STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **LII**, Number 2, 2007

KNOWLEDGESENSE: ENCYCLOPEDIC SYSTEM BASED ON SEMANTIC SEARCH THROUGH NLP

PAUL-VALENTIN BORZA, DANIEL GHIŢĂ, MIHAI NADĂŞ, OVIDIU SABOU, AND SIMONA MOTOGNA

Abstract. The article presents the solution that won the Imagine Cup 2007 National Finals and represented Romania at the Imagine Cup 2007 Software Design Invitational Worldwide Finals in Seoul, Korea August 5-11, 2007. In 2007, the Imagine Cup Theme was "Imagine a world where technology enables a better education for all". We created a unique solution that enables styles of natural interaction with a smart omnipresent humanlike entity, as it exploits human thinking within the education domain on an instant question answering based type of learning - self-education: KnowledgeSense. The solution takes advantage of intelligent devices and technologies that increasingly pervade modern environments - humans interact and learn intuitively while having a rich and complete experience with a virtual human person - Anna - that recognizes speech and gestures. Plus, it is also reachable through widely used communication channels like phones - SMS messages, VoIP or PSTN calls - and instant messaging networks. This approach emphasizes natural behavior by providing the opportunity to adapt to human needs and not reverse.

1. Introduction

Information is usually defined as "organized data", "data endowed with relevance and purpose" or "interpreted data" etc.; these definitions point to the fact that information includes human participation in the purposeful organization of raw data. Knowledge can only reside in one's mind and is the result of human experience and reflection based on a set of beliefs that are at the same time individual and collective; for instance, Nonaka and Takeuchi

Received by the editors: September 15, 2007.

2000 *Mathematics Subject Classification.* 68N01, 68T50.

1998 *CR Categories and Descriptors.* H.3.1 [**Information storage and retrieval**]: Content Analysis and Indexing – *Dictionaries, Indexing methods, linguistic processing*; H.3.3 [**Information storage and retrieval**]: Information Search and Retrieval – *retrievel model, search process*; H.4.m [**Information System Applications**]: Miscellaneous – .

define knowledge as "true and justified belief". The key difference between knowledge compared to information can be summarized by the role played by human beings. In the case of knowledge, as simple as it may seem, individuals play a prominent role as creators, carriers, conveyors and users; in contrast, in the case of information, these same functions can happen outside humans and without their direct influence.

KnowledgeSense uses the human knowledge that is continuously published on safe information sources through the Internet - Wikipedia, Encarta, Britannica and The World FactBook etc. - ensuring the trustworthy property of the vast content that the system is prepared to process. Finding answers through keywords - an approach that is often used in search engines like Windows Live and Google etc. - is extremely unintuitive; sometimes, choosing the right keywords becomes a time consuming challenge even to the most experienced user. People have always asked questions to quickly solve their misunderstandings, so the most natural way of solving the "international system units smallest prefix" puzzle, is to ask "What's the smallest metric prefix unit¿' - The answer is "yocto".

2. Architecture

Take a look at the architecture level of abstraction of the Enterprise Solution Patterns [5] - KnowledgeSense Pattern Frame from figure 1. From the application viewpoint, as Figure 2 shows, KnowledgeSense is an Object-Oriented Application that is logically structured as a Three-Layered Services Application. From the database viewpoint, the application is based on the OLTP processing model. From the infrastructure viewpoint, the hardware and network architecture are based on Four-Tiered Distribution (see figure 3), which calls for separate physical tiers for Web server and application server functionality. And finally, from the deployment viewpoint, we have created a Deployment Plan to map components to servers, based on a Smart Client Application.

Running on Windows Vista, the desktop application was designed and built upon Windows Presentation Foundation to offer a modern GUI and Experience. With the clear separation between the logic and the look of the program, WPF enabled quick development of a good looking interface that remains intuitive and serves its purpose - show information in a fast and convenient way - while maintaining the Windows Vista UX guidelines.

