Exploring the benefits of tagging forum posts based on a hierarchical domain model of the course content in online forums

by

Simarpreet Kaur Gill

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Computer Science

Guelph, Ontario, Canada
© Simarpreet Kaur Gill, December, 2014
EXPLORING THE BENEFITS OF TAGGING FORUM POSTS 
BASED ON A HIERARCHICAL DOMAIN MODEL OF THE COURSE CONTENT IN ONLINE FORUMS

Simarpreet Kaur Gill
University of Guelph, 2014

Advisor:
Dr. Judi McCuaig

Online forums are extensively used for educational purposes by students and teachers. Apart from numerous advantages, online educational forums have some limitations too. It is a common scenario that students have a certain query in their mind and they want to find an answer for that query from the online educational forum. Their intended information could be present in some forum posts, but the students could not find it due to the numerous branching out forum threads. The students thus get frustrated after some time and post their query as a new question on the forum, due to the difficulty in finding the relevant information from the online educational forum. This thesis proposes a solution to solve the problem of difficulty in finding most relevant forum posts for students’ queries by presenting an approach of content-based tagging of forum posts based on the hierarchical domain model of the course content.
Contents

List of Tables vi

List of Figures vii

1 Introduction 1

2 Background : Literature Review 4

2.1 Online Forums 4

2.2 Online Educational Forums 7

2.2.1 Success Factors for Online Educational Forums 7

2.3 Uses of Online Educational Forums 8

2.3.1 Virtual Online Communities 10

2.3.2 Collaborative Learning 13

2.3.3 Summary of Benefits of Online Educational Forums 15

2.4 Problems with Online Educational Forums 18

2.4.1 Lack of Information Synthesis and Summarization 19

2.4.2 Difficulty in Finding Relevant Information 20

2.4.3 Continuation of Misinforming Threads 21

2.4.4 Non-Contribution by Some Forum Members 22

2.5 Solutions to Improve Online Educational Forums 23

2.5.1 Recommender Systems 23
2.5.2 Text Summarization ........................................ 26
2.5.3 Classification-Based Approaches ....................... 28
2.5.4 Content-Based Tagging of Forum Posts ................. 30
2.5.5 Mediation by the Instructors ............................... 31
2.6 Conclusion of Literature Review ............................. 32

3 Proposed Solution ............................................. 34
3.1 Design of the Tagging Mechanism ......................... 35
3.2 Domain Model Addition and Use ............................. 36
  3.2.1 Taggedforum Use ........................................ 38
  3.2.2 Taggedforum Implementation Details .................. 41

4 Design of Experiment ......................................... 44

5 Data Analysis .................................................. 48
  5.1 Hypothesis 1: The success rate of participants in finding the most relevant posts for a particular query is higher in case of Taggedforum as compared to the Forum ..................... 51
  5.2 Hypothesis 2: The time taken by a participant in finding the most relevant forum posts for ones queries using taggedforum is less than the time taken by the participant using forum. ..... 55
  5.3 Hypothesis 3: The index of task load on each participant to find the most relevant posts for their queries is less in case of Taggedforum as compared to the Forum ..................... 57

6 Conclusions and Future Work ................................. 61

A Consent Form .................................................. 63

B Task-Sheet (for Taggedforum) ................................. 67
List of Tables

5.1 Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 1 ................................. 49
5.2 Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 2 ................................. 49
5.3 Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 3 ................................. 50
5.4 Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 4 ................................. 50
5.5 Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 5 ................................. 51
5.6 Success Rate and Total Time taken to complete all Tasks by each Participant ................................................................. 51
5.7 p-values for Wilcoxon Rank Sum Tests for Tasks 1, 2, 3, 4 and 5 .......................................................... 56
5.8 Summary of Results of Data Analysis .......................................................... 59
List of Figures

2.1 Web-interface for an Online Forum (Thomas and Molnar, 2014) . . . 5

3.1 Sample domain model . . . . . . . . . . . . . . . . . . . . . . . . . 37
3.2 Selection of Tag for a Post . . . . . . . . . . . . . . . . . . . . . . 38
3.3 Selection of Tag from the Nested List . . . . . . . . . . . . . . . . 40
3.4 Display of Posts According to the Selected Tag . . . . . . . . . . 40
Chapter 1

Introduction

Virtual online communities are online communities that facilitate communication between the members of a community without any physical meeting with each other. There are no time or location constraints in virtual online communities. Virtual online communities are used by the community members for group discussions and collaborative work without face-to-face meetings with each other (Damer and Bruckman, 1996).

Virtual online communities are often very rich in information, because they contain contributions from a number of community members. The discussion started by one individual keeps on growing by contributions from other community members. This situation sometimes leads to the problem of information overload in virtual online communities, which means that the community members are sometimes not able to find the exact intended information from enormous amount of information present in a virtual online community (Dwivedi and Bharadwaj, 2012).

Online forums provide an asynchronous online platform for a group of users to communicate and exchange ideas with each other in the group, thus leading
Online forums are used in a variety of domains, such as open source projects (Ahmed et al., 2009), social networking (Kio and Negreiros, 2013; Weninger et al., 2013), marketing (Andrlic, 2011), social computing (Ghosh and Hummel, 2012), customer support (Hong and Davison, 2009; Kim et al., 2010), community development (Hong and Davison, 2009), interactive reporting (Hong and Davison, 2009), education (Baxter et al., 2011b,a), etc. Online discussion forums have interesting applications in the domain of social networking and news like reddit (Weninger et al., 2013). This thesis focuses on the online forums used for the purpose of education. Online social communities like Usenet (Rowe et al., 2012) are precursors of online discussion forums.

Online educational forums are the online forums used for educational purposes. Online educational forums are the online forums used by students and teachers to communicate about the course-related concepts and activities. Online educational forums are used by students to share their knowledge about the course with their fellow-students.

This thesis focuses on the problem of difficulty in finding the most relevant forum posts for particular queries. Online educational forums have numerous discussion threads on a variety of discussion topics. Discussions in online educational forums often exhibit incoherence and have a non-linear structure. Students sometimes find the enormous amount of information to be problematic, as finding the most relevant information becomes cumbersome.
The thesis proposes a solution that consists of content-based tagging of forum posts based on the hierarchical domain model of the course content. A new Moodle plugin (Taggedforum) has been developed for the research, with the feature of content-based tagging of forum posts based on the hierarchical domain model of the course content. The new plugin (Taggedforum) is then evaluated by conducting a research experiment to check if the Taggedforum is better than the Forum (the original Moodle Forum plugin) or not.

The research experiment consists of making a group of university students use both the Moodle plugins (Taggedforum and the regular Forum), to perform a set of tasks involving finding relevant information from the forum. The data collected is statistically analysed using R. The results indicate that Taggedforum is better than the Forum in terms of success rate, time taken, mental load, temporal load, effort and frustration involved in finding the most relevant information from an online educational forum.

The chapter 2 contains the in-depth literature review of the area of online educational forums. Chapter 3 consists of the details of the proposed solution and chapter 4 consists of the details of the experiment conducted to evaluate the proposed solution. Results and analysis of the experiment are discussed in chapter 5 and chapter 6 discusses conclusions and future scope of the research.
Chapter 2

Background : Literature Review

This chapter explains what online forums are, how they are used in the domain of education and benefits of using online forums for education. The chapter also contains information regarding the problems with existing online forums being used for education and prior work done in improving them. The chapter concludes by stating the proposed solution and its relevance.

2.1 Online Forums

Online forums are online tools that provide platform for interaction and exchange of ideas among a group of users by allowing the users to generate their own content and read and respond to the content generated by other users (Baxter et al., 2011b,a; Hong and Davison, 2009). Messages are posted by users of an online forum, which are read and responded by other forum members (Baxter et al., 2011a; Robinson, 2011). The term ‘discussion threads’ refers to the topics posted (Baxter et al., 2011a; Marks, 2011; Morzy, 2009) and the term ‘posts’ is used for the replies or responses on online forums (Baxter et al., 2011a; Marks, 2011). Participants of online forum post their queries
I have been upgrading various mods / blocks via the plugins overview page. Most of them I have upgraded successfully but I am stuck on two mods both of which give me the “Plugin files not writable” error.

However, the permission on the mod directory and its sub directories are correct. In fact they are exactly the same as the blocks directory and I was able to upgrade a couple of blocks without issue.

I have tried changing the permissions to a+rw but they are still unwritable according to Moodle.

What is causing this and how can I fix it?

Thanks

- 

You need to change the entire directory where Moodle is installed to the ownership (or group) that apache is being run as. For example if you are on an Ubuntu install you would do:

```
chown -R www-data /home/lms/moodle
```

That is assuming your moodle is installed in /home/lms/moodle

problem is then solved.

---

Figure 2.1: Web-interface for an Online Forum (Thomas and Molnar, 2014)

on the forum, which are answered by the forum members who are interested in the discussion thread, as shown in figure 2.1, which shows the web-interface of an online forum, containing the main discussion topic started by a forum user and the responses by the forum members (Hong and Davison, 2009).

Online discussion forums have evolved from the email lists or email groups used for discussions in the group with a number of subscribed users. Online
discussion boards or online forums have advanced from the email lists to cater to the communication needs for a group of users (Tumbull, 2002). Online forums often revolve around a theme or a particular subject area, and so they can be defined as a “meeting place for people who are at least to a certain extent interested in or familiar with the topic” (Baxter et al., 2011b). Another comprehensive definition of online forums is:

“many-to-many communication space where participants can post a new topic and reply to an existing one. This communication is archived, and all of the threads are always available for reading and posting. Online forums may be public or private” (Baxter et al., 2011b).

Online forums are asynchronous means of computer-mediated communication (Patel and Aghayere, 2006; Minichiello and Hailey, 2013; Organero and Kloos, 2007; Camarero et al., 2012; Robinson, 2011; Meyer, 2003; Abel et al., 2010; Ioannou, 2011; Sani et al., 2013; Hong and Davison, 2009), and serve as an open and virtual discussion community that could be accessed anytime from any place (Minichiello and Hailey, 2013; Camarero et al., 2012; Robinson, 2011; Pieterse and van Rooyen, 2011; Thomas, 2002). Online forums thus allow more flexibility in schedules for participants. There is no requirement of simultaneous or concurrent logins (Patel and Aghayere, 2006; Marks, 2011), as they could be available for users 24/7 (Minichiello and Hailey, 2013; Robinson, 2011; Pieterse and van Rooyen, 2011).

Online threaded discussion forums also serve as communication environments where ideas are reciprocated, which means ideas put forward by one person in an online discussion are responded by other forum members who present their own views and ideas on the discussion topic (Thomas, 2002). Participants influence each others’ ideas (Curtis, 2001) by discussing topics on
online discussion board. Individual ideas and thoughts are thus reciprocated in a forum (Thomas, 2002).

In a nutshell, online discussion forums provide platform for anywhere/anytime interaction and information-sharing among forum users. Users of an online forum learn from each other and also take advantage from knowledge and ideas of other forum users. Online forums are used in a multitude of domains, and this thesis deals with online educational forums.

2.2 Online Educational Forums

Online educational forums are the online forums used for the purpose of education. Online forums are used by students to gain knowledge through collaboration, team-work and sharing information (Baxter et al., 2011b). In fact, as stated by Baxter et al. (2011a) and Godara et al. (2009), online forums have been used as tools for collaborative learning and knowledge sharing since ages.

2.2.1 Success Factors for Online Educational Forums

Making an online educational forum available for students does not guarantee that the educational forum will be successful. An educational forum is said to be successful only if it is accepted by students (Pieterse and van Rooyen, 2011). The success of an online educational forum thus relies on the rate of students' participation in the forum and the mediation by the instructors.

The use of an online educational forum by students for communication and knowledge-sharing is crucial for any online educational forum to be successful (Saltz et al., 2007). The adoption of an online educational forum by students further relies on the forum website layout, which should be free from technical bugs and support a navigation structure easy to use and convenient for
information-retrieval. The participation of students in an online educational forum also relies on usefulness of the educational forum, which is measured in terms of the relevance of the educational forum content and the forum’s ability to make students better understand the course concepts and enhance deeper learning. The forum should have the ability to facilitate online debates and dialogues, and support asynchronous interaction between students and teachers (Camarero et al., 2012). An online educational forum must possess all these necessary features to make students keep visiting the forum.

Instructors should act as communication facilitators in an online educational forum (Pieterse and van Rooyen, 2011). Conversation on online bulletin boards is mostly student-centred, but the instructors need to play an important role as mediators or moderators (Robinson, 2011). They need to perform important tasks such as controlling the continuation of threads containing some misinformation (Picciano, 2002), motivating and encouraging students to participate in online discussions (Chai et al., 2009).

It can be summarised that for an online educational forum to be successful, it must possess the features that are required for the forum to be extensively used by students under the control and mediation by the course instructors. Online educational forum should have the capability to make students accept it as a learning tool (Camarero et al., 2012). Students should find the forum useful for their course-related activities.

2.3 Uses of Online Educational Forums

Online educational forums are highly adaptable (Baxter et al., 2011a) online tools that can enhance students’ learning experiences (Baxter et al., 2011b;
Thomas, 2002; Pieterse and van Rooyen, 2011; Ioannou, 2011; Marks, 2011; Patel and Aghayere, 2006) by providing a platform for interaction and information-exchange among students (Baxter et al., 2011a). Students use online educational forums to ask questions about the course, which their peers or instructors could answer (Organero and Kloos, 2007). These online tools allow knowledge creation and sharing through forum threads and posts (Baxter et al., 2011a). Online discussion forums thus support team-work and collaborative learning (Baxter et al., 2011a; Marks, 2011; Organero and Kloos, 2007).

Online educational discussion forums are useful for academics, because they are based on a pedagogical model focused on collaborative learning (Hiltz et al., 1999; Hamer et al., 2011), and have a number of advantages for students. These advantages include enhancement in motivation levels (Abel et al., 2010; Minichiello and Hailey, 2013; Robinson, 2011), improvement in critical thinking and analytical skills (Baxter et al., 2011a; Meyer, 2003; Thomas, 2002; Pieterse and van Rooyen, 2011), better understanding of course topics (Thomas, 2002), increase in involvement in the course activities (Pieterse and van Rooyen, 2011; Robinson, 2011; Meyer, 2003) and more opportunities for doing group-work with fellow-students (Pieterse and van Rooyen, 2011; Robinson, 2011; Tan and Jones, 2008). Using online forums for education results in deeper learning and better performance of students (Picciano, 2002; Meyer, 2003; Thomas, 2002; Pieterse and van Rooyen, 2011; Ioannou, 2011; Oliver, 2003; Marks, 2011; Hiltz et al., 1999; Robinson, 2011; Patel and Aghayere, 2006).

