**Review of Related Studies**

1. Named Entity Recognition using Hidden Markov model

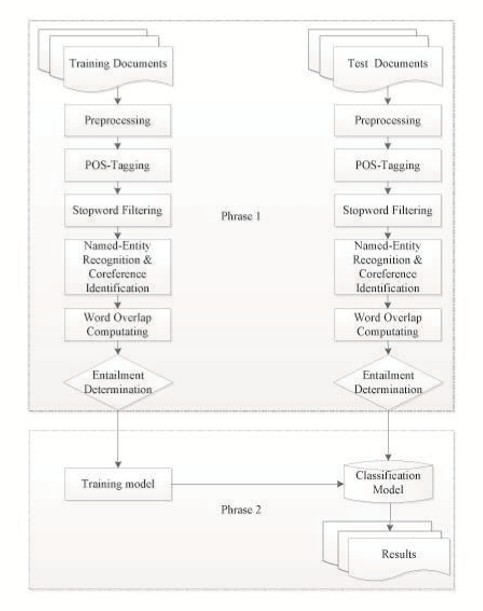
The task of Named Entity Recognition is to classify the given text into some predefined categories, like person, place, date, time, etc. NER serves an important role across the natural language processing, it can be used in information extraction, machine translation and text synthesis. There are few ways to make an NER system, there is a rule-based, which uses handwritten rules by linguist or an expert in the language, and these rules are language dependent which means that the rules are written in that specific language and may or may not be used in other languages for which every language has different rules in writing. Rule-based NER can perform efficiently if the rules are well constructed but this rule-based approach is not versatile in changes and is language dependent. In today’s computing world, machine learning has become a common approach in NER systems. This is because of the advantages of the machine-learning approach. Machine Learning approach does not need rules to be predefined, it needs a *corpora* and the ML system will learn the rules on its own. Also, training in the ML system is easy to implement, with this approach, one ML system can be used in different domains of study and different languages as well because of the requirements of making this approach, a corpora. There are many approaches that can be used in making a machine learning approach in NER, there is Conditional Random Fields (CRF), Maximum Entropy Markov Model (MEMM), Support Vector Machine (SVM) and Hidden Markov Model(HMM). Among these approaches, HMM is the most promising but has not fully developed in NER systems. Hidden Markov Model assigns the joint probability to paired observation and label sequence.Out of all the approached in doing an NER system, HMM have higher evaluation performance than others. HMM based NER is language dependent which it is can be used in different language domains, highly understandable and analyzable, it is dynamic in nature, and importantly it does not require language experts to make the training sets. The proposed system uses learning by example methodology and provided steps in making an HMM based NER system. The first step is data preparation, in this stage, the conversion of raw data into trainable form happens. This needs to be address first so that we can use the transformed training data to feed to the HMM system. This stage requires the own experience in the language, in other words, hand-in-hand or manually annotating the text. Second step would be estimation of parameters for the HMM, these parameters are the state vectors, which contains all the named entity tags candidate interested, the start probability, which indicates the probability that the sentence starts with a certain tag, transition probability indicates the probability of transitioning from one state to another, while the emission probability pertains to probability of assigning particular tag to the word in the corpus or in the document. The final step would be the testing which would be done by the observation or from own experience. In conclusion, they successfully made an NER system based on HMM approach that is dynamic in nature, usefulness to other NLP fields and other language domains. The system acquires 90% accuracy in tagging words in the document, it can be achieved through the use of properly annotated corpora.