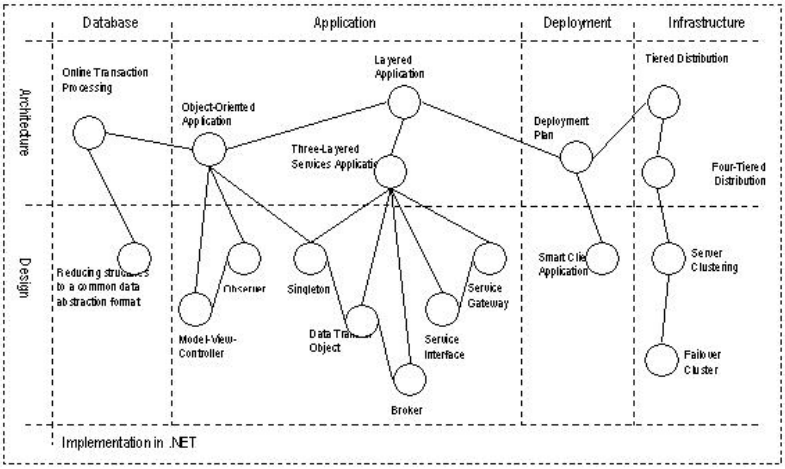

FIGURE 1. Enterprise Solution Pattern Frame

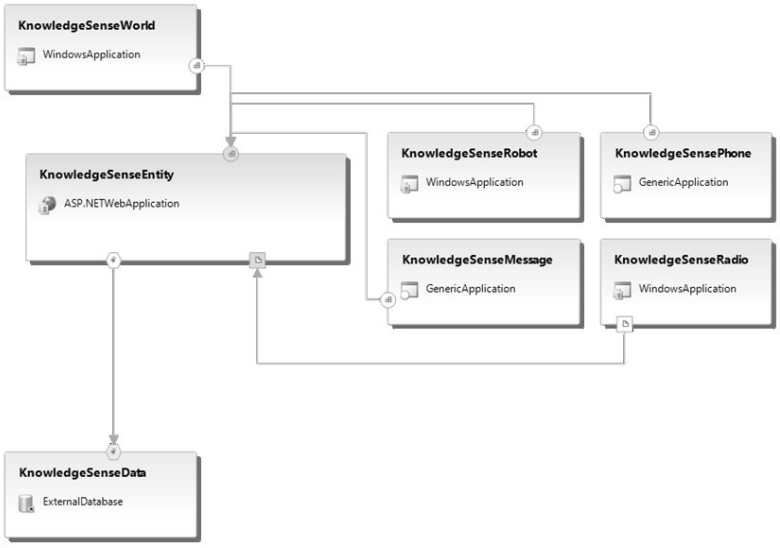

FIGURE 2. Enterprise Solution Pattern Frame

In addition to the classical way of presenting information (text, images, video etc.), a new dimension was introduced, bringing interactivity innovation

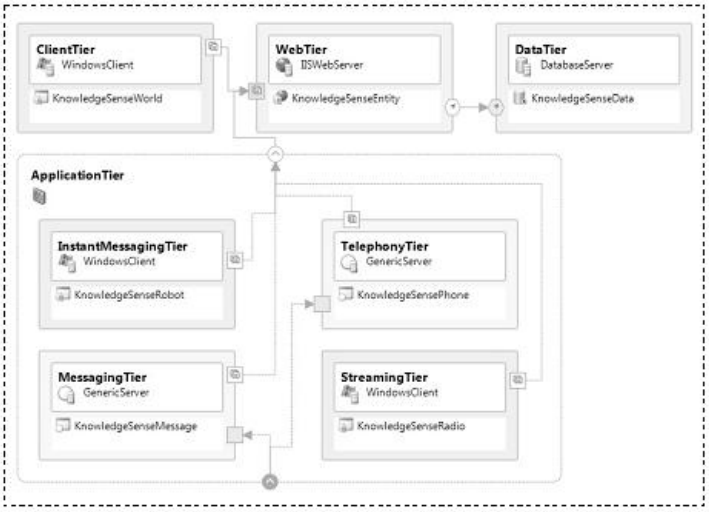

FIGURE 3. Four-Tiered Distribution

in electronic encyclopedias. A virtual assistant and a virtual world were integrated using OpenSceneGraph[2] - a 3D toolkit that enables advanced 3D graphics techniques, already used in many well known 3D graphics solutions like flight simulators and CAD products - achieving a completely new experience.