As discussed by Meyer (2003), online educational discussion forums can be beneficial for course instructors too. They enable the instructors to develop more detailed and better responses for the students’ queries after reflecting on them. In this way, online educational forums ensure better answers to students’ questions.
Online educational forums have tremendous potential in aiding in the learning process. Online discussion forums being used for the purpose of education provide platform for the establishment of a virtual online community outside the classroom (Pieterse and van Rooyen, 2011; Camarero et al., 2012; Organero and Kloos, 2007). Online educational forums also provide support for instructional approaches involving collaborative learning, such as the Constructivist Education Model (Meyer, 2003; Hiltz et al., 1999) and Contributing Student Pedagogy (Hamer et al., 2011).

It can thus be said that, online educational forums are useful for academic purposes. They are useful for both students and teachers. Online educational forums have tremendous potential in enhancing the learning experience of students. Online discussion boards have thus taken a prominent position in students’ lives (Pieterse and van Rooyen, 2011).

2.3.1 Virtual Online Communities

A virtual online community can be defined as a community that allows community members to communicate with each other without meeting physically and with no time or location constraints (Pieterse and van Rooyen, 2011). Online educational discussion forums create a fully-functional virtual social community outside the classroom, where students communicate, share knowledge and ideas with their peers and instructors. Students work together to resolve each others’ issues (Pieterse and van Rooyen, 2011).

Virtual social communities established by online educational forums act as a medium for asynchronous interaction among learners, with the facility that the learners could access these interactions at any desired time. The interactions are saved permanently as forum postings in the online educational forum (Hammond, 2005; Meyer, 2003; Thomas, 2002; Macdonald, 2003; Abel
et al., 2010; Ioannou, 2011; Sani et al., 2013; Godara et al., 2009; Patel and Aghayere, 2006), leading to enhancement of retention levels and decline of frustration levels in students (Patel and Aghayere, 2006). Such communities are adaptable not only to the time and place differences between students and instructors, but also to other individual differences such as the pace of learning differences among learners (Minichiello and Hailey, 2013). In other words, asynchronous nature of online virtual communities allows individuals to communicate collaboratively, without any time or location restrictions, and thus facilitating learning in anywhere/anytime environment (Lewinson, 2005; Robinson, 2011; Meyer, 2003; Thomas, 2002; Godara et al., 2009).

The virtual online communities thus promote student-to-student and student-to-teacher interaction even outside the classroom (Patel and Aghayere, 2006). They aid in reducing the gap among students and course instructor (Meyer, 2003). As discussed by Godara et al. (2009), and Organero and Kloos (2007), virtual online communities formed by online educational forums aid in real-time knowledge sharing and collaboration, and are used as a supplement to class discussions (Brush et al., 2002).

The virtual online communities contribute positively towards student learning (Ellis and Cohen, 2009; Picciano, 2002; Abel et al., 2010; Sani et al., 2013; Patel and Aghayere, 2006; Marks, 2011; Curtis, 2001; Oliver, 2003) by enhancing the satisfaction levels of students with the course leading to students’ achievements. They support the mantra that knowledge will be multiplied on sharing. Individuals exchange knowledge with each other, thus facilitating the formation of group knowledge from each individual’s individual knowledge (Brush et al., 2002).
As per Patel and Aghayere (2006), a student’s participation rate in virtual online communities established by online educational forums is directly proportional to student learning. The same study also establishes that students are more comfortable in asking questions through the virtual online communities than asking them in class, because students could ask questions any time, even before or after the class, and there are many people available to answer those questions. Also, a large amount of class time is saved due to the virtual online communities which serve as virtual question-answer platforms.

Moreover, it has been observed that most of the questions come in students’ minds outside the classroom, while doing homework, and thus virtual online communities are convenient and useful for them (Patel and Aghayere, 2006; Meyer, 2003). In their support for virtual online communities, students also argue that their peers’ responses are sometimes more understandable as compared to the responses from their teachers. The virtual online communities also help students by increasing communication among classmates, and by involving the whole class into one discussion group (Patel and Aghayere, 2006).

Some students might feel out of place or alienated in a classroom discussion, and so the virtual online communities could be of great advantage for such shy students. They can interact with their fellow students and instructors through posts and threads, without any face-to-face interaction. In other words, the virtual online communities could be beneficial for students who are not comfortable interacting in a physical community (Picciano, 2002; Thomas, 2002; Pieterse and van Rooyen, 2011; Patel and Aghayere, 2006), thereby reducing the student drop-out rates (Pieterse and van Rooyen, 2011).

In a nutshell, online educational forums lead to the formation of an online virtual community accessible to students at any time from any place. Students
could access the virtual online community to ask questions and interact with their peers and teachers at any time as per their convenience. The virtual online communities thus provide flexible learning environments always available for students.

2.3.2 Collaborative Learning

Online educational forums also serve as a platform for some useful instructional approaches, such as, the Constructivist Education Model and Contributing Student Pedagogy. They thus aid in the learning process by supporting collaborative learning, which means emerging of knowledge from active communication and team-work among the group members (Camarero et al., 2012). Acquiring knowledge through communication could be much more effective and beneficial than merely learning through individual efforts (Thomas, 2002). Collaborative learning is beneficial not only for one individual, but for the entire communicating group in promoting deeper learning (Camarero et al., 2012; Macdonald, 2003; Willey and Gardner, 2012).

Collaborative learning allows students to benefit from other students’ knowledge and intellect in the group (Ling, 2013; Willey and Gardner, 2012). Complex problems can be solved easily while working in groups (Hiltz et al., 1999; Willey and Gardner, 2012). Each brain participating in collaborative learning is part of a distributed networked system working towards solving a complex problem. The cognitive load is divided among all the participants, which is in contrast to the individual learning system, where the entire cognitive load involved in solving a complex problem falls on one individual brain (Willey and Gardner, 2012).

Online educational forums provide platform for collaborative learning (Curtis, 2001; Sani et al., 2013) by supporting discussion, which is the most vital
component for collaborative learning (Curtis, 2001). Students discuss their ideas and queries with their peers and teachers, and construct knowledge from each others’ knowledge (Thomas, 2002; Abel et al., 2010). Moreover, this knowledge constructed from collaborative learning through online educational forums is better defended as it is accompanied by evidences from a number of users and sources (Meyer, 2003).

The Constructivist Education Model serves as a frame for collaborative learning. According to this model, each student in a group is responsible for creating or constructing knowledge. As per this model, each learner in a group is responsible for discovering knowledge and validating others’ knowledge (Hiltz et al., 1999). The constructivist or active approach to learning focuses on learning by communication within a group of individuals, and thus, establishes that learning is a process involving the society. Each individual puts forward ones own ideas, which are then either further established or changed following the reactions and replies from other people in the group.

Contributing student pedagogy is an instructional approach that supports collaborative learning. Contributing Student Pedagogy (CSP) is a pedagogical approach that involves each student in a group contributing towards others’ learning, and each student benefits from other students’ contributions in a group (Hamer et al., 2011). Students answer each others’ queries and work as a team with other group members. As part of CSP, the learning material created by one student is shared by all others in the group. This group of people involved in knowledge-sharing through online educational forum somewhat acts like a “Community of Practice” (CoP) (Hamer et al., 2011). CoPs are defined as communities or groups involving people tied together by means of shared expertise and a common goal (Baxter et al., 2011a; Godara et al., 2009).
As stated by Godara et al. (2009), the CoPs are centred around common goals and themes, facilitating communication between the community members. The CoPs thus serve as repositories of knowledge, and play a crucial role for knowledge sharing, knowledge maintenance and knowledge extension. New CoP members eventually become experienced members, and this is how a CoP evolves. Knowledge of each CoP member is shared with other CoP members, maintained and also extended by combining ones own knowledge with the knowledge of other CoP members.

It can be inferred that online educational forums aid in the learning process by providing support for collaborative learning and asynchronous anywhere/anytime learning. Online educational discussion forums provide a learning environment where each individual benefits from the knowledge and ideas of other individuals in a group. Knowledge is constructed by the combined efforts of all the contributing members in the forum.

2.3.3 Summary of Benefits of Online Educational Forums

Using online forums as interaction and collaboration tools for educational purposes have a number of advantages. These advantages include enhanced motivation levels of students (Abel et al., 2010; Minichiello and Hailey, 2013; Robinson, 2011), improved critical thinking and analytic skills of students (Baxter et al., 2011a; Meyer, 2003; Thomas, 2002; Pieterse and van Rooyen, 2011), better understanding of course topics by students (Thomas, 2002), increased involvement of students in course activities (Pieterse and van Rooyen, 2011; Robinson, 2011; Meyer, 2003) and opportunities for students to do group-work with their fellow-students (Pieterse and van Rooyen, 2011; Robinson, 2011; Tan and Jones, 2008). Online educational forums lead to enhanced learning
outcomes for students (Oliver, 2003; Meyer, 2003; Pieterse and van Rooyen, 2011).

Online educational forums help in keeping students motivated (Minichiello and Hailey, 2013; Robinson, 2011) to learn in a course by providing sufficient and timely feedback, either from instructors, teaching assistants, or fellow students. Also, students’ perception of helping others leads to enhancement of motivation to learn and contribute more to the forum (Minichiello and Hailey, 2013). Moreover, as stated by Abel et al. (2010), a participating student generally likes to defend against others the opinions put up by him on the forum, and thus, will regularly visit the forum and contribute in the discussions defending his/her opinions. The active participants will be motivated to learn more about the course in order to contribute regularly to the online discussions (Abel et al., 2010).

Online educational forums encourage students to develop critical thinking about the course contents and topics. Students think deeply and critically about the topics being discussed on the online educational discussion board. They critically evaluate others’ postings and reflect on them (Baxter et al., 2011a; Meyer, 2003; Thomas, 2002). Participants of online educational forums take their time to go through the available posts, thinking about the replies to the posted questions before replying, and linking forum posts to bring out some knowledge (Meyer, 2003).

Students can devote much more time to any topic in a threaded discussion as compared to the face-to-face classroom discussion (Meyer, 2003; Pieterse and van Rooyen, 2011), and are thus involved in the critical review of information through the process of reflecting on their fellow-students’ postings in online educational forums. They thus construct ideas that are not individualistic, but a result of collective or combined efforts of peers (Thomas, 2002). Students
thus get involved in critical thinking about the matter put up on the forum, and might have to change their attitudes or beliefs, as a result of this (Thomas, 2002).

Students better understand the course topics with the use of online educational forums as tools for communicating with their peers and instructors and tend to post more multi-structural and relational messages on the forum. This behaviour is an indicator of an increase in the cognitive engagement levels for students. Students develop relationships between course topics (Thomas, 2002).

Students get more involved in the course activities by participating in the online educational discussion boards. Students feel encouraged to share their knowledge, problems and progress with their peers. Their interest and involvement in the course activities thus increases (Pieterse and van Rooyen, 2011; Robinson, 2011).

As per the study by Robinson (2011), a vast majority of students find online educational forums as useful engagement and participative tools. The facts that the students could get some time to reflect on the posted messages, before drafting replies for them, and students who have some queries, do not have to wait for their turn to ask questions in class, but could simply post their queries on the forum, are responsible for engaging students in the course. The findings as part of the study described by Meyer (2003) indicate that discussions occurring on online forums continue for several hours or even for several days. In this way, students devote a large amount of time to a particular course. The amount of time devoted by a student to a particular course can further indicate the level of that student’s commitment towards learning that course.
Online discussion forums act as means for providing opportunities for working in groups or teams aiming towards the goal of solving course-related problems and understanding the course topics better (Pieterse and van Rooyen, 2011). As per the study by Robinson (2011), students perceive online forums to be convenient for group discussions and interactions regarding group projects. The study by Tan and Jones (2008) claims that communication and collaboration among the group members is vital for the success of any group project. The authors of the study found that the online forum was used as a basic tool for discussion, posting of documents, task coordination, and communication among the group members.

It can be inferred that online educational forums are beneficial for education and have a great potential in enhancing student learning (Oliver, 2003; Meyer, 2003; Pieterse and van Rooyen, 2011), and leading to better performance of students (Pieterse and van Rooyen, 2011). They keep the students motivated and engaged in the course activities, and improve their critical thinking and analytic skills. Students gain knowledge through online educational forums not by any declarative or lecturing means, but by collaborative learning. Online educational forums have the capability of making students learn to the same extent from their classmates as from the teachers or course materials (Thomas, 2002; Hiltz et al., 1999; Ioannou, 2011; Picciano, 2002).

2.4 Problems with Online Educational Forums

Online educational forums are beneficial for education, but they have certain limitations too that affect the student use of online forums for education. As per Robinson (2011), although a large population of students finds online forums to be of great value for education, and actively participate in online discussions, some students find online educational forums to be of limited use, and as such
do not actively participate in them. The problems identified in the existing online educational forums include incoherent and non-converging nature of online discussions (Minichiello and Hailey, 2013; Thomas, 2002; Ioannou, 2011), non-linear structure of online educational forums (Picciano, 2002; Thomas, 2002), difficulty in finding the most relevant forum posts for one’s queries (Picciano, 2002; Minichiello and Hailey, 2013; Thomas, 2002; Gottipati et al., 2011; Sani et al., 2013; Hong and Davison, 2009; Si et al., 2011; Krishnamani et al., 2013), threads containing some misinformation can keep on running (Picciano, 2002) and non-contribution by some forum members (Oliver, 2003; Thomas, 2002).

2.4.1 Lack of Information Synthesis and Summarization

The discussions in online discussion forums are often incoherent and non-converging (Thomas, 2002; Ioannou, 2011). Students keep on contributing to the discussion thread freely, which sometimes might lead to incoherence in an online forum. This means one post might not be built upon the previous one (Thomas, 2002). The discussion thread thus starts growing along a number of divergent paths, and there is no synthesis and summarization of information, which means that discussions in online forums are non-converging or diverging (Ioannou, 2011). As stated by Minichiello and Hailey (2013), and Ioannou (2011), the diverging nature of online threaded discussions acts as a limitation for online forums to be used for education. Learners are prevented from getting the central idea or the core of the discussion.

Online discussion forums thus have a non-linear or divergent structure, having multiple threads, where a large volume of discussions continue to grow and branch out simultaneously (Picciano, 2002; Thomas, 2002). Thus the communication is not truly conversational, which means that the communication is not linear and revolving around one particular theme (Thomas, 2002).
online threaded discussions keep on branching out and fragmenting as they evolve (Thomas, 2002). Each student is observed to follow his/her own path in a discussion, and thus students are not united but diverging out along a number of paths (Thomas, 2002).

Face-to-face discussions are linear, centred around a single discussion theme (Picciano, 2002). Online forums are sometimes not able to support the same kind of communication as is possible in face-to-face interaction. Face-to-face discussions are truly conversational as compared to the asynchronous online discussions.

2.4.2 Difficulty in Finding Relevant Information

The enormous amount of student interaction in an online educational forum can sometimes lead to overload of information (Picciano, 2002; Si et al., 2011; Krishnamani et al., 2013). This has been cautioned by Picciano (2002), by stating that too much information could sometimes be a problem. Students face problems in finding the information they want from an online forum (Gottipati et al., 2011; Sani et al., 2013; Hong and Davison, 2009; Si et al., 2011; Krishnamani et al., 2013).