1. An Experiment in Semantic Tagging using Hidden Markov Model Tagging

Understanding the meaning of a word has been a challenge to psychology and artificial intelligence research communities. The use of Hidden Markov Model in part of-speech tagging has been proven to be a useful approach in eliminating language ambiguity. The idea of semantic tagging is to add word sense markings so that it can be used to process in some automatic processor of languages to choose the proper meaning of words in a given context. Researchers can agree that semantic tagging is much more difficult than performing part-of-speech tagging. But because there are approaches that are proven to be effective in part-of-speech, the researchers experimented how these approaches behave when used in semantic tagging. They have used the WordNet’s 45 semantic tags, these got some advantages, first the size of the semantic tag is scalable, meaning the training data does not have to be large, secondly the semantically tagged corpora has been provided also. The part-of-speech tagging is easily understood than semantic tagging, it is because semantic tagging has no consensus or general representation or agreement while the POS tagging has general understanding and ways of interpreting. But despite the lack of consensus, the WordNet team takes the challenge to construct general semantic tagging scheme. These 45 semantic tags consist of 3 tags for adjectives, 1 for all adverbs, 26 tags for nouns and 15 tags for verbs. It is a fact that HMM approach in POS tagging is very promising, so the researchers try the HMM approach on semantic tagging if it will result to higher accuracy in tagging meanings to the context. If using the Hidden Markov Model, there are preparation activity. This preparation includes making a *lexicon* which contains all possible semantic tags. Also, we need to prepare training corpus and a test corpus with the same size. After this compute the HMM model based on training sets by comparing the tags by semantic tagger and in the test corpus. They include three major test, the first test’s test corpus the overall accuracy was 86% and the accuracy over ambiguous tokens of 71%. The second test, they experimented with the use of POS pre-tagging, by using this, the number of ambiguity classes decreases, with this the overall accuracy of 89% but the accuracy over ambiguous tokens is identical at 71%. In the third test, they assigned the most common part of speech for each lexicon which got the overall accuracy of 90%. Researchers find it surprising that the processes and techniques to improve the accuracy of the POS tagger when applied to semantic tagger, the overall accuracy also improves.

1. Applying CRFs and SVM to Textual Entailment Recognizing

Textual entailment is a useful for systems that includes inferencing over sentences in natural language. Entailment knowledge is important in any knowledge-based systems. The Recognizing Textual Entailment(RTE) in a simple sense is if you are given two text or phrase, the system should recognize if the either one of the text can be entailed, i.e. deriving or inferring one text from the other text. There are few challenges in developing a systems that focused in RTE. One problem is the incompleteness of sentences, incorrect spelling, grammar error, and abbreviations, which will make a problem soon because these errors and abbreviations could mean something or anything else making it hard for the system to recognize entailments. Second if some parts of the text are written in the context but the references or sources of information is defined explicitly or implicitly in reference. dates, places , pertaining to the corpus. RTE defined as the direct relationship or connection between two given text. Support Vector Machine aims to seek or make decision boundary to separate the given training data into two classes and their corresponding classification. It is actually based in the structural risk minimization principle of computational machine learning theory. The input of the SVM is a training examples and the machine will try to find the classification of every training data and maps them to classes. The structural risk minimization aims to find the hypothesis that we can rest assure that the lowest probability bound for generalization error. While the Conditional Random Fields (CRF), it gets the strong independence assumptions of HMM and solving the problem of label-bias problem of the MEMM. To model that real world data that the conditional probability of a sequence can depend on features of the observation sequence.



There are preprocesses that are involve before performing the textual entailment recognizing. This includes converting uppercase words and letters to lower cases, removing stop words (the, to, etc.) because they have little lexical content, they also remove words with length 1, there are also named entity recognition which will give identification to words. After the preprocessing procedure, the classification starts. The use of machine learning methods in recognizing entailments between a text and a hypothesis is effective however the dataset’s nature will help to achieve higher accuracy as a feature in testing.

1. Representation of Knowledge in a Program for Solving Physics Problem

Any word problem can be solved efficiently through the use of modelling the problem itself. The need for a machine that solves problems automatically are rising in today’s world, numerous than ever. A computer program that solves physics word problems that is stated in English has been made in this regard. The sentences of the physics word problem are transformed into semantic network form to generate objects that are from canonical object frames. The use of specialized representation using the procedural knowledge that is required to convert one representation into another because of some information is unspecified implicitly by the source representation, with this, it simplifies many of the processes which must be performed by the program. The program is called ISAAC, it is able to read, understand, solve, and generate objects based on the problem stated. The program first needs to understand the English sentence of the problem. This needs the parsing of various semantic parser network form, which helped by large number of semantic programs. One of the important semantic process is referent identification, this identify the relationship between objects based on developing model of the problem. The next major step is to identify the type of conceptual entity, such as the location of the object, its attributes, etc. The final step in the semantic process is the execution of the verb semantics, the arguments of the verb are represented as semantic frames, this causes the transfer of new information to existing objects in the internal model of the problem. After all the semantic processing, the next step is to process by the language-free internal model, this model makes the representation of the objects such as ladders, tables, ropes, the person, also its features, and their relationship. Then this objects are transformed into Canonical Object Frame, which is as idealization or abstraction of certain features of an actual object that is the behavior of real objects. This representation are used as a reference to make a geometric model that is a model for representing he positions of the objects in the problem to a common coordinate system. Combining geometric models of course needs the features or attachment to each other of the objects. The program solves the problem in the formal system and returns the result back in its original form. After the canonical objects have been generated from geometric models, equations are written based accordingly on physical laws. The last step is the generation of the picture and the diagram of the problem, this includes reasonable sizes in the drawing object. The program proven that the notion of canonical object frame is a powerful technique for constructing problem systems [?]. Also the use of canonical object frame makes the problem search space smaller because the irrelevant information are not being accessed.

