IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 20, Issue 3, Ver. II (May. - June. 2018), PP 01-05
www.iosrjournals.org

Generating UML Diagram Using Natural Language Processing And Use Case Diagram

Sandip kumar Shrivastav[1],Bramah Hazela[2], SheenuRizvi[3]
Department of Computer Science & Engineering,Amity University, Uttar Pradesh, India
Corresponding Author:Sandip kr. Shrivastav

Abstract : *This research paper presents a natural language processing based automated system for generating UML diagrams after analyzing the given Use case diagram in the current scenario. This model represent to analyzing natural language processing and extract the relative information from the given UML diagram which are given by the user . User defines the all requirement in simple input language of machine language which are context free language and grammar language. After combining all the information we prepare a relative model in the grammar language. Natural language processing convert this language in machine language for understood to the machine. After this information we can generate the various UML diagram as activity diagram, sequence diagram, use case diagram, class diagram and other many case tools. These outline and customary CASE tool devices require a great deal of additional time and endeavors from the framework examiner amid the way toward making, masterminding, naming and completing the UML graphs. This is solid approach to produce UML charts to spare the time and spending plan of both the client and framework examiner.*

Keywords *-Natural language processing, UML diagram, Use case diagram, Data flow diagram, Software Requirement Specification.*

Date of Submission: 18-05-2018 Date of acceptance: 02-06-2018

I. Introduction

In the modern generation the looks and style of programming designing has been completely changed. Presently every progression of programming designing takes after the guidelines of Object Oriented configuration designs. As of late, there is no product which gives administrations to draw UML outlines more productively aside from Rational Rose, Smart Draw and so forth and there is almost certainly that these are sensibly great programming.

Customarily, the framework expert needs to complete a great deal of work for deriving model rationale and comprehension the client necessities for UML graphs by utilizing universal CASE instruments. Consequently, there is large of wastage time because of the dull idea of the accessible CASE devices for the required situation. In this day and age everyone needs a fast and solid administration. So it was required that there ought to be a type of canny programming for producing UML based documentation to spare time and spending plan of both the client and framework expert. With a specific end goal to determine every single such issue, we require programming, which encourages the two clients and programming engineers. As far as this product is concerns the time, it takes to investigate all the offices and administrations, ought to be very not exactly a moment and this data is very helpful for the clients.

The point of this paper is to show the utilization of NLP and area cosmology procedures for the extraction of UML graphs from casual common dialect necessities by executing a model device that uses the specified systems.

II. Summary Of problem

The issue particularly tended to in this exploration is basically identified with the product examination and outline period of the product advancement process. Scarcely any years back information stream graph's (DFD) were being utilized to symbolize the stream of information and speak to the client's necessities. Yet, in current age, brought together displaying dialect is utilized to model and guide the client prerequisites,which is more far reaching e and real approach to of portrayal and it is valuable for the later phases of programming improvement.

The product in the present market which give this office are simply paint like apparatuses as Visual UML, GD Pro, Shrewd Draw, Rational Rose and so forth. Every one of them have dull nature. To utilize the broadly over-burden interface of these CASE instruments is a vexing issue. The procedure of creating the UML graphs through these product designing devices is extremely troublesome and this is a tad more repetitive and

long procedure to perform. Consequently, it was required that any distinctive individual included mandatory in programming improvement may get his required yield with most extreme exactness in least time devoured.

III. Related Solution

The usefulness of the directed research was space particular yet it can be upgraded effectively later on as per the prerequisites. Current composed framework consolidate the capacity of mapping client prerequisites in the wake of perusing the given prerequisites in plain content and represent in UML and class diagram, Action Diagram, Sequence Diagram, Use case chart what's more, Component Diagram.
An effective Environment would likewise be accommodated User Interaction and effective Input and yield.

IV. Natural Language Processing

In the NLP we give any input any language NLP are convert that input which language are machine understand and machine understand the given input and give the feedback their own machine language and NLP change that the output language in our understanding language.

The comprehension and multi-viewpoint preparing of the normal dialects that are additionally named as "discourse dialects", is really one of the contentions of more noteworthy enthusiasm for the field manmade brainpower field. The characteristic dialects are unpredictable and lopsided. Customarily, characteristic dialects depend on un-formal punctuations.