Gesture recognition data is gathered using a Nintendo Wii Remote Controller[3] (a $40 value) that provides six degrees of freedom through the use of an analog accelerometer device and an optical sensor, and then processed through the Hidden Markov Models[4] with the Forward-backward, Viterbi and Baum-Welch algorithms that recognizes gesture patterns and executes the associated actions. In addition to the sensors, the solution uses the other unique functionalities of the remote: rumble, speaker, direction pad etc.

The VoIP component establishes connections over the TCP implementation of the Session Initiation Protocol. Running on a standard protocol, a multitude of SIP-enabled softphones and hardphones can dial the Windows Workflow Foundation based application that runs inside Speech Server 2007 Beta. Unlike voice services that use the PSTN - with low audio quality - the voice recognizer and synthesizer are able to perform their tasks much better,

as they receive and produce high quality sound. PSTN phone calls are handled by the Cantata Technology TR1000 for MSS Speech - telephony and voice processing board.

The DotMSN Messaging Library[6] is used as a wrapper over the Windows Live Messenger network protocol, being able to trigger all instant messaging events required to chat with a person; the user might not even know it is talking with a machine.

The full product grid that provides a non-detailed, yet comprehensive snapshot of used technologies is presented in Appendix A.

3. Hibride Search Algorithm

KnowledgeSense is integrating a hybrid search mechanism , as in figure 4 that enables high accuracy in the process of getting the right answer for a certain question. We are combining traditional search technologies like the Microsoft SQL Server's Full Text Search feature and the Windows Live Search, with our custom build artificial intelligence driven natural language processing engine.

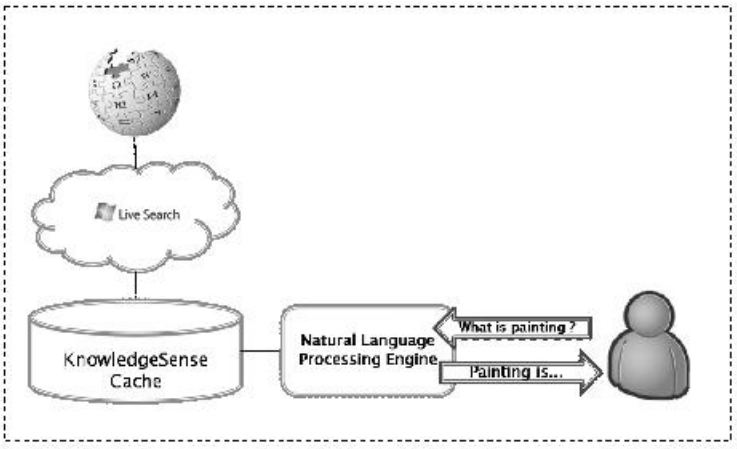

Figure 4. QA Flow

Our system uses a database in which the relevant encyclopedic information is stored. But as the system relies on the knowledge gathered on the Internet to provide users with answers, we need an intelligent process responsible for acquiring data from all the trusted sources assigned to our application.

From a technological point of view, our algorithm is capable of solving three different question answering situations. All of these situations start with the same procedure of breaking up the question into relevant keywords. For example if a user asks What is painting?" the engine will process the question and it will generate the keywords is painting". From this point on, the algorithm works differently for the mentioned situations.