The organization of topics in the order of their times of origin, and not according to the conceptual ideas of each topic, inhibits the process of deep learning by students because they are not able to find relationships between a number of concepts being discussed on the forum (Minichiello and Hailey, 2013; Ioannou, 2011). Students are not able to place the forum contributions in correct contexts. They are not able to make out the central theme of the discussion because of historical organisation of forum posts (Ioannou, 2011).

Students need to scroll through each post of an extremely long discussion thread till they find a relevant post for their query. And if the information
they want is towards the end of the discussion thread, they may get frustrated and tend to quit before reaching there (Robinson, 2011; Thomas, 2002). As discussed by Thomas (2002), as the number of forum posts in a discussion thread increases, there is a proportional decline in the hit rate (the number of times a post is read) for each post. It can be inferred that students tend to explore or read less before contributing to the forum, leading to a large number of duplications or repetitions in the forum. They end up posting their queries as new posts instead, thus starting a new divergent path and leading to duplication in forum posts.

There is also a possibility that a post might be available in an online forum which could be relevant for the student’s query, but that post could be in a wrong thread. And thus, such a post could not have been found by the student (Abel et al., 2010).

When reaching to the relevant posts for ones queries by traversing the discussion threads is cumbersome, keyword-based search could come to the rescue of the students. But keyword-based search is not considered a complete solution because students find it difficult to use keywords to express their search queries (Abel et al., 2010).

2.4.3 Continuation of Misinforming Threads

Each student has complete freedom to start a discussion thread. Students also have full freedom to create a new post or respond to existing posts. This freedom is good, because it is in line with the open structure of online educational forums, but such complete freedom and independence could also be problematic in some situations (Picciano, 2002).
The complete independence to contribute leads to the possibility that some students might knowingly or unknowingly contribute some incorrect information in the forum. The posts containing misinformation have the possibility of being read by other students, who might learn the incorrect information believing it to be correct. Moreover, they will also respond to some of the existing posts with wrong information they learnt from misinforming threads.

The discussions containing some misinformation could thus keep on growing, unless interrupted by the instructor or some knowledgeable student (Picciano, 2002). Students could thus learn some incorrect information by participating in online discussions. So online discussions could contribute negatively towards student learning, if threads containing incorrect information are allowed to grow.

2.4.4 Non- Contribution by Some Forum Members

Some students don’t make contributions in online educational forums. Such students only tend to read or find information from the discussions in an online educational forum. They do not make their own contributions to the discussions (Oliver, 2003).

There could be many reasons for such lurking or non-contributing behaviour. Some students might just feel lost or out of pace in the ever-growing online discussions. Some students might face the problem of lack of proper understanding of the concepts being discussed. Delay in receiving timely feedback for ones queries (Castro-Herrera, 2010; Curtis, 2001), also leads to dissatisfaction among students leading to their non-contributing behaviour in the forum. Such students lose confidence to create their own posts and contribute to the online discussions.
The behaviour of lurking or avoiding to make own contributions to the forum, hinders the process of collaborative learning (Oliver, 2003), because collaborative learning requires contributions from each student in a group (Camarero et al., 2012; Hiltz et al., 1999; Thomas, 2002). An important purpose of having an online educational forum is thus unfulfilled. It can thus be said that non-contribution in online discussions by some of the students could ultimately lead to the failure of the online forum.

2.5 Solutions to Improve Online Educational Forums

There is a scope for improvement in online educational forums, so that they could serve the students’ needs better. Use of recommender systems or collaborative filtering techniques (Abel et al., 2010; Castro-Herrera, 2010; Budalakoti et al., 2009; Kardan, 2013), content-based tagging of forum posts (Gottipati et al., 2011; Kim et al., 2010; Bateman et al., 2007), text-summarization techniques (Sani et al., 2013; Krishnamani et al., 2013; Ezzat et al., 2012), classification-based approaches (Hong and Davison, 2009; Cong et al., 2008; Ezzat et al., 2012) and mediation by the instructors (Picciano, 2002; Chai et al., 2009) could be solutions to improve online educational forums. These solutions could help in making online forums better suited for educational purposes.

2.5.1 Recommender Systems

Recommender systems are the systems used for filtering of unseen information. They are also used for predicting if a particular resource would be liked by a particular user or not. Machine learning approaches and techniques are used by recommender systems to do their functions (Ghazanfar and Prugel-Bennett, 2010).
Recommender systems can be used to find the relevant posts for a particular query (Abel et al., 2010). Recommender systems can also be used to find the relevant forum users who could answer a particular query posted on the forum, thereby reducing the delay in replying to queries posted on the forum. Students will then get timely feedback for their queries (Castro-Herrera, 2010; Budalakoti et al., 2009; Kardan, 2013).

A recommender system for educational discussion forum ‘Comtella-D’ was evaluated by Abel et al. (2010). They studied various sources through which users could provide feedback in an online educational discussion forum and what should be the appropriate amount of interactions among users to generate recommendations. Their evaluation resulted in a personalization based rule, which could choose the best suitable strategy for recommendation on the basis of input data generated by users.

A recommender system was proposed by Castro-Herrera (2010) to dig out potential forum members to answer particular queries put up on the forum. The author evaluated his proposed system using a cross validation statistical technique, and his results suggested that his system had set up a benchmark in terms of accuracy and recall. He also presented a relationship analysis between forum threads and users.

A recommender system for finding the appropriate forum members to reply to particular queries was also proposed by Budalakoti et al. (2009). The forum users were found to be more satisfied with the forum, because their queries were being answered quickly. The authors developed two metrics to evaluate their proposed solution, which included, load on the responders and satisfaction levels of persons asking questions. They used the dataset from Yahoo! Answers for their experiment. They conducted experiments covering a wide range of
domains and concluded that their proposed system was able to find appropriate responders for specific queries quickly and efficiently.

The recommender system’s application for digging out perspective users of an online educational forum for replying to particular queries was also proposed by Kardan (2013). The algorithm proposed by the author to find out appropriate forum users (experts) to answer specific queries classified all forum members based on their knowledge levels and then used filtering mechanism on the dataset derived. The evaluation of the proposed solution provided evidence that his proposed algorithm had made the process of getting replies for ones queries quicker and efficient.

The recommender systems require users to rate each others postings for generating recommendations. But the users who seldom post on the forum, may not even rate others’ postings. Due to the heterogeneous nature and amount of data generated by users of an online forum, an individual recommendation algorithm might not be appropriate for generating recommendations (Abel et al., 2010).

Moreover, most of the educational discussion forums are either built from scratch or adapted with certain specifications, as most of the online education platforms cannot use generic discussion forums because of data structures, style of programming or licensing issues. Due to this, tightly integrated recommendation algorithms are of limited use for generating quality recommendations. A loosely-coupled Web Service-based algorithm has been proposed as a solution by Abel et al. (2010), using which different discussion forums could retain their original features along with implementation of recommendation algorithm.

Collaborative filtering systems are types of recommender systems. These systems are used for recommending items by considering the particular users’ preferences for specific items. Collaborative filtering systems function with
two assumptions. The first assumption is that the users under consideration are interested in the specified items. The second assumption is that the users similar to the users under consideration have highly rated the specified items (Ghazanfar and Prugel-Bennett, 2010).

A number of recommendation techniques are based on collaborative filtering, which are of limited use in the domain of online educational systems. The primary reasons for this are that either some users don’t rate each others’ postings or there are not sufficient users using the system (Abel et al., 2010). To deal with these challenges, Abel et al. (2010) have proposed systems that either are extensions or replacements of collaborative filtering techniques, such as clustering, rule-based techniques or stereotyping. The appropriate techniques for social structures similar to discussion forums, could even be a combination of collaborative filtering and natural language processing (Abel et al., 2010).

It has been concluded by Abel et al. (2010) and Castro-Herrera (2010) that appropriate solution for online educational discussion forums could be a recommender system that will be hybrid and flexible on the basis of input data. Such a recommender system could be a combination of a variety of different personalisation techniques (tau Yih et al., 2004), in order to counter the limitations of personalisation techniques that are single. The example system could be a combination of collaborative and content-based recommender systems, using Web Ontology Language (OWL).

### 2.5.2 Text Summarization

As discussed by Sani et al. (2013), Krishnamani et al. (2013), and Ezzat et al. (2012), text summarization could be helpful in managing the enormous amount of information in online discussion threads, and finding relevant information in an online forum. Text summarization could thus be useful in finding answers to
queries posted by users. Text summarization tools have been proposed by Sani et al. (2013) and Krishnamani et al. (2013), with the capability of presenting the users with a summary of the best answers selected for their queries or questions.

In general, text-summarization techniques for a large volume of forum posts available for a user query are of two different types. There is one approach that involves extracting information or sentences from the enormous volume of information to present summarised information to the users. The other approach is to generate sentences or tags for conveying the central idea of a large volume of information (Sani et al., 2013). The first approach was used by Sani et al. (2013) in their tool, as their tool involved extracting the summarised information from a large volume of posts.

The tool proposed by Sani et al. (2013) was evaluated by collecting data of the tool usage with 30 discussion contents. The authors found that their text summarization tool was successful in making them accept their hypothesis, which means, that the users were able to find the most relevant information for their queries. Hence, the users were able to learn better by using the forum (Sani et al., 2013).

A method for automatic text-summarization of forum posts was proposed by Krishnamani et al. (2013). Their method was based on models of the topics and content-based clustering. The method proposed by Krishnamani et al. (2013) involved document summarization, which is a data mining technique to generate summaries of documents. Document summarization could be used for online forums by considering each forum post as a separate document. The implementation challenges in their approach were posed by the structural differences between forum messages and traditional documents. In spite of the
challenges, the authors were able to propose a novel method to summarize online forum posts.

The authors evaluated their proposed solution by using the database provided by University of Arizona (Krishnamani et al., 2013). The database included over 25 web forums. Each web forum consisted of nearly 40000 discussion threads, and each thread consisted of about 100 messages. Only forum messages with minimum 5 words were chosen for summarization. Summaries generated by human beings were used to compare the summaries generated by their proposed method, and the authors concluded that their proposed solution was useful.

Text-based summarization of forum posts could thus be a solution to solve the problem of difficulty in finding most relevant forum posts for ones queries. Users could find relevant forum posts for their queries by going through or searching by means of text-based summaries of forum posts. Text-summarization could also lead to improvement in online forums.

2.5.3 Classification-Based Approaches

Classification-based techniques or approaches were used by Hong and Davison (2009), and Cong et al. (2008) to determine if a particular post was a question or an answer to a question. The authors tested their techniques using various datasets. The results of their experiments suggested that their approaches were successful in finding whether a specific post was a question or an answer.

The classification-approach used by Hong and Davison (2009) was able to determine the category of post (question or answer) using both content-based as well as non-content based post features. The features used were whether the post was ending with a question mark, type of words used, count of posts
in a thread and author of the post. The authors did a comparison of their approaches with other popular approaches in the same domain.

The experimental results suggested that their approach was successful in extracting useful information from the online forum. Their approach could find the questions and answers to those answers. This was helpful for online forum users to find their target information (Hong and Davison, 2009).

A classification method based on sequential patterns to find out questions in discussion forum threads has been used by Cong et al. (2008). The authors also used a graphical propagation method to dig out forum posts that were answers to certain questions in the same discussion thread. The authors evaluated their approach using three different discussion forums of different scales and concluded that their approach was useful in finding out questions and answers in discussion threads.

The use of a classification-based approach to handle enormous amount of data in online discussion forums was proposed by Ezzat et al. (2012). Their approach involved categorization of text based on ontologies and multiple labels. The authors evaluated their approach using data sets involving large numbers of posts, and found that their proposed solution showed high performance and accuracy.

The classification-based approaches could thus be used to classify or categorise forum posts using content based or non-content based approaches. The classification techniques could thus simplify the search of desired information. The classification-based approaches could thus lead to improvement in online forums.
2.5.4 Content-Based Tagging of Forum Posts

Tagging forum posts based on the content of posts could be a solution to the problem of difficulty in finding the most relevant forum posts for one's queries (Gottipati et al., 2011; Kim et al., 2010; Bateman et al., 2007). Forum posts can be grouped together according to their content-specific tags (Kim et al., 2010; Bateman et al., 2007). The process of finding relevant posts for one's queries could thus be simplified (Gottipati et al., 2011; Kim et al., 2010; Bateman et al., 2007).

A content-based semantic tagging approach was adopted by Gottipati et al. (2011), and Kim et al. (2010) with the aim of simplifying the problem of finding relevant posts for queries posted by users. The tagging approach described by Gottipati et al. (2011), and Kim et al. (2010) tagged different forum posts into different categories based on a set of labels created for forum posts. The posts were tagged to be either questions, answers, positive feedback, negative feedback, junk, clarifying answers etc., based on the contents of posts.

A text classification based engine was used to tag the forum posts. The individual words of each of the posts were used in the text classification approach. The authors then analysed 6068 forum posts from three different online forums (Gottipati et al., 2011). The authors were able to achieve an average precision of 67 percent, recall of 71 percent and F-measure of 69 percent. They also conducted a study assisted by users, which suggested that their proposed solution could elevate the mean precision in finding relevant posts to 71 percent from 17 percent.

The authors of (Kim et al., 2010) evaluated their proposed solution by conducting an experiment with 320 discussion threads. The 320 discussion threads contained 1332 posts in total, across a wide range of domains. The
results achieved were above the baseline decided by the authors in finding the relevant posts for particular queries.

A collaborative content-based tagging approach was proposed by Bateman et al. (2007). Each forum member could add tags to the forum posts. The experiment conducted to evaluate the approach consisted of 200 participants who were asked to tag the forum posts based on the contents of posts. In addition to the tagging by students, forum posts were also tagged by domain experts and automatic tag generator (Bateman et al., 2007).

Students generated 211 tags in total, out of which unique tags were 122. The experts generated 60 tags, out of which the unique tags were 38. The automatic tag generator created 25 tags in total, out of which the unique tags were 24. The tags generated by students, experts and the automatic tag generator were analysed and compared with each other (Bateman et al., 2007). The authors of (Bateman et al., 2007) concluded that generation of a set of tags by experts, and then making the set of tags available to students for tagging could be an ideal solution (Bateman et al., 2007).

It can thus be inferred that content-based tagging of forum posts could be a solution to solve the problem of difficulty in finding relevant posts or information in an online forum. Tagging forum posts based on their content and displaying the posts according to tags simplifies the information-retrieval from an online forum. Content-based tagging of forum posts could thus improve online forums.

2.5.5 Mediation by the Instructors

Course instructors or teachers could play an important role as discussion moderators. They could play a vital part in mediating the continuation of threads
containing some misinformation (Picciano, 2002). They could also help in motivating and encouraging students to participate in online discussions (Chai et al., 2009).

Course instructors or teachers have all administrative privileges to control or moderate discussions in a forum (Ellis and Cohen, 2009). Students can sometimes contribute incorrect information to the online discussions based on their knowledge or understanding of the concepts (Picciano, 2002). Course instructors should then mediate by monitoring discussions and discontinue threads containing some misinformation (Picciano, 2002).

