robert and the travel support system - much as a journalist uses them to concisely relate a story: to encode and transfer knowledge.

3.1 Who?

While in this note we do not focus on customer-oriented knowledge, it must be stressed that our knowledge about Robert is the most important for satisfying his needs. The system has kept a detailed record of Robert's past trips: the options he considered from an array of choices, the ones he selected and those he rejected are all important in forming a picture of who Robert is – as a consumer of travel services [4]. This information has been transformed into Robert's profile, which has been initially constructed based on one of the four possible approaches described in [4] and later modified from knowledge and information mined from the records of his interactions with the system. For a full discussion of this process, as well as some of the pitfalls in managing user profiles, see [4, 14].

We cannot present Robert with all of the courses near his destination; there are far too many. Instead we must carefully trim the initially displayed response to his request, to include only those golf courses which pass through three personalization "filters": (1) possibility (real-world limitations on available choices), (2) acceptability (what Robert considers a viable choice) and (3) affability (what Robert likes). To perform this personalization selection we need to consider a number of other decisions. For example: on business trips, does Robert like to play the expensive private courses, the more crowded but less costly public ones, or will he settle for whatever is available? etc. In this way the knowledge about Robert as a traveler plays the role of a filter of the travel data.

3.2 What?

Robert has submitted a very specific request, and so it is easy to determine the "what": a golf course. In general, finding an answer to this question is one of the most important functions of the system, especially if the customer's request is as vague as "something relaxing" or "something cultural." The system should know that "cultural" includes the typical choices, like concerts, museums and local historical sites (which are the travel categories represented in the system). In addition other options labeled "cultural" in accordance with the user's previous behavior and associations need to be considered; and here the who-question influences the selection process (some people consider square dancing "cultural," and the system has to be able to deal with this fact *about them*; the who-question).

We can therefore say that the what-question indicates the process through which the system will select, out of all the high-level categories known to it, those that are of potential interest to the user. This also points out that relationships between various travel categories need to be taken into account to be able to make additional suggestions to the user (e.g. someone looking for a kayaking vacation needs to be able to reach the destination etc.).

3.3 When?

The when-question points to two different determinations that have to be made. First, the time of Robert's stay in Florida, which filters availability of some courses. Second, the tee time availability that has to be cross-referenced with the list of courses that Robert rates highest (the who-question). These and other temporal factors (which may or may not be predictable) must be considered when first presenting Robert with available courses, including the order in which the choices are listed. However, because the information on golf courses in a given area is only *indexed* by the travel support system, we cannot know before hand the exact tee times open for the duration of Robert's trip. We deal with this problem in Section 5.

3.4 Where?

Certainly we know "where" Robert is going, but the knowledge involved in this aspect of the request is not so simplistic. We must consider the relative proximity of all available courses to Robert's base/location (we assume that the lodging has been chosen beforehand). This proximity is also relative because it depends on Robert's means of transportation, e.g. if he does not opt to rent a car in Florida some of the courses may not be reachable (and this, again, points to the interrelations between travel categories represented in the system). This points to an important relationship between geospatial knowledge and travel knowledge that needs to be taken into account. Note also that the question if Robert may be willing to take a taxi or public transportation (if it is available) to a certain course is again a who-question.

3.5 Why?

On the surface, this question may seem somewhat irrelevant, but this is not the case. During one trip, Robert may want to play golf for business reasons, i.e. with current or potential partners, employers or employees, etc. During another trip, even though it is for the purposes of business, Robert may want to play by himself, i.e. for recreation only. The type of courses he would choose in different situations may be vary as greatly as his motives for the trip, so the distinction is entirely relevant, though it must be determined subtly, if we are to keep the customer from feeling like his privacy has been invaded [4]. In order to provide *truly* personalized service, the travel support system *must* know "why."

4 Representing Knowledge Within the System

Every system utilizing large quantities of information must find a suitable way of managing and organizing this information in a useable format. We decided to use XML for information exchange, input and output because the tools of parsing, production, transfer and storage of XML are well known, widely implemented and constantly improved [15]. However, we must still decide how this content will be represented in XML within the system. This internal representation is the cornerstone of knowledge processing. Its format must lend itself to search, extension and maintenance, and, most of all, interoperability.

4.1 OTA

The Open Travel Alliance provides an ebXML-based, hierarchical vocabulary for describing knowledge and processes of the travel industry, including booking and availability [2, 10]. The internal knowledge structure of our travel support system will closely resemble the schemas defined by the Open Travel Alliance. It will, however, have to be extended as the OTA standard, for the time being is limited to a number of relatively standard travel services, i.e. air, car, railroad travel, travel insurance or golfing. Among others it does not include support for historical information, information about local currencies, weather information etc. While, over time, these categories may be included in the OTA specifications, at this moment we can include them in our system, as the schema is quite flexible in this regard.

4.2 Expert Rules

The travel support system requires expert systems to enforce temporal and s-patial reasoning, real-world limitations and other "reality" imperatives on the process of filtering available travel options. Clearly, Robert cannot be in two places at once, and the system must have an understanding of this fundamental fact. Beyond the reality-enforcing rules, there is a need for the system to recognize patterns of user behavior and encode them in expert systems. This reasoning enables the system to make on-target recommendations, based upon the user's profile as well as patterns recognized in the general population [4]. Some of these restrictions exist already in the OTA specifications [10], e.g. the fact that a person cannot be in two places at the same time. The remaining rules will become a part of the travel expert and advertising expert.