The first case in when the database does not contain a relevant article for the keywords generated by the engine. We know if a set of keywords match an article by using the database's Full Text Search feature that returns for each article stored a match accuracy, which tells us if the article is relevant or not. And because we said above that in this case the database will not contain relevant articles for the given question, the algorithm continues the search on the next level, searching the Internet. The algorithm uses Windows Live Search to search the trusted data sources present on the Internet (e.g. Wikipedia) in order to retrieve articles that match the search criteria. Once this process is over, the engine parses the text of the articles thus creating a semantic tree of the information present in the article and after this is done all the data is stored into the database providing information for further queries. When the parsing process is over, the engine uses the semantic trees from each relevant article and retrieves the answer for the given question (e.g. "Painting is the process of applying color to a surface").

The second case is when the database already contains the semantic tree for the article relevant to the given question. In this case the semantic tree associated with the article is sent back to the engine, where it extracts the relevant answers for the given question. And to assure that our information is updated permanently, we run a comparing test between the information of the source and the article that is present in our database. If there are differences, the article is reparsed in order to deliver the most relevant answers. The third case is possibly the easiest of all. After the system delivers an answer for a question, the answer is stored into the database for a better performance the next time that exact question is asked. Thus, if the asked question is already answered, the system will return the answer without analyzing the semantic tree or searching for new articles on the Internet.

The engine that processes natural language identifies each question and parses the input to construct a tree with word relations[8] ; the WH-part - who, what, how etc. - and statement segment are extracted to classify the input accordingly. Keywords are generated and matched on articles where the

system looks for possible answers; the highest rated answer is rephrased with the additional available information and provided back to the user in human language[9].

4. Anna, the 3D animated assistant

The humanoid assistant from our application had to be as realistic as possible. In order to achieve this, we had to use the best real time rendering technologies available to us at that time. Since we already had some experience with OpenGL and OpenSceneGraph, we chose these technologies for rendering Anna and the objects in the virtual environment.

The first major problem we encountered when developing KnowledgeSense World (the desktop client application) was the integration between WPF (Windows Presentation Foundation) and a 3D graphics API. Although WPF provides some support for rendering 3D objects, it wasn't enough for us. We had to render tens of thousands of polygons at interactive frame rates with the assistant's 3D model included. WPF isn't suitable for this because it doesn't provide a low level way of specifying geometric data and it also doesn't provide hardware skinning support (for animating the character directly on the video card).

In order to render data with OpenGL, a special type of window has to be created. Since access to the operating system and hardware resources is better handled with C++, and OpenSceneGraph is a C++ toolkit, we needed something that could glue together the C# part (user interface and the logic of the application) and the C++ part (for rendering the scene). C++/CLI is a language designed for creating systems that interact with .NET and native components at the same time. The rendering part, which was implemented in native C++, was wrapped by a thin C++/CLI module that provided a simple .NET API for controlling the objects within the virtual world (with features like changing the position and rotation of objects, animating the character, changing the camera properties, etc.).

The realism of the scene was improved by adding static shadows (baked with 3D Studio Max) at high resolutions. It wasn't a problem because the environment was supposed to be changed rarely so the static approach for lighting the virtual world was enough (the resulting images had very high quality).

Throughout the development process, three models were used as the virtual body of Anna. The first two models were custom models made by

us, while the third model was professionally built. For character animation, we used the most popular open source library, named Cal3D, which is available at http://cal3d.sourceforge.net. We have acquired a license to use the third model (named Masha) from Turbosquid.com and integrated it into our scene after it was rigged (rigging is the process of attaching a skeleton to a mesh) by a professional Blender artist. Because the Cal3d exporter for Blender was the only one that worked for us (the 3Dsmax version was out of date), Blender was the main tool for manipulating the 3d content related to the virtual assistant. The final rendering of the character was done using osgCal, an OpenSceneGraph plug-in that renders Cal3d models. The development of the models and theirs transformations can be seen at http://www.borza.ro/demo/knowledgesense.