Participation of students in online discussions is crucial for an online educational forum to be successful (Saltz et al., 2007). Course instructors could thus contribute to the success of the forum by motivating students to participate in discussions in an online forum (Chai et al., 2009). Online educational forums with appropriate mediation by course instructors were found to be successful and accepted by students (Sani et al., 2013).

2.6 Conclusion of Literature Review

Online forums are important educational tools but have certain limitations. This thesis aims to propose and test a solution for the problem in finding most relevant forum posts for one’s query in an online forum. It can be inferred from the literature review that content-based tagging could be a solution to improve online educational forums. Also, making the central theme of the forum posts easily visible to students, could make the search for relevant information faster.

As part of this thesis, we thus propose a solution involving content-based tagging of forum posts based on a hierarchical domain model of the course content constructed using xml-owl. The display of forum posts according to
their specific tags could be a form of text-summarization technique, as students will be able to look for answers to their queries by just going to the specific tag corresponding to the domain of their problem.

The proposed solution is tested by developing a Moodle plugin (Tagged-forum) with the feature of tagging forum posts based on their content. The plugin is then tested by making a group of university students use both the plugins (Original Forums plugin and the new Taggedforum plugin) to do a set of tasks. The data collected from their experiences is then analysed to reach to the results. The details of the proposed solution are explained in the next chapter.
Chapter 3

Proposed Solution

The literature review suggests that content-based tagging of forum posts can improve online educational forums. The literature review also suggests that tagging systems work better when students are provided with a set of tags to use. The proposed solution, thus, consists of adding a provision of tagging forum posts based on course content and then displaying the posts according to the tags. Each student and teacher can add tag to their own posts as well as others’ posts. Each post would thus be associated with one or more tags, and the students will be able to search forum posts according to the content-based tags.

A plugin, named ‘Taggedforum’, has been developed for Moodle (a Learning Management System), which includes a feature of tagging forum posts based on a hierarchical domain model of the course content. The new plugin contains the entire basic functionality of the regular Forums plugin. Taggedforum has an additional provision of tagging forum posts by students and teachers.

Taggedforum is different from the conventional tagging systems, as the set of tags are generated using the hierarchical domain model of the course content.
Taggedforum makes use of ontologies. The screen-shots of the user-interface of Taggedforum are attached as Appendix H.

3.1 Design of the Tagging Mechanism

Students and teachers can add single or multiple tags to their own posts as well as posts created by other users. The course instructor has the permission to delete any tags if he/she finds irrelevant tags associated with any post. Students can delete tags only from their own posts.

Tags can be selected only from the set of tags obtained by parsing of the hierarchical domain model of the course content. The domain model is created and added to the Taggedforum by the course instructor. The domain model is used because the literature review indicates that tagging mechanism works better if students are given a set of tags to use, instead of students freely creating their own tags.

The domain model contains the hierarchy of tags in a specific domain. The domain model consists of the course topics and their respective sub-topics arranged in a hierarchical way. Each course topic and sub-topic included in the domain model is included in the set of tags to be used for tagging.

The set of tags generated by parsing the domain model can be used by students and teachers to tag the forum posts, based on the content of the posts. Students and teachers can apply tags to their posts while creating the posts. Single or multiple tags can be applied to the forum post by selecting any number of tags from the set of tags to be used for tagging. The tags associated with a forum post are highlighted when any user wants to edit the post or reply to that post.
Tags can also be applied to the posts created by other users, by both students and teachers. Tags can be applied to the posts created by other users by editing the posts, and then selecting new tags to be added from the set of tags. Students can only add new tags, if they are editing a post not created by them.

Teachers have the permission of deleting tags from posts, if irrelevant tags are associated with any post. Students can delete tags only from their own posts. Tags can be deleted by editing the posts and selecting the tags to be deleted from the set of tags associated with the posts.

The application of tags to the forum posts based on the course domain model can be used to display or search the forum posts based on their content. The forum posts can be displayed according to their respective tags. The posts can be displayed according to the hierarchy of tags. The tags can even be displayed in a hierarchical way and users can simply click or check on the appropriate tag to find forum posts related to a particular course topic or sub-topic.

3.2 Domain Model Addition and Use

While adding a new Taggedforum, the domain model is added to the Taggedforum by the course instructor using the Moodle file picker. The domain model is parsed by the Taggedforum plugin to produce a set of tags to be used by the forum users to tag forum posts. Adding the domain model is thus necessary for the taggedforum functionality, and so adding the domain model is made mandatory.

The domain model is an xml-owl file. The classes in the domain model are the course topics, and each of the sub-classes are the course sub-topics. A part
of domain model of a course on C programming language is shown in figure 3.1, which shows the parent-child relationship and hierarchy between some possible tags related to C language. The complete domain model used for research is included as Appendix G.

It can thus be inferred that a course domain model contains the course topics and their respective sub-topics arranged in a hierarchy. The domain model is parsed to generate a set of tags to be used for content-based tagging, and then display forum posts according to their tags. The domain model is used to order forum posts according to the content of posts.
3.2.1 Taggedforum Use

The set of tags generated by parsing the domain model are displayed in a drop-down menu when a student or teacher is creating new post or editing any post. The student or teacher creating a new post or editing a previously created post can select one or multiple tags from the drop-down menu, as shown in figure 3.2. Figure 3.2 shows one tag selected for the post from the drop-down menu. Selection of at least one tag is obligatory while creating a new post. Each post is thus associated with one or more tags.

Students and teachers can reply to any post. The selected tags for the parent post are highlighted for the users to see the set of tags associated with the parent post. The users can then either add new tags to this selected set of tags, or delete some tags from this selected set (tags would not be deleted from the parent post) to generate a new set of tags for the reply post. Users thus can generate a new set of tags for the reply post while retaining some, all or none of the tags from the parent post.

Students and teachers can edit any post. Teachers have the authority of editing the content of any post, but students can only add new tags to the post and do nothing else if the post is not created by the editing student. The tags selected for the post are highlighted, so that the student or teacher editing the
post knows the tags already associated with the post. Students and teachers can create a new set of tags for the post by adding tags to the selected set of tags. If the user is a teacher, he/she can even delete some or all of the tags previously associated with the post, but students can delete tags only from their own posts.

Taggedforum supports all the forms of display of posts supported by the regular Forum. This means that the display of posts flat, with oldest or newest first, display of posts in nested and threaded form are supported by Taggedforum. In addition to these conventional displays, Taggedforum supports three additional forms of displays. These additional display forms include flat display ordered by tags, nested display ordered by tags and display of posts selected by particular tags.

Flat display of posts ordered by tags displays all the posts in a Taggedforum according to their tags. Posts with a particular tag are placed together, which means posts are grouped together according to their tags. Students can thus find the most relevant forum posts for their queries by looking for the specific tag/tags or domain associated with their query, and then the posts underneath those tag(s).

The nested display of posts ordered by tags, displays the tags associated with all the posts in Taggedforum according to their hierarchy in the domain model. The posts are displayed underneath their respective tags. Users can find the most relevant forum posts for their queries more easily using this display as compared to the flat display ordered by tags. This display displays the tags according to their hierarchy, which makes it easier for the user to reach to the desired tag.

The display of posts selected by particular tags displays the perspective tags as generated by parsing the course domain model in the form of nested
lists according to their hierarchy. Figure 3.3 shows a part of the nested list, with one tag selected. The user can select one or multiple tags, and posts corresponding to those tags will be displayed on clicking the 'Submit' Button. Users can find the most relevant forum posts for their queries using this display, by just clicking on the appropriate tag(s). Figure 3.4 shows the posts displayed by selection of a tag from the nested list.
It can thus be inferred that Taggedforum plugin aims to make finding the most relevant forum posts for user queries easier by associating one or more content-specific tags with each post. Students and teachers can tag their own as well as others’ posts. The posts can be displayed or retrieved according to the tags displayed in hierarchical order.

### 3.2.2 Taggedforum Implementation Details

The taggedforum plugin is a Moodle plugin, so it is coded using Moodle APIs and Libraries. The programming language used for development is PHP. HTML and JavaScript are used for designing the user-interface, and the database used is MySQL.

Moodle Filepicker element is used to upload the domain model while creating a new Taggedforum. The contents of the domain model are saved in the database when the details of the new Taggedforum are saved in the database. The domain model content is stored in a string variable, which is then converted into an array. The following piece of code shows the storage of domain model contents into a string variable, which is then converted into an array.

```php
$cont = $DB->get_field('taggedforum_dmcontents', 'dmcontents', array('taggedforumid' => $taggedforum->id));

$cont2 = str_replace('>', ' ', $cont);
$cont3 = str_replace(']', ' ', $cont2);
$cont4 = str_replace('−−', ' ', $cont3);
$cont5 = str_replace('/', ' ', $cont4);
$cont6 = str_replace('′', ' ', $cont5);
```
The domain model contents are retrieved from the database table and stored in a string variable. Some specific symbols are removed from the string variable, and the string is then converted into an array by using the PHP explode function, such that wherever the symbol ‘<’ is encountered in the string, a new array element is created. The whole domain model is now split into elements of an array, such that each line of the domain model is a new array element.

For each element of the array, a loop runs that extracts the tag by retrieving the text after the hash symbol from each of the domain model lines containing tags. The following piece of code is the part of the loop that extracts tags and saves them in the database.

```php
foreach ($strarr as $y=>$y_value) {
    if (( (strstr(trim($y_value)," Class") !=") and
        (strstr(trim($y_value)," about") !=") ) ||
        (strstr(trim($y_value)," subClassOf") !=") ){
        $record->pscontents =
            ltrim(strstr(trim($y_value)," #")," #");
    }
}
```

The above piece of code extracts the tags and saves them in the database. The unique set of tags generated is displayed using Moodle multiple-select element, when a user is creating a new post or editing a previously created post. The tags selected by the users for each post are stored in a database table.

In order to display the tags in hierarchical form, the tags generated are stored in a multi-dimensional array. The first element of each column of the
multi-dimensional array is the root element (tag with no parents), and the subsequent column elements are the children of the root element in hierarchical form. The multi-dimensional array has been used to find parents, children and siblings of each tag to construct the hierarchy of tags, and the Moodle display APIs are used to build the display of posts.

The Taggedforum plugin is tested according to the experiment, explained in Chapter 4, to determine if the proposed solution makes finding the most relevant forum posts for one’s queries easier and efficient.
Chapter 4

Design of Experiment

The goal of this thesis is to determine whether Taggedforum is better than the regular Forum or not. As such, an experiment has been designed to evaluate the following hypothesis:

1. The success rate of participants in finding the most relevant posts for a particular query is higher for Taggedforum as compared to the Forum.
2. Participants take less time to find the most relevant posts for their query using Taggedforum compared to the Forum.
3. The index of task load on each participant to find the most relevant posts for their query is lower for Taggedforum compared to the Forum.

In general, the experiment requires participants to perform tasks using Taggedforum and the Forum. The two types of forums are then compared based on success rate and time taken to complete each task.

Participants recruited for the experiment had knowledge of the basic concepts of C programming because the domain model for the experiment was designed for a course on C programming language. Participants also had some experience using online forums. Participation in the research was voluntary,
and participants were recruited through email. The ideal number of participants for such experiment could be nearly 100. Over 600 students were emailed, with the target of getting at least 20 participants. But only 8 students gave consent to participate. The participants were free to ask questions any time during the experiment. On the arrival of the participant, he/she was given an instruction sheet after a brief round of instructions. Participants were required to fill a consent form prior to the experiment. The consent form has been included as Appendix A. Participants were free to withdraw from the experiment any time according to their will. A pizza lunch was organized for all participants where they were also provided with the summary of results of the experiment.

Participants were alternately divided into two groups; Group A and Group B, based on their order of being called in for the experiment. Participant who came in first belonged to Group A, second participant belonged to Group B, third participant to Group A, fourth to Group B, and so on. Group A members used Taggedforum first and then Forum, while Group B members used Forum first and then Taggedforum. The purpose of dividing participants into two groups and the group members using both the plugins alternately was that after finding posts for a set of queries using a particular plugin, the participant would have gained some knowledge about the posts to be found. Since the posts to be found were actually the same for both the plugins, so this prior knowledge about the posts would have influenced the participants’ performance with other plugin. The purpose of changing the order of using plugins for the two groups was that one group used the Taggedforum with prior knowledge about the posts and the other group used the Forum with prior knowledge about the posts. In this way, the effect of prior knowledge on the use of both the plugins was balanced.
If the participant belonged to Group A, he/she was given the task-sheet for Taggedforum first and then for the Forum. If the participant belonged to Group B, he/she was given the task-sheet for Forum first and then for Taggedforum. The task-sheets contained a set of queries, for which the users were requested to find the most relevant posts assuming that they did not know the answers to those queries. The task-sheets for both plugins have been included as Appendix B (Task-Sheet for Taggedforum) and Appendix C (Task-Sheet for Forum). The queries in both the task-sheets were similar, and the posts to be found were the same. The participants could perform the tasks in any order, but could take maximum 10 minutes to perform each task.

The data for the experiment were collected using two duplicate sets of posts created for both the plugins (Forum and Taggedforum). The set of posts (used for both the plugins) is included as Appendix E. Each post was given a unique number (post number) by the person conducting the experiment. The post numbers of the posts found by the research participants as part of the experiment, were recorded in a record-sheet. The time taken to perform each task by each of the participants was recorded using a stop watch and then recorded in a record-sheet. NASA (National Aeronautics and Space Administration) Task Load Index sheets were used to record the mental load involved in performing each task by each of the participants. NASA Task Load Index Sheet used has been included as Appendix D.

The post numbers of the posts found by the participants for each of the tasks were recorded in the record-sheet. If a participant was not able to find any post for a particular task, the query corresponding to the respective task was recorded in the record-sheet. This data was used as metrics for Hypothesis 1, as this data could determine the success rate for each participant. The time taken by the participant to perform each task was recorded using a stop-watch.
and noted in the record-sheet by the person conducting the experiment as a metric for Hypothesis 2. The participant was requested to fill in the NASA Task Load Index sheet after performing a set of tasks with each plugin as metrics for Hypothesis 3. The participant was free to leave after this.

The participants were requested to use both the plugins (Forum and Tagged-forum) to perform some designated tasks, and then fill in the NASA Task Load Index sheets (one for each task-sheet) given to them. The data collected was recorded in the record-sheet and the NASA Task Load Index Sheets, which were analysed to derive results. The data in the record-sheet was recorded by the person conducting the experiment (author of this thesis).

After conducting the experiment with all the participants, the observations and data were analysed to derive results. The three types of tests were performed for the statistical analysis of the data collected using R (Kehrer et al., 2012) (version 3.11). The three types of tests included the Test of Two Proportions (E.L., 2005), Mann Whitney U-Test (also known as Wilcoxon sum test) and the Wilcoxon signed rank test (E.L., 2005). Test of two-proportions and Wilcoxon sum test require the two populations to be independent, and the Wilcoxon signed rank test requires the two populations to be dependant on each other (E.L., 2005). Analysis of data and results of the experiment are discussed in chapter 5.
Chapter 5

Data Analysis

This chapter contains the detailed analysis of the experimental data collected from eight participating students. The data was analysed to find out whether TaggedForum is better than Forum in terms of success rates in finding relevant information, the time spent in finding the required information, along with the mental load, effort and frustration involved. The analysis of the data provides evidence that TaggedForum has an edge over Forum in finding the most relevant information for one’s queries.