1. Reading Coaching for Math Word Problems

Young student nowadays have challenges in comprehending math word problems, even though they know what mathematical operations to perform. The reason for this problem is they can’t comprehend what the problem is actually asking them to do, because young readers is distracted by everyday language, math words, or combination of both. Those who struggles in reading comprehension and the math computation faces the biggest challenge. The researchers make use of the fourth grade math questions because that grade is a crucial part in student’s life in school, learning in this stage establishes the foundation of learning into upcoming school years.

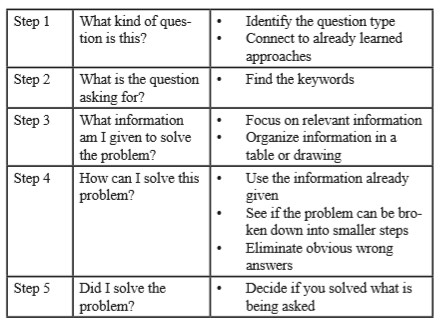
It is challenging for students who are solving math word problems when they are facing or reading words that are unknown or unfamiliar to them.

*“Haley swam 22 laps each day for 18 days. Then she swam 25 laps each day for 10 days. What was the total number of laps she swam over the 28 days?” (Massachusetts Department of Education, 2006)*

Some students may not know the word ‘swam laps’ which makes them have difficulties in solving this problem. The strategy is to ignore unfamiliar or confusing words and try solving the problem with the words the child knows [?]. In above example question, a young individual can ignore the words *swam* and *laps.* The readers need to recognize that *Haley* did something 22 times for 18 days, and 25 times for 10 days then adding the product of these two then you have the final answer.

Word problems include proper names that may be unfamiliar for young readers. Proper names like people names may distract readers to the essential mathematical information in the problem. When young readers are solving word problems, if they encounter an unknown word, the easily classify this as a name. That’s why when readers get confused with these names, teachers encourage them to substitute them with the name of their parents, friends, brother, etc. Then it is easier now to comprehend the given word problem.

Another young reader faces is the sentence structure and syntax, because word problems are written in compositional and not in conversational English. Mathematician George Polya’s (1973) classic problem solving approach use of a model that teachers can teach the students how to properly comprehend word problems.



Polya’s framework need the readers first to determine the type of problem, then the next thing they need to look is what is the problem trying to find. Then look for the given relevant variables in the problem, from this reflect how do you solve the problem with the given data. Then finally ask if you have satisfied the problem is trying you to look for.

Math terminologies makes it difficult for young readers to solve the word problems because of they need to recall mathematical terms and how to solve them. For example, *total, odd and even, greatest and least, etc.* First strategy here is constantly teach the student how to solve these mathematical terms. Second strategy is to encourage students to formulate their own problems to familiarize them in math terms.

Another challenge is problems that have multiple math operations, which hinders students to understand how to formulate solution to the given problem, the strategy here is that change word variables with another names that a student can relate too (e.g. basketball points, scores). Charts, graphs and visual about the question also makes the problem for the students uncomprehensive, the solution is to make charts and graphs about different topics so that students can ask their friends and family.

Math word problems is not just about mathematical computation, it needs the use of literacy and proper comprehensive. Literacy coaches and teachers need wide-ranging strategies in order to support children as they improve their skills in reading and mathematics [?].