Table 1: Technologies

Technology Name and KnowledgeSense Component	Core	Message	Phone	IE Add-In	Robot	World	Team W
Microsoft Windows							
• Windows Vista Business Edition				X	X	X	
• Windows Internet Explorer 7.0				X			
• Windows Live Messenger 8.1					X		
• Windows Mobile 5.1 Pocket PC Phone Edition		X	X		X		
Microsoft Servers							
• Windows Server 2003 R2 Standard Edition	X		X		X		X
• Internet Information Services 6	X		X				X
• Windows SharePoint Services 2							X
• SQL Server 2005 Standard Edition	X						X
• Database Engine	X						
• Speech Server 2004 R2			X				
• Office Communications Server 2007 Speech Server			X				
• Visual Studio 2005 Team Foundation Server							X
Microsoft Developer Tools							
• Visual Studio 2005 Team Suite	X	X	X	X	X	X	X
• Visual C#	X	X	X		X	X	
• Visual C++						X	
• Expression Blend						X	
Additional Developer Tools							
• Autodesk 3ds Max 9						X	
• Blender 2.43						X	
• Counterpath X-Lite 3.0			X				
Microsoft Frameworks, APIs and SDKs							
• .NET Framework 2.0	X		X		X	X	
• ADO.NET	X						
• ASP.NET	X						
• .NET Compact Framework 2.0		X					
• .NET Framework 3.0			X			X	

• Windows Presentation Foundation						X	
• Speech API 5.3						X	
• Windows Workflow Foundation			X				
• Windows Mobile 5.0 for Pocket PC SDK		X					
• Windows Live Search API 1.1	X						
Additional APIs and SDKs							
• Xih Solutions DotMSN API 2.0.2					X		
• Princeton WordNet API 3.0	X						
• Proxem Antelope API 0.7.1	X						
• Stanford Parser API 1.5.1	X						
• Wii Remote Controller API 1.1						X	
• OpenSceneGraph SDK 1.2						X	
Devices							
• Cantata Technology TR1000 for MSS			X				
• HP iPAQ hw6915		X					
• MSI Star Key 2.0 Long Range Bluetooth Dongle						X	
• Nintendo Wii Remote Controller						X	

5. Conclusion

The solution delivers trustworthy educational knowledge in an intuitive manner that enables questions to be stated in natural language; subjects are presented in an interactive graphical environment. A complete set of platforms are available for application use.

References

[1] Dion Hinchcliffe: Patterns for High-Integrity Data Consumption and Composition - The Architecture Journal, Data by Design, http://msdn2.microsoft.com/en-us/arcjournal/bb245676.aspx

[2] Robert Osfield and Don Burns: OpenSceneGraph, http://www.openscenegraph.org/

[3] Brian Peek: Managed Library for Nintendo's Wiimote - MSDN Coding4Fun, http://blogs.msdn.com/coding4fun/archive/2007/03/14/1879033.aspx

[4] Lawrence R. Rabiner: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition - Proceeding of the IEEE, February 1989 http://www.caip.rutgers.edu/∼lrr/Reprints/tutorial on hmm and applications.pdf

[5] David Trowbridge, Dave Mancini, Dave Quick, Gregor Hohpe, James Newkirk and David Lavigne: Enterprise Solution Patterns Using Microsoft .NET - MSDN Patterns and Practices, http://msdn2.microsoft.com/en-us/library/ms998469.aspx

[6] *** Xih Solutions: DotMSN, http://www.xihsolutions.net/dotmsn/

[7] *** National Institute of Standards and Technology: Question Answering Collections, http://trec.nist.gov/data/qa.html

[8] *** Princeton University Cognitive Science Laboratory: WordNet, http://wordnet.princeton.edu/

[9] *** Proxem: Advanced Natural Language Object-oriented Processing Environment, http://www.proxem.com/Antelope/tabid/55/Default.aspx
[10] *** Stanford Natural Language Processing Group: Stanford Parser, http://nlp.stanford.edu/software/lex-parser.shtml

Department of Computer Science, Faculty of Mathematics and Computer Science, Babeş-Bolyai University, 1 M. Kogălniceanu St., 400084 Cluj-Napoca, Romania

E-mail address: bp20416,gd20439,nm20474,so20492@scs.ubbcluj.ro, motogna@cs.ubbcluj.ro