4.3 Cross references

One important aspect of the travel system is the necessity to cross reference information from the XML-based database against information contained in a geospatial index. The system will then be able to filter its query responses based upon geospatial significance, which leaves the focus on the relevant choices only (see above). To store and represent the geospatial data we will follow the approach introduced by the Open GIS Consortium [9]. Finally, there may be other indexes in the system as well, including those that rank absolute popularity (as determined by the number of followed-through selections) and other perspectives on the basic travel ontology defined by the Open Travel Alliance. They will be cross-referenced using standard database techniques [12].

5 Messaging Technologies Within the System

The selection of the OTA ebXML-based standard allows us an accurate way of representing and responding to customers' requests. This is because the ebXML protocol provides a standardized interface to sources of business information, making it possible to send formal request and accept formal responses, as well as error messages, warnings, status updates, etc., while the OTA standard provides with the semantics of these messages. The travel support system will therefore utilize the ebXML standards as one means of communicating with various indexed information resources within the system; that is to say, if an ebXML resource is available, our system will be able to utilize it. This messaging standard will also be considered the preferred method of exchanging all travel related information. Using ebXML not only lifts the burden of design from our development team, but also ensures that large components of the system that will be compatible with international standards and other ebXML-compliant travel support systems.

We will use Robert's wish to enjoy a round of golf on his trip as demonstrative example of ebXML message interchange. At this point, Robert's desired destination has been locked down, so it is time to begin the search. First a query is executed on the internal system index [1], an ebXML resource(s) for golf courses in the Palm Springs area is(are) returned. The steps in the exchange are very straightforward:

- request a list of golf courses in the area,
- receive the results,
- select a course and a tee time and see if it is available,
- receive confirmation that the time is available or receive an error message stating that the time is taken,
- if the tee time is available, make a reservation for that tee time,
- if the reservation is accepted, a response is return stating so.

As an example, the request for nearby courses will have a form similar to that depicted in Figure 1.

insert figure 1 around here

Figure 1. Golf course service request as an ebXML message.

This request is parsed and translated into a database query for the pertinent indexed information sources. The "<Criteria Name>" tags will be used to search database columns of similar names for the values indicated. The results of this query are translated back into the OTA response syntax and eventually transported back for the user to view [13]. The OTA Course Response message will have a form similar to that presented in Figure 2.

insert figure 2 around here

Figure 2. Response to the golf course service request as an ebXML message.

In this case, two courses meet the criteria Robert was seeking: the location (Palm Springs) and the slope rating less than 113. The Open Travel Alliance does not define the specific business rules for determining responses; this is the defining function of the implementing system. The OTA XML schemas only describe travel resources and the actions (search, request, reserve, et al.), which may be performed upon them. Nevertheless, they allow us to accurately represent knowledge of the travel(ed) world in an eminently useful form.

6 Concluding Remarks

In this note we have sketched the main issue related to the travel knowledge representation in an agent-based Internet Travel Agency. It was suggested that after extending its scope, the ebXML-based Open Travel Alliance description of the world of travel will become the main travel knowledge repository. We are currently in the process of implementing a demonstrator system and will be reporting in more details about the encountered issues and applied solutions shortly.

References

1. Angryk, R., Galant, V., Paprzycki, M., Gordon, M.: Travel Support System - an Agent-Based Framework, in: H. R. Arabnia, Y. Mun (ed.), Proceedings of the International Conference on Internet Computing (IC'02), CSREA Press, Las Vegas, NV, 2002, 719-725
2. ebXML: http://www.ebxml.org
   http://www.fipa.org (2001)
3. Galant, V., Jakubczyc, J., Paprzycki, M.: Infrastructure for E-Commerce, in: M. Nycz, M. L. Owoc (eds.), Proceedings of the 10th Conference Extracting Knowledge from Databases, Wroclaw University of Economics Press, 2002, 32-47
4. Galant V., Paprzycki M: Information Personalization in an Internet Based Travel Support System, in: Abramowicz (ed.), Proceedings of the BIS'2002 Conference, Poznan University of Economics Press, Poznan, Poland, 2002, 191-202
5. Jakubczyc, J., Galant, V., Paprzycki, M., Gordon, M.: Knowledge Management in an E-commerce System, submitted for publication
6. Java Agent DEvelopment Framework (JADE): Telecom Lab Italia, http://jade.cselt.it and http://jade.cselt.it/papers.htm (2001)

7. Jennings, N. R.: An Agent-based Approach for Building Complex Software Systems, CACM, 44 (4), 2001, 35-41
8. Martins K., Heising P., Vorbeck J.: Knowledge Management, Springer-Verlag, Berlin, 2001
9. OGC: http://www.ogc.org
10. OTA: http://www.opentravel.org
11. Paprzycki M., Angryk R., Kolodziej K., Fiedorowicz I., Cobb M., Ali D., Rahimi S.: Development of a Travel Support System Based on Intelligent Agent Technology, in: S. Niwinski (ed.), Proceedings of the PIONIER 2001 Conference, Technical University of Poznan Press, Poznan, Poland, 2001, 243-255
12. Paprzycki M., Kalczynski P. J., Fiedorowicz I., Abramowicz W., Cobb M.: Personalized Traveler Information System, in: B. F. Kubiak and A. Korowicki(eds.), Proceedings of the 5th International Conference Human-Computer Interaction, Akwila Press, Gdansk, Poland, 2001, 445-456
13. Paprzycki M., Gordon M., GalantV.: Agent-Client Interaction in a Web-based E-commerce System, in: D. Grigoras (ed.), Proceedings of the International Symposium on Parallel and Distributed Computing, University of Iasi Press, Iasi, Romania, 2002, 1-10
14. Paprzycki M., Gordon M., Galant V.: Knowledge Management in an Agent-based E-Commerce System, submitted for publication
15. XML: http://www.w3c.org
16. Valera F., et al.: Communication Management Experiences in E-Commerce, CACM, 44(4), April 2001