Fine-tuning blackboard architecture for question answering systems

Rivindu Perera

Department of Computer Science
Informatics Institute of Technology
Colombo, Sri Lanka
pererarivindu@hotmail.com

Abstract. Question answering is an emerging field which combines Information Retrieval (IR) and Natural Language Processing (NLP) by providing answers for questions formed in natural language. But when observing the trends in the question answering domain for past few years, it is noted that there is a lack of standardized architecture or design pattern which can be used in system implementation. Furthermore, classic architectural models like three-tier or service oriented architectures cannot provide advanced features required by question answering systems. Therefore, this paper is focusing on introducing customized blackboard architecture for question answering systems. Author tries to modify classic blackboard model architecture features accordingly and preserve their advanced qualities.

Keywords: Question answering systems, blackboard architecture, natural language processing, information retrieval

I. Introduction

Question answering is a field that require vast amount of knowledge to be manipulated and deeply analyzed to extract the exact answer for a given question. Therefore, underlying architecture of a question answering system must support knowledge management thus providing range of functionalities which can be performed on the knowledge. Currently, most of the question answering systems are developed using the layered architecture and therefore do not concentrate on the five most important knowledge management techniques:

- Knowledge acquiring
- Knowledge representation and modeling
- Knowledge validation
- Inference
- Justification

Therefore, research suggests the blackboard architecture which concentrates on the knowledge manipulation to be used for question answering systems. But classic blackboard model is still has no ability to innovate the question answering systems with novel features and require several modification to fine tune the pattern to be used with question answering systems. Blackboard architecture is consisted of four major layers:

- Control Shell
- Blackboard
- Knowledge source
- Agent or frontend system

It is required to customize each of these layers to focus on the question answering process by providing efficient way of answer extraction.

However, so far few research studies have investigated architecture standardization for question answering systems. But these research studies only concentrate on current industry specific architecture models and customizing them to be used in question answering system. Therefore, main attributes which are required by question answering system from the underlying architecture are not considered, rather these research studies try to incorporate the identified architecture to the problem.

Therefore, this paper seeks to explore the blackboard architecture customization for question answering systems to analyze the effect that it can perform on the answer extraction process and knowledge maintenance.
II. Background of the study

a. Qualities considered in question answering

Question answer is a complex process which involves several algorithms and processes which are associated with IR and NLP. Additionally, end users expect various qualities from a question answering system [3]. Therefore, qualities of a question answering system can be discussed under two categories,

- Qualities required by processes
- Qualities required by end users

Algorithmic process running inside a question answering system usually communicate with each other. Due to this, architecture by its own must support this communication [2]. When ontology or graphs are associated with a question answering system Extensible Markup Language (XML) or Resource Description Framework (RDF/N3) play major roles in inter-process communication [6].

Common problem that computer scientists face when developing a question answering system is that lack of software libraries for the selected platform [4]. For an example, Java has rich set of software libraries ready to be used with NLP and IR based application. Therefore, OpenNLP and Unstructured Information Management Architecture (UIMA) powered application stack are ready to be integrated to Java based question answering systems [1]. But when considering other frameworks like .Net and PHP, there is a lack of software libraries to be used in question answering system development domain. Possible alternative method that most computer scientists use is the software library conversion, but this method can arise several unknown issues at the runtime. But a better alternative is to use communication bridges to communicate with external software libraries. To perform this underlying architecture must support this feature.

b. Architectures for question answering systems

Most important factor noticed during this research is that there is no specific architecture for question answering system development. So far, no research study focuses this context to identify the architecture. But in this research, author identified 5 different architectural patterns that can be applied in a question answering system. Not stopping there this research performed an evaluation of these architectures through development of a question answering system. Result of this research can be found in Table 1.

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Method of evaluation</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service oriented architecture</td>
<td>Analysis and prototyping</td>
<td>For standalone application development, request and response behavior causes a negative effect.</td>
</tr>
<tr>
<td>Blackboard architecture</td>
<td>Analysis and prototyping</td>
<td>No extended support for knowledge management.</td>
</tr>
<tr>
<td>Pipes and filters architecture</td>
<td>Analysis and prototyping</td>
<td>Workers should be able to data flows according to the interface request. Predefined connection flows are not possible.</td>
</tr>
<tr>
<td>Execution Orchestrator architecture</td>
<td>Analysis and prototyping</td>
<td>Communication between workers is difficult when dynamic question processing is happening.</td>
</tr>
<tr>
<td>Customized layered architecture</td>
<td>Analysis and prototyping</td>
<td>Knowledge base systems interfaces must be designed separately.</td>
</tr>
</tbody>
</table>
Key point identified in this evaluation is that out of 5 architectures, blackboard has several important features which can be used in a question answering system. Therefore, author tried to customize the blackboard architecture to be used in question answering systems.

### III. Customized blackboard architecture

Though blackboard architecture is noticed as an appropriate architecture for question answering system, still there is a drawback that can be identified in the architectural pattern. It is the need of a knowledge management procedure to be applied to the architecture. Therefore, this research focuses on integrating knowledge management functionalities to blackboard architecture to make it more suitable for question answering systems.

Blackboard architecture currently supports three of the knowledge management techniques, knowledge acquiring, knowledge representation and inference. Knowledge acquiring is done through the agents that architecture provides and representation of knowledge can be modeled by the designer. Inference is usually be implemented with scheduling programs which controls the knowledge source of the architecture. But still justification and validation is not performed for the knowledge acquired.

Justification is a human ability where represented knowledge can be justified by giving reasons to the selection. Therefore, it is difficult to transform this ability to an automated program and even it is transformed partially, an automated program cannot justify knowledge through tacit knowledge it has. When considering these issues, the best way to perform this is to usage of a collaborative effort towards the knowledge validation [6]. To build this collaborative knowledge validation and justification process there must be domain experts associated with the system and they must be given the full range of access to the entire knowledge base.

Another key point to notice is ways of improving knowledge representation and acquiring techniques, therefore it is vital to consider ways of knowledge representation which can support further query on the knowledge. Furthermore, knowledge acquiring can be implemented with machine learning techniques to makes it even more efficient, but in this scenario to manage accuracy again proper knowledge validation and justification technique is required.

### IV. Discussion and interpretation

Research findings suggest that blackboard architecture can be used in a question answering systems with few customizations that can bring the entire question answering process to a more effective and productive stage. But results or findings of this research cannot be generalized as it limits the generalization through focusing on question answering systems. Also research revealed several important qualities that blackboard architecture has which empower the architecture to be used in question answering systems.

### References