There are the topographical, mental and sociological elements which impact the practices of common dialects. There are vague arrangement of words what's more, they likewise change and differ region to zone and time to time. Due to these varieties and irregularities, the characteristic dialects have diverse flavors as English dialect has the greater part dozen prestigious flavors all over the world. These flavors have distinctive accents, set of vocabularies and phonological viewpoints. These foreboding furthermore, threatening errors and irregularities in normal dialects make it a troublesome assignment to process them when contrasted with the formal languages. During the time spent investigating and understanding the regular dialects, different issues are generally looked by the specialists. The issues associated with the more noteworthy many-sided quality of the regular dialect are verb's conjugation, emphasis, lexical abundancy, issue of vagueness, and so forth. From this arrangement of issues the issue which ever causes more challenges is issue of vagueness. Equivocalness could be effortlessly unraveled at the sentence structure and semantic level by utilizing a sound and hearty run based framework.

V. Component of NLP

The two major components of Natural Language Processing are,

1. Natural Language Understanding (NLU)
2. Natural Language Generation (NLG)

I. Natural Language Understanding (NLU)

Here the speech input gets transformed into useful representation in order to analyze the various aspects of language.A natural language can be very ambiguous (different meaning of same sentence)

II. Lexical Ambiguity

The issue of words composed or articulated not effectively is discarded in this issue situation. These sorts of blunders are just feasible through the examination with the articulations contained in a word reference. Lexical uncertainty is made when a same word expect different implications. For this situation that vagueness is produced from the reality that which implications will be joined in which situation. For instance we consider the "ice" modifier in the accompanying sentences:
"That room is ice."
"That person is cold."
It turns out clear that the same "frosty" descriptive word expect, in the two expressions diverse implications. In the initially sentence it show a temperature, in the second one a specific character of a man.

III. Syntactical Ambiguity

Language structure examination is performed on world level to perceive the word classification. The syntactic examination of the projects would need to be in a situation to segregate subject, verbs, objects and different supplements. It is minimal complex method.
"Mario eats the apple."

In this case, the genuine implications are that "Mario eats apple" yet vagueness can be affirmed if this sentence is imagined as "the apple eats Mario".

IV. Semantic Ambiguity

To break down an expression from the semantic perspective intends to give it a significance. This should let you comprehend we touched base to a pivotal point. Semantic ambiguities are most regular because of the way that by and large a PC isn't in a situation to recognize the legitimate circumstances. "The auto hit the post while it was moving." Every one of us would doubtlessly decipher the expression like "The auto, while moving, hit the post.", while no one would be envisioned to credit to the sentence the signified "the auto hit the post while the shaft was moving ".

V. Pragmatic Ambiguity

Sober minded ambiguities conceived when the correspondence occurs between two people who don't have the same setting. As following illustration:
"I will touch base to the air terminal at 12 o'clock."
In this illustration, if the subject individual has a place with a diverse mainland, the implications can be completely changed.

VI. Disclosure Ambiguity

Here the meaning of the sentence is verified with sentence before it.

VI. Natural Language Generation (NLG)

I. Text Planning

It includes the extracting knowledge from knowledge base.

II. Sentence Planning

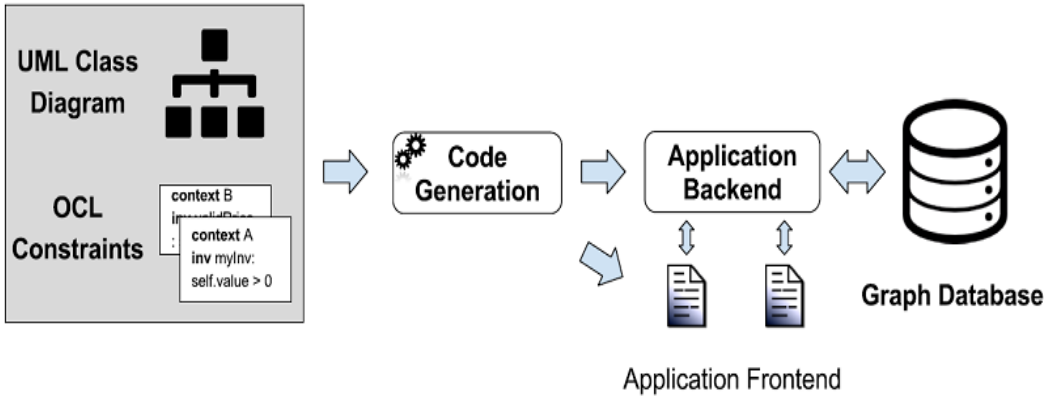
This includes selection of correct words and forming sentence which follow the grammar.