The experimental data consists of data collected from 8 participants, who volunteered to participate in the experiment. The success rates of participants in finding the most relevant forum posts for the queries, the time in seconds taken by participants to find the desired information, and the data in the NASA Task Load Index Sheets were recorded as experimental data. The tables 5.1 to 5.5 show the experimental data, which show the success rates and the times taken (in seconds) to accomplish a set of tasks involving finding the most relevant information for the queries from the forum. Tables 5.1 to 5.5 present data collected for each of the tasks (Task 1 to Task 5) from 8 participants,
divided into 2 groups- Group A and Group B. The set of tasks to be performed by the participants are included as Appendices B and C.

Table 5.1: Outcome(Successful or Unsuccessful) and Time Taken by each Participant to perform Task 1

<table>
<thead>
<tr>
<th>Taggedforum</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Forum</th>
<th>Time Taken</th>
<th>Successful or Not</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1 (GroupA)</td>
<td>32.50</td>
<td>Successful</td>
<td>2.08</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 3 (GroupA)</td>
<td>49.80</td>
<td>Successful</td>
<td>67.60</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 5 (GroupA)</td>
<td>39.40</td>
<td>Successful</td>
<td>28.40</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 7 (GroupA)</td>
<td>113.60</td>
<td>Successful</td>
<td>79.10</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 2 (GroupB)</td>
<td>2.40</td>
<td>Successful</td>
<td>2.330</td>
<td>Unsuccessful</td>
<td></td>
</tr>
<tr>
<td>Participant 4 (GroupB)</td>
<td>24.50</td>
<td>Successful</td>
<td>159.40</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 6 (GroupB)</td>
<td>24.30</td>
<td>Successful</td>
<td>257.60</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 8 (GroupB)</td>
<td>20.80</td>
<td>Successful</td>
<td>87.30</td>
<td>Successful</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2: Outcome(Successful or Unsuccessful) and Time Taken by each Participant to perform Task 2

<table>
<thead>
<tr>
<th>Taggedforum</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Forum</th>
<th>Time Taken</th>
<th>Successful or Not</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1 (GroupA)</td>
<td>1.13</td>
<td>Unsuccessful</td>
<td>43.50</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 3 (GroupA)</td>
<td>82.30</td>
<td>Successful</td>
<td>7.40</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 5 (GroupA)</td>
<td>70.80</td>
<td>Successful</td>
<td>6.40</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 7 (GroupA)</td>
<td>62.80</td>
<td>Successful</td>
<td>61.80</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 2 (GroupB)</td>
<td>60.40</td>
<td>Successful</td>
<td>117.40</td>
<td>Successful</td>
<td></td>
</tr>
<tr>
<td>Participant 4 (GroupB)</td>
<td>58.50</td>
<td>Successful</td>
<td>325.20</td>
<td>Unsuccessful</td>
<td></td>
</tr>
<tr>
<td>Participant 6 (GroupB)</td>
<td>56.10</td>
<td>Successful</td>
<td>132.80</td>
<td>Unsuccessful</td>
<td></td>
</tr>
<tr>
<td>Participant 8 (GroupB)</td>
<td>22.20</td>
<td>Successful</td>
<td>95.40</td>
<td>Successful</td>
<td></td>
</tr>
</tbody>
</table>

Each participant had an individual success rate (rate of success in completing tasks) and the total time taken to complete the tasks. Table 5.6 shows the success rate (out of 10) of each individual participant. The table also shows the total time (in seconds) taken by each participant.

As stated in Chapter 4, the data collected was analysed to determine whether the three hypotheses for the experiment could be rejected or not. The first hypothesis for the research experiment was that the success rate of participants in finding the most relevant posts for a particular query
Table 5.3: Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 3

<table>
<thead>
<tr>
<th>Participant</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Taggedforum</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Forum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1 (GroupA)</td>
<td>1.03</td>
<td>Successful</td>
<td>27.00</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 3 (GroupA)</td>
<td>25.60</td>
<td>Successful</td>
<td>10.00</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 5 (GroupA)</td>
<td>18.60</td>
<td>Successful</td>
<td>68.60</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 7 (GroupA)</td>
<td>19.90</td>
<td>Successful</td>
<td>16.30</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 2 (GroupB)</td>
<td>1.16</td>
<td>Successful</td>
<td>20.70</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 4 (GroupB)</td>
<td>25.40</td>
<td>Successful</td>
<td>45.40</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 6 (GroupB)</td>
<td>34.50</td>
<td>Successful</td>
<td>20.60</td>
<td>Partially Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 8 (GroupB)</td>
<td>0.00</td>
<td>Successful</td>
<td>189.00</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4: Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 4

<table>
<thead>
<tr>
<th>Participant</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Taggedforum</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Forum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1 (GroupA)</td>
<td>54.00</td>
<td>Successful</td>
<td>21.00</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 3 (GroupA)</td>
<td>26.10</td>
<td>Successful</td>
<td>8.10</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 5 (GroupA)</td>
<td>64.40</td>
<td>Successful</td>
<td>7.40</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 7 (GroupA)</td>
<td>21.70</td>
<td>Successful</td>
<td>66.50</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 2 (GroupB)</td>
<td>40.10</td>
<td>Successful</td>
<td>27.50</td>
<td>Unsuccessful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 4 (GroupB)</td>
<td>24.90</td>
<td>Successful</td>
<td>90.30</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 6 (GroupB)</td>
<td>57.20</td>
<td>Successful</td>
<td>55.70</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 8 (GroupB)</td>
<td>68.40</td>
<td>Successful</td>
<td>20.60</td>
<td>Successful</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

is higher in case of Taggedforum as compared to the Forum. The second hypothesis stated that the participants take less time to find the most relevant posts for their query using Taggedforum as compared to Forum. The third hypothesis was that the index of task load on each participant to find the most relevant posts for their query is less in case of Taggedforum as compared to the Forum.
Table 5.5: Outcome (Successful or Unsuccessful) and Time Taken by each Participant to perform Task 5

<table>
<thead>
<tr>
<th>Participant</th>
<th>Time Taken</th>
<th>Successful or Not</th>
<th>Participant</th>
<th>Time Taken</th>
<th>Successful or Not</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1 (GroupA)</td>
<td>21.00</td>
<td>Successful</td>
<td>Participant 3 (GroupA)</td>
<td>17.20</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td>Participant 5 (GroupA)</td>
<td>43.40</td>
<td>Successful</td>
<td>Participant 7 (GroupA)</td>
<td>76.90</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td>Participant 2 (GroupB)</td>
<td>1.10</td>
<td>Successful</td>
<td>Participant 4 (GroupB)</td>
<td>89.40</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td>Participant 6 (GroupB)</td>
<td>57.60</td>
<td>Unsuccessful</td>
<td>Participant 8 (GroupB)</td>
<td>76.30</td>
<td>Unsuccessful</td>
</tr>
</tbody>
</table>

Table 5.6: Success Rate and Total Time taken to complete all Tasks by each Participant

<table>
<thead>
<tr>
<th>Group</th>
<th>Participant id</th>
<th>Success Rate</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>8</td>
<td>203.24</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>9</td>
<td>329.9</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
<td>9</td>
<td>388.2</td>
</tr>
<tr>
<td>A</td>
<td>7</td>
<td>9</td>
<td>522</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>8</td>
<td>362.49</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>8</td>
<td>1128.7</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>6.5</td>
<td>800.7</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>9</td>
<td>611</td>
</tr>
</tbody>
</table>

5.1 Hypothesis 1: The success rate of participants in finding the most relevant posts for a particular query is higher in case of Taggedforum as compared to the Forum

In the process of accepting or declining Hypothesis 1, four statistical tests were conducted. These tests included a test of two-proportions, 2 Mann Whitney U-tests and 1 Wilcoxon signed rank test. A test of two proportions assumes that the two proportions are independent. Mann-Whitney U-test assumes that the two populations are independent and Wilcoxon signed rank test assumes that the two populations are related to each other. All tests discussed in this
A test of two-proportions was performed to determine if there was statistically significant difference between the average success rates of Group A and Group B. The test was performed using prop.test function in R.

Null hypothesis: The probable average rate of success of the two groups (Group A and Group B) is statistically not significantly different.

Alternative hypothesis: The probable average rate of success of the two groups (Group A and Group B) is significantly different.

The test gave a p-value of 0.4555, which is greater than the significance level of 0.05, so there was insufficient evidence to reject the null hypothesis, and so it could be said that the 2 populations did not have statistically significantly different proportion of successes.

It could thus be inferred that there is statistically no significant difference between the average total success rates of participants of the two groups.

It was thus interpreted that the average success rates of the two groups were statistically not significantly different, so the next step was to determine whether the distributions consisting of each participant’s success rate (by group), were statistically significantly different or not. This was done using Mann Whitney U-test.

Null hypothesis: The distributions of success rates of participants of Group A and Group B are statistically not significantly different.

Alternative hypothesis: The distributions of success rates of participants belonging to Group A and Group B, are statistically significantly different.
The test was performed using the `wilcox.test` function in R.

The test gave a p-value of 0.2059, which was higher than the significance level of 0.05. So there was insufficient evidence to reject the null hypothesis, and thus, it could be said that success rates of GroupA and success rates of GroupB were statistically not significantly different populations.

It could thus be inferred that the success rates achieved by using Taggedforum and the success rates achieved using forum are independent of the order in which the two plugins are used. Mann Whitney U-test was then performed on the difference between participant success rate with Taggedforum and the Forum (Success rate with taggedforum - Success Rate with Forum).

Null hypothesis: The distribution of differences between the success rates using Taggedforum and the success rates using forum is statistically not significantly different for the two groups

Alternative hypothesis: The distribution of differences between the success rates using Taggedforum and the success rates using forum is statistically significantly different for the two groups.

The test gave a p-value of 0.06317, which was again greater than the significance level of 0.05. So there was insufficient evidence to reject the null hypothesis, and thus, it could be said that the distribution of differences between the success rates using Taggedforum and the success rates using forum were statistically not significantly different for the two groups.

It could thus be inferred that the differences between the success rates achieved by forum and the success rates achieved by Taggedforum are independent of the order of using the two types of forums.
From the three tests conducted so far, it could be inferred that success rates of Group A and success rates of Group B are statistically not significantly different populations. It could thus be said that the order of using the two plugins (Taggedforum and forum) does not affect the success rate of a participant. The learning effect (the affect of knowing about the posts to be found while using a plugin) is thus balanced for both the groups. The two populations could thus be combined to determine whether the success rates achieved by Taggedforum and the success rates achieved by forum are statistically significantly different or not. Wilcoxon signed rank test was thus the next step in the analysis.

Specifically this test was performed to compare the two related populations (success rates of participants using Taggedforum and the success rates of participants using forum).

Null hypothesis: The two populations Taggedforum and forum are statistically not significantly different.

Alternate hypothesis: The two populations Taggedforum and forum are statistically significantly different

The test gave a p-value of 0.01991, which was less than the significance level of 0.05, so there was sufficient evidence to reject the null hypothesis. So it could be said that the two populations, Taggedforum and forum were statistically significantly different in terms of success rates.

It could thus be inferred that the success rates achieved by Taggedforum are different from the success rates achieved by forum. The literature review supports that content-based tagging of forum posts could be beneficial in finding the most relevant forum posts for ones queries, and it could thus be inferred that the success rates achieved by Taggedforum are higher as compared to the success rates achieved by forum.
There is thus insufficient evidence to reject the Hypothesis 1 that the success rates in finding the most relevant forum posts for one’s queries achieved by Taggedforum are higher as compared to the success rates achieved by forum. It could thus be concluded that the success rates achieved by Taggedforum are higher as compared to the forum, and the success rate of a participant in finding the most relevant information from the forum is independent of the order of using the two types of forums. This conclusion is based on our sample data which is small in size, but the statistical analysis has sufficient evidence to indicate the trend that makes us not to reject our Hypothesis 1.

5.2 Hypothesis 2: The time taken by a participant in finding the most relevant forum posts for one’s queries using Taggedforum is less than the time taken by the participant using forum.

Hypothesis 2 was tested using the Wilcoxon rank sum test between the two independent populations - total time taken by each Group A participant and the total time taken by each Group B participant in performing all five tasks. This test showed that there was no statistically significant difference between the two populations (p-value=0.1143), so a second Wilcoxon sum test was performed between the two independent populations- difference between the total time taken using forum versus the total time taken using Taggedforum for each participant of Group A and the difference between the total time taken using forum versus the total time taken using Taggedforum for each participant of Group B. This test indicated that there was statistically significant difference between the two populations (p-value=0.02857) and Group A and Group B.
were statistically significantly different groups in terms of times taken to complete tasks. So tests were conducted for each of the 5 tasks, to determine where the differences were. The details of all the tests are included in Appendix F.

Specifically, Wilcoxon rank sum tests were performed to test the following hypothesis.

Null hypothesis The total time taken by GroupA participants to complete task X (using Taggedforum and forum) is statistically not significantly different from the total time taken by GroupB participants to complete task X using both kinds of forums.

Alternate hypothesis- The total time taken by GroupA participants to complete task X (using Taggedforum and forum) is statistically significantly different from the total time taken by GroupB participants to complete task X using both kinds of forums.

X= 1,2,3,4,5

The p-values for each of the tests are shown in table 5.7.

Table 5.7: p-values for Wilcoxon Rank Sum Tests for Tasks 1,2,3,4,5

<table>
<thead>
<tr>
<th>Task</th>
<th>Wilcoxon Rank Sum Test1</th>
<th>Wilcoxon Rank Sum Test2</th>
<th>Wilcoxon Signed Rank Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8857</td>
<td>0.05714</td>
<td>0.4609</td>
</tr>
<tr>
<td>2</td>
<td>0.05714</td>
<td>0.02857</td>
<td>Not Performed</td>
</tr>
<tr>
<td>3</td>
<td>0.6857</td>
<td>0.8857</td>
<td>0.1094</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.4857</td>
<td>0.5469</td>
</tr>
<tr>
<td>5</td>
<td>0.02857</td>
<td>0.1143</td>
<td>Not Performed</td>
</tr>
</tbody>
</table>

From the individual tests conducted for all five tasks, it was found that the differences existed only in case of task 2 and task 5. The reason for the differences in case of task 2 could be attributed to the observation during the experiment that while performing task 2 using Taggedforum, some participants had difficulty in placing the problem in correct domain. So, this was a slight variation as compared to other tasks. The differences in case of task 5
could be attributed to the observation that for answering task 5 query using Taggedforum, participants could do it directly by seeing the display of tags. So, there was again a slight variation in case of task 5. It could thus be inferred that typically there was no significant difference between the time taken by a participant to complete a task using Taggedforum and the time taken by the participant to complete the similar task using forum.

