March 26, 2010

I would like to endorse the thesis proposal of Briane Paul Samson, entitled “Automatic Extraction of Conceptual Relations from Children’s Stories”. Listed below are the dates of consultation between Briane and me on the said thesis proposal:

|  |  |
| --- | --- |
| Date | Description |
| Jan 7 | Upon receiving an email request from Briane, I sent him a document describing a proposed research project on automatic relation extraction |
| Jan 26 | Briane consulted with me to clarify certain issues in the proposed topic and to determine items to be included in the RRL |
| Feb 5 | Briane attended the CELT-Story Generation Workshop to understand the complete framework of the automatic story generation (SG) project and to present his proposed research to the rest of the members of the SG team |
| Feb 15 | First draft of RRL was emailed to me by Briane for review. I instructed him to “read more papers as the draft RRL did not consider new researches in *Machine Reading”*, and “to review the following additional items:  1) other relation extraction works beyond partof and isa  2) other domains where RE has been applied  3) approach of openmind and knowitall |
| Mar 14 | Second draft of RRL was emailed to me by Briane for review, and gave the following comments, which were reiterated in the Mar 16 meeting:   * 2.2 of your RRL should focus on various types of relations. In the SemEval 2010 call for participation, if you follow the hyperlinks on the relations, they provide more details on what the relations are. For RRL, provide only a summary, and the detailed discussion can be placed later in Theoretical Framework. * 2.3 of your RRL should focus on knowledge representations, such as WordNet, ConceptNet, VerbNet, FrameNet, Cyc, etc., which are repositories of semantic relations. * A summary table for your RRL will focus on a list of relations presented in 2.2 that are within the scope of the systems presented in 2.1, as well as the domain. * A research objective will then be to identify the relations to be included in your study (suitable for your domain). |
| Mar 16 | Instructed Briane to add the following sections: 2.2 semantic relations; and 2.3 knowledge representations (ConceptNet, WordNet, Cyc, FrameNet) |
| Mar 20 | I received a copy of the complete thesis proposal for review, and gave the following comments:   * So some items in the URCO's Introduction do not map to our Overview of Current State, specifically the last few paragraphs prior to the "statement of the problem" paragraph discussing the significance of a knowledge base. These can be moved to the Significance of your Research portion instead. * Your RRL is very well-written, so nothing to change there. * For Research Methodology, 3.1 why Filipino linguist? Isn't Picture Books part of what you have to study (to satisfy this sentence in your S&L ==> "reviewing the design of the Picture Books ontology, which is patterned after the design of ConceptNet, to validate the presence of the appropriate relations against those identified from the collected corpus"). Second paragraph of 3.1 should become 3.2 Requirements Specification. In 3.4, again there was a mention of a Filipino linguist. Your language is English, not Filipino. * Architectural Design phase should involve identifying the modules of your proposed research, including external tools and resources that may be utilized. This also includes not only analyzing but also designing the data structures to be used to represent the semantic relations for story generation. * For future revision -- theoretical framework should be split to different sections; e.g., Storytelling Knowledge (e.g., Picture Books); Semantic Ontology (e.g., ConceptNet); Rhetorical Structure Theory, etc. * For the mock presentation for XMETHODS, you have to include sample sentences and possible relations that you can extract, preferably those related to stories and Picture Books. * While Jeff Cua's doing his research on mapping Pic BOoks knowledge to SUMO, we were able to identify more limitations of the current ontology of Pic Books which the new batch (Pic Books 2) is trying to address. So it is possible that you have to work with the relations identified by Pic Books 2 (which covers Pic Books 1 + extensions to fix the identified problems) rather than Pic Books 1; otherwise, we won't be able to test some of your extracted relations). For now, you can leave your existing proposal as is (reference to Picture Books), but you may need to add one more section in RRL -- representing storytelling (or narrative) knowledge - and focus on works of Picture Books 1 and 2. I'll get the latter from the new batch -- they are targetting to finish their ontology design by end of this week. |
| Mar 26 | Received a revised copy of the proposal with the requested revisions in place. |

Reviewed by:

Ethel ChuaJoy Ong (sgd)

AAC – Center for Human Language Technologies

Software Technology Department

College of Computer Studies

De La Salle University – Manila