5.3 Hypothesis 3: The index of task load on each participant to find the most relevant posts for their queries is less in case of Taggedforum as compared to the Forum

The NASA Task Load Index Sheet were filled by each of the participants for each of the plugins- Taggedforum and Forum. The Task Load Index had five components - Mental Demand, Temporal Demand, Performance, Effort and Frustration. The information filled by the participants was analysed manually to make out observations and reach to some inferences or conclusions.

7 out of 8 participants felt that less mental demand was involved in completing tasks using Taggedforum as compared to forum. One participant felt that he had less mental load when doing tasks with forum, but he already had knowledge about the posts to be found before, because he had used Taggedforum already to find the same set of posts. Majority of the Group A participants (who used Taggedforum before forum) were of the view that lower mental load was involved with Taggedforum. It can thus be said that finding relevant posts for ones queries using Taggedforum is less mentally demanding as compared to the forum, which is consistent with our literature.
4 out of 8 participants were of the view that lower temporal demand was involved with Taggedforum as compared to the forum. One Group A participant was of the view that same temporal demand was involved with Taggedforum and forum, and two Group A students were of the view that lower temporal demand was involved with forum as compared to the Taggedforum. But this could be attributed to the fact the Group A participants had already used Taggedforum before using forum, and so they had the prior knowledge about the post. One participant of Group B was of the view that the temporal demand was same with both types of forums. So, it could be said that content-based tagging of forum posts could be better in terms of temporal demand involved in finding the most relevant forum posts for ones queries, which is consistent with our literature review.

6 out of 8 participants were of the view that their performance was better with Taggedforum as compared to the forum. One participant of Group A was of the view that his performance was better with forum was better as compared to Taggedforum, because he already had knowledge of the posts to be found while using forum. Only one participant of Group B was of the view that his performance was higher with forum, but it could still be said that participants showed higher performance with Taggedforum as compared to forum, which is consistent with our literature review too that content-based tagging forum posts could be a potential solution to the problem of difficulty in finding the most relevant forum posts for ones queries.

6 out of 8 students observed that less effort was involved in completing tasks using Taggedforum as compared to forum. Two Group A students were of the view that lower effort was involved with forum, because they were using forum after using Taggedforum, and so they already had some prior knowledge about the posts to be found, it could thus be said that tagging forum posts involves
using less effort in finding the most relevant forum posts for one's queries, which is consistent with our literature review.

6 out of 8 participants were of the view that their frustration levels were lower with Taggedforum as compared to the forum. One participant of Group A was of the view that the frustration levels with both types of forums were same and another student of Group A was of the view that frustration levels were lower with forum, but this could be attributed to the fact that they had used Taggedforum before using forum. So, it could be said that participants had lower frustration levels with Taggedforum as compared to the forum, which is consistent with our literature review that content-based tagging of forum posts could be helpful in finding the most relevant forum posts for one's queries.

It can thus be inferred that, the analysis of the NASA Task Load Index Sheets filled by our small number of participants, indicates the pattern that finding the most relevant information using Taggedforum involves lower mental demand, lower temporal demand and lower effort as compared to the regular forum. Taggedforum also leads to enhanced performance levels and lower frustration levels as compared to the regular forum.

The results of the data analysis are summarised in Table 5.8.

<table>
<thead>
<tr>
<th>Factor Measured</th>
<th>Taggedforum</th>
<th>Forum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success Rate</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>Time Taken</td>
<td>Same</td>
<td>Same</td>
</tr>
<tr>
<td>Mental Demand</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Temporal Demand</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Effort</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Performance Level</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>Frustration Level</td>
<td>Lower</td>
<td>Higher</td>
</tr>
</tbody>
</table>

These data suggest that performing a set of tasks involving finding relevant posts for particular queries could be done with higher success rate, lower mental
and temporal demands, lower effort and frustration levels and higher perceived
performance with Taggedforum as compared to forum. Although the times
taken for both types of forums were statistically not significantly different, but
we can still say that Taggedforum is better than forum because of the higher
success rate and performance levels with lower mental demand, lower temporal
demand, lesser effort and lower frustration levels as compared to the regular
forum.
Chapter 6

Conclusions and Future Work

The goal of the research was to find out whether content-based tagging of forum posts based on hierarchical domain model of the course content could improve online educational forums or not. The Moodle plugin developed, Taggedforum, was compared with the regular Forum Moodle plugin. The hypothesis for the research included that the success rate of participants in finding the most relevant posts for a particular query is higher in case of Taggedforum as compared to the Forum, participants take less time to find the most relevant posts for their query using Taggedforum as compared to the Forum, and the index of task load on each participant to find the most relevant posts for their query is less in case of Taggedforum as compared to the Forum.

It has been found from the data analysis that the success rate of participants in finding the most relevant posts for a particular query is higher in case of Taggedforum as compared to the Forum. The time taken to find the most relevant forum posts for a particular query has been found to be same for both types of forums, and the index of task load on each participant has been found to be lower in case of Taggedforum as compared to the Forum. The data analysis has thus revealed that Taggedforum is better than Forum.
It was also observed that the participants of the experiment were more satisfied and comfortable with Taggedforum as compared to the Forum. Some participants were of the view that Taggedforum should be made available as an online forum for a course.

The limitation of the research that the sample size is very small. But although the sample size for the study was small, but the statistical analysis indicated the pattern in favour of the proposed solution. Another limitation of the research is that the order of performing tasks by participants was not randomised, and all participants performed tasks in the same order.

Like every research, this research also has a potential for future work. The plugin (Taggedforum) can be tested with a larger number of students. Testing can also be done to find out if the proposed solution can also help in solving the problems of incoherence, diverging nature and delay in receiving replies for ones queries in online educational forums. New features like allowing students to create their own tags could also be added in future.
Appendix A

Consent Form
Exploring the benefits of tagging forum posts based on a hierarchical domain model of the course content in online forums

You are invited to participate in a research study conducted by Dr. Judi McCuaig and Ms. Simarpreet Gill, from the School of Computer Science at the University of Guelph. The results of this study will be used in the completion of Simarpreet Gill’s MSc thesis.

If you have any questions or concerns about the research, please feel free to contact the person conducting the experiment.

PURPOSE OF THE STUDY
This explores the benefits of tagging forum posts based on a hierarchical domain model of the course. It has been observed from the literature review in the area of using forums as an educational tool, that students find problems in finding the most relevant posts for their questions. Instead of finding the relevant posts, they post their query as a new forum post. The study accesses whether tagging forum posts based on hierarchical domain model of the course solves this problem.

PROCEDURES
If you volunteer to participate in this study, we would ask you to do the following things:
Perform some tasks as designated by the person conducting the experiment, following the instructions in the instruction sheet given to you. You are also requested to fill in the NASA Task Load Index sheets (for Forum and Taggedforum) based on your experience in doing the tasks. You are also invited for a pizza lunch that would be organized after the experiment would be done with all participants, where a summary of results of the study would be provided to you. The lunch will be for maximum 1 hour.

POTENTIAL RISKS AND DISCOMFORTS
There are no foreseeable risks associated with participating in this study.

POTENTIAL BENEFITS TO PARTICIPANTS AND/OR TO SOCIETY
There are no direct benefits to participants in this study. The results from this study will be used to improve future implementations of online forums.

PAYMENT FOR PARTICIPATION
There is no payment in cash for participation. Each participant will be eligible for the pizza and pop lunch as discussed above. The lunch will be organized in the Reynolds Building (Reynolds312).
CONFIDENTIALITY
We will not collect any identifiers during this research. You will be identified by a number only.

PARTICIPATION AND WITHDRAWAL
You can choose whether to be in this study or not.

RIGHTS OF RESEARCH PARTICIPANTS
You may withdraw your consent at any time and discontinue participation without penalty. You could ask for the withdrawal of your data any time during the experiment session. If you would want to withdraw your data after your experiment session, you must be knowing your participant id to have your data withdrawn. This study has been reviewed and received ethics clearance through the University of Guelph Research Ethics Board. If you have questions regarding your rights as a research participant, contact

Director, Research Ethics Telephone: (519) 824-4120, ext. 56606
University of Guelph E-mail: sauld@uoguelph.ca
437 University Centre Fax: (519) 821-5236
Guelph, ON N1G 2W1
I have read the information provided for the study “Studying the use of online forums” as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

Name of Participant (please print)

Signature of Participant

Date

Name of Witness (please print)

Signature of Witness

Date
Appendix B

Task-Sheet (for Taggedforum)
Task –Sheet (for Taggedforum)

Please find the most relevant forum posts using Taggedforum for answering these questions (the actual answers for these questions are not required), assuming that you don’t know the answers to these questions.

1) What will happen if the loop counter reaches the upper boundary value of its data type and is incremented?

2) What is the position of the first occurrence of ‘an’ in the following sentence- ‘Another man also has an umbrella’?

3) How can we stop the execution of a loop?

4) If I write x=y in a piece of code, and then increment the value of y. Would the value of x also be incremented?

5) What are the basic data types in C?

You will be given maximum 10 minutes for completing each task.
Appendix C

Task-Sheet (for Forum)
Please find the most relevant forum posts using Forum for answering these questions (the actual answers for these questions are not required), assuming that you don’t know the answers to these questions.

1) What will happen if the loop counter reaches the lower boundary value of its data type and is decremented?

2) What is the position of the first occurrence of ‘a’ in the following sentence- ‘An apple a day keeps the doctor away’?

3) How can we halt the execution of a loop?

4) If I write a=b in a piece of code, and then increment the value of b. Would the value of a also be incremented?

5) What are the basic data types in C?

You will be given maximum 10 minutes for completing each task.
Appendix D

NASA Task Load Index Sheet
NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

<table>
<thead>
<tr>
<th>Participant #</th>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
</table>

**Mental Demand**

How mentally demanding was the task?

- Very Low
- Very High

**Physical Demand**

How physically demanding was the task?

- Very Low
- Very High

**Temporal Demand**

How hurried or rushed was the pace of the task?

- Very Low
- Very High

**Performance**

How successful were you in accomplishing what you were asked to do?

- Perfect
- Failure

**Effort**

How hard did you have to work to accomplish your level of performance?

- Very Low
- Very High

**Frustration**

How insecure, discouraged, irritated, stressed, and annoyed were you?

- Very Low
- Very High
Appendix E

Set of Posts (Same for both Plugins)
What is an if block?

The group of statements following an if condition.

What is the default scope of if statement?

The statement immediately after if statement.

I wrote the following program:

```c
#include<stdio.h>
```

#include<stdio.h>
main()
{
    int i;
    printf("Enter the value of i");
    scanf("%d", &i);
    if(i=5)
        printf("You entered 5");
    else
        printf("You entered something other than 5");
}
Even if I enter the number 200 or any other number, I get the output "You entered 5". Please explain.

Re: 6-C Fundamentals
by Simarpreet Gill - Tuesday, 17 June 2014, 10:13 AM
You have used assignment operator(=) in the if condition. Use the relational operator(==).

Re: 7-C Fundamentals
by Simarpreet Gill - Tuesday, 17 June 2014, 10:11 AM
if(5)
Is this statement valid?

Re: 8-C Fundamentals
I wrote the following program:

```c
#include<stdio.h>
main()
{
 int i;
 printf("Enter the value of i");
 scanf("%d", &i);
 if(i==5);
 printf("You entered 5");
}
```

The program is not producing any output. Please help.
I wrote the following program, which crashes on running instead of printing numbers from 1 to 32767. Please help.

```c
#include<stdio.h>

void main()
{
    int i=1;
    while(i<32767)
    {
        printf("%d\n", i);
    }
```
T - His father is the professor.

INDEX(T,THE) = 7 or 15?

---

7

---

I wrote the following piece of code:

```c
#include<stdio.h>

void main(){
  int x=0, y=0;
  x=y;
  x=x+1;
  x=y;
  y=y+1;
  printf("%d,%d", x,y);
}
```
It should produce the output:

\[ x = 1, \]
\[ y = 1, \]

but it is producing:

\[ x = 0 \]
\[ y = 1 \]

Please explain.
Appendix F

Details and R Code of Statistical Tests done for Hypothesis 1 and Hypothesis 2

Hypothesis 1: The success rate of participants in finding the most relevant posts for a particular query is higher in case of Taggedforum as compared to the Forum In the process of accepting or declining Hypothesis 1, four statistical tests were conducted. These tests included a test of two-proportions, 2 Mann Whitney U-tests and 1 Wilcoxon signed rank test. Test of Two-Proportions: A test of two-proportions was performed to determine if there was statistically significant difference between the average success rates of Group A and Group B. Null hypothesis : The probable average rate of success of the two groups (Group A and Group B) is statistically not significantly different. Alternative hypothesis: The probable average rate of success of the two groups (Group A and Group B) is significantly different. The test was performed using prop.test function in R, using two vectors. The first vector contained the average total
success rates for each of the groups and the second vector contained the number of trials or tasks for each of the groups. successes <-c(35,31.5) trials <-c(40,40) prop.test(successes, trials,conf.level=0.05,correct=TRUE) In the success vector, average total success rate of Group A is 35 and the average total success rate of Group B is 31.5. The test gave a p-value of 0.4555, which is greater than the significance level of 0.05, so there was insufficient evidence to reject the null hypothesis, and so it could be said that the 2 populations did not have statistically significantly different proportion of successes.

It could thus be inferred that there is statistically no significant difference between the average total success rates of participants of the two groups.

It was thus interpreted that the average success rates of the two groups were statistically not significantly different, so the next step was to determine whether the distributions consisting of each participant’s success rate (by group), were statistically significantly different or not. This was done using Mann Whitney U-test.

Null hypothesis: The distributions of success rates of participants of Group A and Group B are statistically not significantly different.

Alternative hypothesis : The distributions of success rates of participants belonging to Group A and Group B, are statistically significantly different. The test was performed using the Wicoxon.test function in R. The first vector contained the success rate of each of the Group A participants and the second vector contained the success rate of each of the Group B participants.

```r
groupA <-c (8,9,9,9)
groupB <-c (8,8,6.5,9)
Wilcoxon.test(groupA,groupB)
```
The test gave a p-value of 0.2059, which was higher than the significance level of 0.05. So there was insufficient evidence to reject the null hypothesis, and thus, it could be said that success rates of GroupA and success rates of GroupB were statistically not significantly different populations.

It could thus be inferred that the success rates achieved by using Tagged-forum and the success rates achieved using forum are independent of the order in which the two plugins are used. Mann Whitney U-test was then performed on the difference between participant success rate with Taggedforum and the Forum (Success rate with taggedforum - Success Rate with Forum).

Null hypothesis: The distribution of differences between the success rates using Taggedforum and the success rates using forum is statistically not significantly different for the two groups

Alternative hypothesis: The distribution of differences between the success rates using Taggedforum and the success rates using forum is statistically significantly different for the two groups.

The test was conducted using R using two vectors. The first vector contained the success rate achieved by forum minus the success rate achieved by taggedforum for each participant of Group A, and the second vector contained the success rate achieved by forum minus the success rate achieved by taggedforum for Group B participant.

groupA <- c(0,1,1,1)
groupB <- c(2,2,1.5,1)

wilcox.test(groupA,groupB)

The test gave a p-value of 0.06317, which was again greater than the significance level of 0.05. So there was insufficient evidence to reject the null
hypothesis, and thus, it could be said that the distribution of differences between the success rates using Taggedforum and the success rates using forum were statistically not significantly different for the two groups.

It could thus be inferred that the differences between the success rates achieved by forum and the success rates achieved by Taggedforum are independent of the order of using the two types of forums.

From the three tests conducted so far, it could be inferred that success rates of Group A and success rates of Group B are statistically not significantly different populations. It could thus be said that the order of using the two plugins (Taggedforum and forum) does not affect the success rate of a participant. The learning effect (the affect of knowing about the posts to be found while using a plugin) is thus balanced for both the groups. The two populations could thus be combined to determine whether the success rates achieved by Taggedforum and the success rates achieved by forum are statistically significantly different or not. Wilcoxon signed rank test was thus the next step in the analysis.

Specifically this test was performed to compare the two related populations (success rates of participants using Taggedforum and the success rates of participants using forum).

Null hypothesis: The two populations Taggedforum and forum are statistically not significantly different.

Alternate hypothesis: The two populations Taggedforum and forum are statistically significantly different

The test was conducted using R using two vectors. The first vector consisted of the success rate of each of the participants using Taggedforum and the second vector consisted of the success rate of each of the participants using forum.

Taggedforum <-c(4,5,5,5,4,5,5)
Forum <- c(4,3,4,3,4,2.5,4,4)

Wilcox.test(taggedforum, forum, paired = TRUE)

The test gave a p-value of 0.01991, which was less than the significance level of 0.05, so there was sufficient evidence to reject the null hypothesis. So it could be said that the two populations, Taggedforum and forum were statistically significantly different in terms of success rates.

It could thus be inferred that the success rates achieved by Taggedforum are different from the success rates achieved by forum. The literature review supports that content-based tagging of forum posts could be beneficial in finding the most relevant forum posts for ones queries, and it could thus be inferred that the success rates achieved by Taggedforum are higher as compared to the success rates achieved by forum.

There is thus insufficient evidence to reject the Hypothesis 1 that the success rates in finding the most relevant forum posts for ones queries achieved by Taggedforum are higher as compared to the success rates achieved by forum. It could thus be concluded that the success rates achieved by Taggedforum are higher as compared to the forum, and the success rate of a participant in finding the most relevant information from the forum is independent of the order of using the two types of forums. This conclusion is based on our sample data which is small in size, but the statistical analysis has sufficient evidence to indicate the trend that makes us not to reject our Hypothesis 1.

Hypothesis 2: The time taken by a participant in finding the most relevant forum posts for ones queries using taggedforum is less than the time taken by the participant using forum.
Wilcoxon rank sum test was performed between the two independent populations - total time taken by each Group A participant and the total time taken by each Group B participant in performing all five tasks.

The test was conducted using R using the two vectors. The first vector contained the total time taken by each Group A participant and the second vector consisted of the total time taken by each Group B participant.

\[
\text{TotTimeA} \leftarrow \text{c}(203.24, 329.9, 388.2, 522)
\]

\[
\text{TotTimeB} \leftarrow \text{c}(362.49, 1128.7, 800.7, 611)
\]

\[
\text{Wilcox.test}\left(\text{TotTimeA}, \text{TotTimeB}\right)
\]

The test gave the p-value of 0.1143, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so, it could be said that the total time taken by each of the GroupA participants to perform all tasks and the total time taken by each of the GroupB participants to perform all tasks were statistically not significantly different populations.

It could thus be inferred that the total time taken by a participant to complete all tasks (using forum and taggedforum) is independent of the order of using the plugins.

The second Wilcoxon sum test for Hypothesis 2 was conducted to determine if the difference between total time taken by each participant using forum and the total time taken by each participant using taggedforum for each of the groups was statistically significantly different or not.

Null hypothesis : The difference between total time taken by each participant using forum and the total time taken by each participant using taggedforum is statistically not significantly different for GroupA and GroupB.
Alternate hypothesis: The difference between total time taken by each participant using forum and the total time taken by each participant using tagged-forum is statistically significantly different for GroupA and GroupB.

The test was conducted using R using the two vectors. The first vector contained the time taken by forum minus the time taken by tagged-forum for each Group A participant. The second vector contained the time taken by forum minus the time taken by tagged-forum for each Group B participant.

\[
\text{DtimeA} \leftarrow (-16.08, -109.3, -85.79, 2) \\
\text{DtimeB} \leftarrow (152.17, 762.3, 341.3, 326.2)
\]

The test gave the p-value of 0.02857, which was lower than the significance level of 0.05. There was thus sufficient evidence to reject the null hypothesis, and so it could be said that the difference between total time taken by each participant using forum and the total time taken by each participant using tagged-forum was statistically significantly different for GroupA and GroupB.

It could thus be interpreted that the difference between the total time taken by each participant to complete a set of tasks using forum and the total time taken by each participant to complete a set of tasks using tagged-forum is affected by the order of using the plugins.

Groups A and B were thus statistically significantly different in terms of time taken to complete tasks. The Wilcoxon sum tests were then performed for each individual task to find out where the differences were.

Wilcoxon sum test for task 1: The Wilcoxon sum test was conducted to determine if the population consisting of the total time taken by each Group A participant to complete task 1 was statistically significantly different population as compared to the population consisting of the total time taken by each Group B participant to complete task 1.
The test was conducted using R using two vectors. The first vector contained the total time taken by each Group A participant in completing task 1 and the second vector consisted of the total time taken by each Group B participant in completing task 1.

\[
\text{Tottime1A} \leftarrow c(34.58, 117.4, 67.8, 192.7)
\]

\[
\text{Tottime1B} \leftarrow c(4.73, 183.9, 281.9, 108.1)
\]

\[
\text{Wilcox.test(Tottime1A, Tottime1B)}
\]

The test gave the p-value of 0.8857, which was higher than the significance level of 0.05. There was thus insufficient evidence to reject the null hypothesis, and so it could be said that the total time taken by GroupA participants to complete task 1 was not statistically significantly different than the total time taken to complete the same task by Group B participants.

It could thus be interpreted that the time taken to complete task 1 is independent of the order of doing it with two different kinds of forums.

A second Wilcoxon sum test was performed for task 1 to determine if the difference between the time taken to complete task 1 by each participant using forum and the time taken to complete task 1 by each participant using taggedforum was statistically significantly different for each of the groups.

Null hypothesis: The difference between the time taken to complete task 1 by each participant using forum and the time taken to complete task 1 by each participant using taggedforum is statistically not significantly different for GroupA and GroupB.

Alternate hypothesis: The difference between the time taken to complete task 1 by each participant using forum and the time taken to complete task 1 by each participant using taggedforum is statistically significantly different for GroupA and GroupB.
The test was conducted using R using two vectors. The first vector contained the time taken to complete task 1 using forum minus the time taken to complete task 1 using taggedforum for each of the Group A participants. The second vector consisted of the time taken to complete task 1 using forum minus the time taken to complete task 1 using taggedforum for each of the Group B participants.

\[ \text{Dtime1A} < -c(-30.42, 17.8, -11, -34.5) \]

\[ \text{Dtime1B} < -c(-0.07, 134.9, 233.3, 66.5) \]

\[ \text{Wilcox.test(Dtime1A, Dtime1B)} \]

The test gave the p-value of 0.05714, which was higher than the significance level of 0.05. So there was insufficient evidence to reject the null hypothesis, and so it could be said that the difference between the time taken by each participant to complete task 1 using forum and the time taken by each participant to complete the same task using taggedforum was statistically not significantly different for Group A and Group B.

It could thus be interpreted that the difference between the time taken by each participant to complete task 1 using forum and the time taken by each participant to complete task 1 using taggedforum is independent of order of using the plugins.

From the two tests conducted for task 1, it could be inferred that the two groups were statistically not significantly different in terms of times taken to complete task 1. The populations for two groups could thus be combined, so the next step was Wilcoxon signed rank test for task 1.

\[ \text{Wilcoxon signed rank test for task 1: This test was conducted to determine if the times taken by participants to complete task 1 using taggedforum was} \]
Null hypothesis: The times taken by participants to complete task 1 using taggedforum is statistically not significantly different population than the times taken by participants to complete the same task using forum.

Alternate hypothesis: The times taken by participants to complete task 1 using taggedforum is statistically significantly different population than the times taken by participants to complete the same task using forum.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 1 using taggedforum by each participant and the second vector consisted of the the time taken to complete task 1 using forum by each participant.

\[
\text{TimeTg1} \leftarrow \{32.5, 2.4, 4.9, 8.24, 5.4, 24.3, 113.6, 20.8\}
\]

\[
\text{TimeF1} \leftarrow \{2.08, 2.33, 67.6, 159.4, 28.4, 257.6, 79.1, 87.3\}
\]

\[
\text{Wilcoxon.test} (\text{TimeTg1,TimeF1,paired=TRUE})
\]

The test gave the p-value of 0.4609, which was higher than the significance level of 0.05. So there was insufficient evidence to reject the null hypothesis, and so it could be said that the time taken by each participant to complete task 1 using taggedforum and the time taken by each participant to complete the same task using forum were statistically not significantly different.

It could thus be interpreted that the time taken by each participant to complete task 1 using taggedforum and the time taken by each participant to complete task 1 using forum are statistically not significantly different.

Wilcoxon sum test for task 2: The Wicoxon sum test was conducted to determine if the population consisting of the total time taken by each Group A
participant to complete task 2 was statistically significantly different population as compared to the population consisting of the total time taken by each Group B participant to complete task 2.

Null hypothesis The total time taken by GroupA participants to complete task2 (using taggedforum and forum) is statistically not significantly different from the total time taken by GroupB participants to complete task 2 using both kinds of forums.

Alternate hypothesis- The total time taken by GroupA participants to complete task 2 (using taggedforum and forum) is statistically significantly different from the total time taken by GroupB participants to complete task 2 using both kinds of forums.

The test was conducted using R using two vectors. The first vector contained the total time taken by each Group A participant in completing task 2 and the second vector consisted of the total time taken by each Group B participant in completing task 2.

\[
\text{Tottime2A} \leftarrow \text{c}(44.63, 89.7, 77.2, 124.6) \\
\text{Tottime2B} \leftarrow \text{c}(177.8, 383.7, 188.9, 117.6)
\]

\text{Wicox.test(Tottime2A,Tottime2B)}

The test gave the p-value of 0.05714, which was higher than the significance level of 0.05. So there was insufficient evidence to reject the null hypothesis, and so it could be said that the total time taken by GroupA participants to complete task 2 was not statistically significantly different than the total time taken to complete the same task by Group B participants.

It could thus be interpreted that the time taken to complete task 2 is independent of the order of doing it with two different kinds of forums.
A second Wilcoxon sum test was performed for task 2 to determine if the difference between the time taken to complete task 2 by each participant using forum and the time taken to complete task 2 by each participant using taggedforum was statistically significantly different for each of the groups.

Null hypothesis: The difference between the time taken to complete task 2 by each participant using forum and the time taken to complete task 2 by each participant using taggedforum is statistically not significantly different for GroupA and GroupB.

Alternate hypothesis: The difference between the time taken to complete task 2 by each participant using forum and the time taken to complete task 2 by each participant using taggedforum is statistically significantly different for GroupA and GroupB.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 2 using forum minus the time taken to complete task 2 using taggedforum for each of the Group A participants. The second vector consisted of the time taken to complete task 2 using forum minus the time taken to complete task 2 using taggedforum for each of the Group B participants.

\[
D_{time2A} <- c(42.37, -74.9, -64.4, -1)
\]
\[
D_{time2B} <- c(57, 266.7, 76.7, 73.2)
\]

Wilcox.test(Dtime2A,Dtime2B)

The test gave the p-value of 0.02857, which was lower than the significance level of 0.05. So, there was sufficient evidence to reject the null hypothesis, and so it could be said that the difference between the time taken by each participant to complete task 2 using forum and the time taken by each participant
to complete the same task using taggedforum was statistically significantly different for GroupA and GroupB.

It could thus be interpreted that the difference between the time taken by each participant to complete task 2 using forum and the time taken by each participant to complete task 2 using taggedforum is affected by the order of using the plugins.

From the two tests conducted for task 2, it could be inferred that the two groups were statistically significantly different in terms of times taken to complete task 2.

Wilcoxon sum test for task 3: The Wilcoxon sum test was conducted to determine if the population consisting of the total time taken by each Group A participant to complete task 3 was statistically significantly different population as compared to the population consisting of the total time taken by each Group B participant to complete task 3.

Null hypothesis The total time taken by GroupA participants to complete task 3 (using taggedforum and forum) is statistically not significantly different from the total time taken by GroupB participants to complete task 3 using both kinds of forums.

Alternate hypothesis- The total time taken by GroupA participants to complete task 3 (using taggedforum and forum) is statistically significantly different from the total time taken by GroupB participants to complete task 3 using both kinds of forums.

The test was conducted using R using two vectors. The first vector contained the total time taken by each Group A participant in completing task 3 and the second vector consisted of the total time taken by each Group B participant in completing task 3.
The test gave the p-value of 0.6857, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so it could be said that the total time taken by GroupA participants to complete task 3 was not statistically significantly different than the total time taken to complete the same task by Group B participants.

It could thus be interpreted that the time taken to complete task 3 is independent of the order of doing it with two different kinds of forums.

A second Wilcoxon sum test was performed for task 3 to determine if the difference between the time taken to complete task 3 by each participant using forum and the time taken to complete task 3 by each participant using taggedforum was statistically significantly different for each of the groups.

Null hypothesis: The difference between the time taken to complete task 3 by each participant using forum and the time taken to complete task 3 by each participant using taggedforum is statistically not significantly different for GroupA and GroupB.

Alternate hypothesis: The difference between the time taken to complete task 3 by each participant using forum and the time taken to complete task 3 by each participant using taggedforum is statistically significantly different for GroupA and GroupB.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 3 using forum minus the time taken to complete task 3 using taggedforum for each of the Group A participants. The second vector consisted of the time taken to complete task 3 using forum minus
the time taken to complete task 3 using taggedforum for each of the Group B participants.

\[
D_{\text{time3A}} \sim \text{c}(25.97, -15.6, 50, -3.6) \\
D_{\text{time3B}} \sim \text{c}(19.54, 20, -13.9, 189) \\
\text{Wilcoxon.test}(D_{\text{time3A}}, D_{\text{time3B}})
\]

The test gave the p-value of 0.8857, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so it could be said that the difference between the time taken by each participant to complete task 3 using forum and the time taken by each participant to complete the same task using taggedforum was statistically not significantly different for Group A and Group B.

It could thus be interpreted that the difference between the time taken by each participant to complete task 3 using forum and the time taken by each participant to complete task 3 using taggedforum is independent of order of using the plugins.

From the two tests conducted for task 3, it could be inferred that the two groups were statistically not significantly different in terms of times taken to complete task 3. The populations for two groups could thus be combined, so the next step was Wilcoxon signed rank test for task 3.

**Wilcoxon signed rank test for task 3:** This test was conducted to determine if the times taken by participants to complete task 3 using taggedforum was statistically significantly different population than the times taken by participants to complete the same task using forum.

**Null hypothesis:** The times taken by participants to complete task 3 using taggedforum is statistically not significantly different population than the times taken by participants to complete the same task using forum.
Alternate hypothesis: The times taken by participants to complete task 3 using taggedforum is statistically significantly different population than the times taken by participants to complete the same task using forum.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 3 using taggedforum by each participant and the second vector consisted of the time taken to complete task 3 using forum by each participant.

\[
\text{TimeTg3} \leftarrow c(32.5,2.4,49.8,24.5,39.4,24.3,113.6,20.8)
\]

\[
\text{TimeF3} \leftarrow c(2.08,2.33,67.6,159.4,28.4,257.6,79.1,87.3)
\]

\[
\text{Wilcoxon.test(TimeTg3,TimeF3,paired=TRUE)}
\]

The test gave the p-value of 0.1094, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so it could be said that the time taken by each participant to complete task 3 using taggedforum and the time taken by each participant to complete the same task using forum were statistically not significantly different.

It could thus be interpreted that the time taken by each participant to complete task 3 using taggedforum and the time taken by each participant to complete task 3 using forum are statistically not significantly different.

Wilcoxon sum test for task 4: The Wilcoxon sum test was conducted to determine if the population consisting of the total time taken by each Group A participant to complete task 4 was statistically significantly different population as compared to the population consisting of the total time taken by each Group B participant to complete task 4.

Null hypothesis The total time taken by Group A participants to complete task 4 (using taggedforum and forum) is statistically not significantly different
from the total time taken by GroupB participants to complete task 4 using both kinds of forums.

Alternate hypothesis- The total time taken by GroupA participants to complete task 4 (using taggedforum and forum) is statistically significantly different from the total time taken by GroupB participants to complete task 4 using both kinds of forums.

The test was conducted using R using two vectors. The first vector contained the total time taken by each Group A participant in completing task 4 and the second vector consisted of the total time taken by each Group B participant in completing task 4.

\[ \text{Tottime4A} \leftarrow \text{c(75,34.2,71.8,88.2)} \]

\[ \text{Tottime4B} \leftarrow \text{c(67.6,115.2,112.9,89)} \]

\[ \text{wilcox.test(Tottime4A,Tottime4B)} \]

The test gave the p-value of 0.2, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so it could be said that the total time taken by GroupA participants to complete task 4 was not statistically significantly different than the total time taken to complete the same task by Group B participants.

It could thus be interpreted that the time taken to complete task 4 is independent of the order of doing it with two different kinds of forums.

A second Wilcoxon sum test was performed for task 4 to determine if the difference between the time taken to complete task 4 by each participant using forum and the time taken to complete task 4 by each participant using taggedforum was statistically significantly different for each of the groups.
Null hypothesis: The difference between the time taken to complete task 4 by each participant using forum and the time taken to complete task 4 by each participant using taggedforum is statistically not significantly different for GroupA and GroupB.

Alternate hypothesis: The difference between the time taken to complete task 4 by each participant using forum and the time taken to complete task 4 by each participant using taggedforum is statistically significantly different for GroupA and GroupB.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 4 using forum minus the time taken to complete task 4 using taggedforum for each of the Group A participants. The second vector consisted of the time taken to complete task 4 using forum minus the time taken to complete task 4 using taggedforum for each of the Group B participants.

Dtime4A <-c(-33,-18,-57,44.8)
Dtime4B <-c(-12.6,65.4,-1.5,-47.8)
Wilcoxon.test(Dtime4A,Dtime4B)

The test gave the p-value of 0.4857, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so it could be said that the difference between the time taken by each participant to complete task 4 using forum and the time taken by each participant to complete the same task using taggedforum was statistically not significantly different for GroupA and GroupB.

It could thus be interpreted that the difference between the time taken by each participant to complete task 4 using forum and the time taken by each
participant to complete task 4 using taggedforum is independent of order of using the plugins.

From the two tests conducted for task 4, it could be inferred that the two groups were statistically not significantly different in terms of times taken to complete task 4. The populations for two groups could thus be combined, so the next step was Wicoxon signed rank test for task 4.

Wilcoxon signed rank test for task 4: This test was conducted to determine if the times taken by participants to complete task 4 using taggedforum was statistically significantly different population than the times taken by participants to complete the same task using forum.

Null hypothesis: The times taken by participants to complete task 4 using taggedforum is statistically not significantly different population than the times taken by participants to complete the same task using forum.

Alternate hypothesis: The times taken by participants to complete task 4 using taggedforum is statistically significantly different population than the times taken by participants to complete the same task using forum.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 4 using taggedforum by each participant and the second vector consisted of the the time taken to complete task 4 using forum by each participant.

\[
\text{TimeTg4} <- c(54,40.1,26.1,24.9,64.4,57.2,21.7,68.4) \\
\text{TimeF4} <- c(21,27.5,8.1,90.3,7.4,55.7,66.5,20.6) \\
\text{Wilcox.test(TimeTg4,TimeF4,paired=TRUE)}
\]

The test gave the p-value of 0.5469, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis,
and so it could be said that the time taken by each participant to complete task 4 using taggedforum and the time taken by each participant to complete the same task using forum were statistically not significantly different.

It could thus be interpreted that the time taken by each participant to complete task 4 using taggedforum and the time taken by each participant to complete task 4 using forum are statistically not significantly different.

Wilcoxon sum test for task 5: The Wilcoxon sum test was conducted to determine if the population consisting of the total time taken by each Group A participant to complete task 5 was statistically significantly different population as compared to the population consisting of the total time taken by each Group B participant to complete task 5.

Null hypothesis The total time taken by GroupA participants to complete task 5 (using taggedforum and forum) is statistically not significantly different from the total time taken by GroupB participants to complete task 5 using both kinds of forums.

Alternate hypothesis- The total time taken by GroupA participants to complete task 5 (using taggedforum and forum) is statistically significantly different from the total time taken by GroupB participants to complete task 5 using both kinds of forums.

The test was conducted using R using two vectors. The first vector contained the total time taken by each Group A participant in completing task 5 and the second vector consisted of the total time taken by each Group B participant in completing task 5.

\[
\text{Totttime5A} \leftarrow c(21,53,84.2,80.3) \\
\text{Totttime5B} \leftarrow c(90.5,375.1,161.9,107.3)
\]

\[
\text{Wilcox.test(Totttime5A,Totttime5B)}
\]
The test gave the p-value of 0.02857, which was lower than the significance level of 0.05. So, there was sufficient evidence to reject the null hypothesis, and so it could be said that the total time taken by GroupA participants to complete task 5 was statistically significantly different than the total time taken to complete the same task by Group B participants.

It could thus be interpreted that the time taken to complete task 5 is affected by the order of doing it with two different kinds of forums.

A second Wilcoxon sum test was performed for task 5 to determine if the difference between the time taken to complete task 4 by each participant using forum and the time taken to complete task 5 by each participant using taggedforum was statistically significantly different for each of the groups.

Null hypothesis: The difference between the time taken to complete task 5 by each participant using forum and the time taken to complete task 5 by each participant using taggedforum is statistically not significantly different for GroupA and GroupB.

Alternate hypothesis: The difference between the time taken to complete task 5 by each participant using forum and the time taken to complete task 5 by each participant using taggedforum is statistically significantly different for GroupA and GroupB.

The test was conducted using R using two vectors. The first vector contained the time taken to complete task 5 using forum minus the time taken to complete task 5 using taggedforum for each of the Group A participants. The second vector consisted of the time taken to complete task 5 using forum minus the time taken to complete task 5 using taggedforum for each of the Group B participants.

\[ D_{\text{time5A}} = c(-21,-18.6,-2.6,73.5) \]
Dtime5B <-c(88.3,275.3,46.7,45.3)

Wilcox.test(Dtime5A,Dtime5B)

The test gave the p-value of 0.1143, which was higher than the significance level of 0.05. So, there was insufficient evidence to reject the null hypothesis, and so it could be said that the difference between the time taken by each participant to complete task 5 using forum and the time taken by each participant to complete the same task using taggedforum was statistically not significantly different for GroupA and GroupB.

It could thus be interpreted that the difference between the time taken by each participant to complete task 5 using forum and the time taken by each participant to complete task 5 using taggedforum is independent of order of using the plugins.
Appendix G

Domain Model

<?xml version="1.0"?>

<!DOCTYPE rdf:RDF [
  <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
  <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]

<rdf:RDF xmlns="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#"
  xml:base="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Static_Data_Structures">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Data_Structures"/>
</owl:Class>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Dynamic_Data_Structures">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Data_Structures"/>
</owl:Class>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Static_Data_Structures"/>
</owl:Class>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Stacks">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Static_Data_Structures"/>
</owl:Class>

104
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Linked_Lists">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Dynamic_Data_Structures"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#2D_Arrays -->

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#2D_Arrays">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7/
domainModel#Arrays ➔

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays"/>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays_of_Characters ➔

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays_of_Characters">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Assigning_Values ➔

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Assigning_Values">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Assigning_Values"/>
</owl:Class>
ontologies/2013/7/domainModel#Variables" />

</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7/
domainModel#Else_Statements -->

<owl:Class rdf:about="http://
www.semanticweb.org/judi/ontologies/2013/7/
domainModel#Else_Statements">
 <rdfs:subClassOf rdf:resource="http://
www.semanticweb.org/judi
/ontologies/2013/7/domainModel#Flow_Control"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7
/domainModel#Flow_Control -->

<owl:Class rdf:about="http://
www.semanticweb.org/judi/ontologies/2013/7
/domainModel#Flow_Control"/>
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Loops"/>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Program_Comilation"/>

109
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Programming_Style"/>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Running_a_C_program"/>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Strings"/>
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Strings">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Arrays_of_Characters"/>
</owl:Class>

<!---- http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Switch_Statements --->

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Switch_Statements">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Else_Statements"/>
</owl:Class>

<!---- http://www.semanticweb.org/judi/ontologies/2013/7 --->

111
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Using_Functions">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Functions"/>
</owl:Class>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Recursion">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Using_Functions"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Variable_Types -->

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Using_Functions -->
www.semanticweb.org/judi
/ontologies/2013/7/
domainModel#Variable_Types">
  <rdfs:subClassOf
    rdf:resource="http://www.semanticweb.org/judi
/ontologies/2013/7/domainModel#Variables"/>
</owl:Class>

<!-- http://www.semanticweb.org/
judi/ontologies/2013/7
/domainModel#Variable_Types -->

<owl:Class rdf:about="http:
  //www.semanticweb.org/judi/ontologies/2013/7
/domainModel#Integer">
  <rdfs:subClassOf rdf:resource="http:
  //www.semanticweb.org
/judi/ontologies/2013/7/
  domainModel#Variable_Types"/>
</owl:Class>

<!-- http://www.semanticweb.org/
judi/ontologies/2013/7
/domainModel#Variable_Types -->

<owl:Class rdf:about="http:
  //www.semanticweb.org/judi/
ontologies/2013/7/domainModel#Character">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi
  /ontologies/2013/7/domainModel#Variable_Types"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi/ontologies/2013/7
/domainModel#Variable_Types -->

<owl:Class rdf:about="http:
  //www.semanticweb.org/judi/ontologies/2013/7
  /domainModel#Float">
  <rdfs:subClassOf rdf:resource="http://www.semanticweb.org/judi
  /ontologies/2013/7/domainModel#Variable_Types"/>
</owl:Class>

<!-- http://www.semanticweb.org/judi
/ontologies/2013/7
/domainModel#Variable_Types -->

<owl:Class rdf:about="http:
//www.semanticweb.org/judi
/ontologies/2013/7/
domainModel#Double_float">
  <rdfs:subClassOf rdf:resource="http:
<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Variables"/>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#While_Loops"/>

<owl:Class rdf:about="http://www.semanticweb.org/judi/ontologies/2013/7/domainModel#Loops"/>

<!-- 115 -->
Note: Each class and sub-class in the domain model corresponds to one possible tag in Taggedforum.
Appendix H

User-Interface of Taggedforum
Adding the Domain Model using File Picker Element

Selection of tags while creating a post
Selection of tags while editing a post (Tags already associated with the post are highlighted)

Selection of tags while replying to a post (Tags associated with the parent post are highlighted)
Display of posts flat ordered by tags

Static_Data_Structures
Re: 16- Pattern matching problem
by Simarpreet Gill - Wednesday, 18 June 2014, 5:47 PM
7

Static_Data_Structures
Data Structures
by Simarpreet Gill - Thursday, 17 July 2014, 9:39 AM
What are data structures? What are their types?

Data_Structures
by Simarpreet Gill - Thursday, 17 July 2014, 9:39 AM
What are data structures? What are their types?

Dynamic_Data_Structures
Data Structures
by Simarpreet Gill - Thursday, 17 July 2014, 9:39 AM
What are data structures? What are their types?
Display of posts in nested form ordered by tags

Display of posts selected by particular tags
- Data_Structures
  - Static_Data_Structures
    - Arrays
      - 1D_Arrays
      - 2D_Arrays
      - Arrays_of_Characters
    - Strings
  - Stacks
  - Dynamic_Data_Structures
    - Linked_Lists

- Variables
  - Assigning_Values
  - Variable_TYPES
    - Integer
    - Character
    - Float
      - Double_float

- Flow_Control
  - Else_Statements
    - Switch_Statements
  - If_Statements

- Loops
  - For_Loops
  - While_Loops
On clicking the Submit button, the posts with the checked tag are displayed. Single or multiple tags can be checked at a time.
<table>
<thead>
<tr>
<th>Flow_Control</th>
<th>1-C Fundamentals</th>
</tr>
</thead>
<tbody>
<tr>
<td>by Simarpreet Gill - Monday, 16 June 2014, 12:58 PM</td>
<td></td>
</tr>
<tr>
<td>What is an if block?</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flow_Control</th>
<th>Re: 2-C Fundamentals</th>
</tr>
</thead>
<tbody>
<tr>
<td>by Simarpreet Gill - Tuesday, 17 June 2014, 10:08 AM</td>
<td></td>
</tr>
<tr>
<td>The group of statements following an if condition</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flow_Control</th>
<th>Re: 3-C Fundamentals</th>
</tr>
</thead>
<tbody>
<tr>
<td>by Simarpreet Gill - Tuesday, 17 June 2014, 10:08 AM</td>
<td></td>
</tr>
<tr>
<td>What is the default scope of if statement?</td>
<td></td>
</tr>
</tbody>
</table>
Bibliography


125