III. Text Realization

Mapping the planned sentence into reality.

VII. Architecture of Designed System

The outlined UMLG framework has capacity to draw UML outlines in the wake of perusing the content situation gave by the client. This framework attracts outlines four modules: Text input obtaining, content comprehension, information extraction. Following figure show the architecture of UML is given below.



I. Text input acquisition

This module gets input content situation. Client gives the business situation in from of sections of the content. These modules peruse the information message in the frame characters and create the words by linking the input characters. This module is the execution of the lexical stage. Dictionaries and tokens are produced in this module.

II. Text Understanding

This module peruses the contribution from module in the form of words. These words are ordered into different classes as verbs, helping verbs, things, pronouns, modifiers, relational words, conjunctions, and so forth.

III. Knowledge extraction

This module, separates distinctive protests and classes and their particular properties on the basses of the information given by the former module. Things are symbolized as classes and questions and their related properties are named as qualities.

IV. UML diagram generation

This is the last module, which at last uses UML images also, draws different UML graphs by consolidating accessible images as per the data separated of the past module. As isolated situation will be given for different charts as classes, succession, action and utilize cases outlines, so the different capacities are actualized for particular.

VIII. Used Methodology

Regular common dialect preparing based frameworks client run based frameworks. Specialists are another approach to address this problem. In the examination, an administer based calculation has been utilized which has vigorous capacity to peruse,

comprehend and separate the coveted data. As a matter of first importance
essential components of the dialect sentence structure are removed as verbs, things, modifiers, and so forth then based on this extricated data additionally preparing is performed. In etymological terms, verbs frequently indicate activities, and thing phrases the articles that take an interest in the activity. Each thing expression's then part indicates how the protest takes an interest in the activity. As in the accompanying case:
"Robbie hit a ball with a racket."

A system that sees such a sentence must find that Role is the specialist since he plays out the activity of hitting, that the ball as the topical protest since it is the protest hit, and that the racket is an instrument since it is the apparatus with which hitting is done.

In this manner, sentence examination requires, partially, the responses to these activities: The quantity of topical parts grasped by different speculations fluctuates likely. A few people use about about six topical parts. Others utilize more circumstances as many. The correct number does not make a difference much, as long as they will sufficiently awesome to uncover normal limitations on how verbs and topical examples shape sentences.

• **Agent:** The specialist makes the activity happen as in "Robbie hit the ball," Robbie is operator who plays out the errand. Be that as it may, in this illustration an inactive sentence, the operator additionally may show up in a prepositional "The ball was hit by
Robbie."

• **Co Agent:**The word with may present a join expression that serves an accomplice in the key operator. They two do the activity together "Robbie played tennis with Susie."

• **Beneficiary:**The recipient is the individual for whom an activity has honey bee performed: "Robbie purchased the balls for Suzie." In this sentence Suzie is recipient.

• **Thematic object:**The topical question is the question the sentence is extremely about—ordinarily the question, experiencing a change. Frequently the topical protest is the same a-s the syntactic direct protest, as "Robbie hit the ball." Here the ball is topical question.

• **Conveyance:**The movement is something in which or on which voyages: 'Robbie dependably passes via prepare."

• **Trajectory:**Motion from source to goal takes put over at direction. ID difference to the next part conceivable outcomes, a few relational words can serve to present direction thing phrases: "Robbie and Suzie went in through the front entryway: he conveyed her over the edge."

• **Location:**The area is the place an activity happens. As in the direction part, "a few relational words circular segment conceivable, which passes on implies notwithstanding filling in as a flag that an area thing phrase is "Robbie and Suzie considered m the library, at a work area, by the divider, a photo, close to the entryway."

• Time: Time indicates when an activity happens. Relational words such at, previously, and after present thing phrases filling in as time part fill "Robbie and Susie left before twelve."

• **Duration:**Duration determines to what extent a move makes. Relational word, for example, for demonstrate term. "Robbie and Susie ran for 60 minutes."

IX. Conclusion& Future Scope

This examination is about the dynamic age of the UML graphs by perusing and breaking down the given situation in English dialect gave by the client. The composed framework can discover the classes and questions and their properties and tasks utilizing a fake knowledge system, for example, regular dialect preparing. At that point the UML graphs, for example, Activity burrow., Sequence burrow., Component burrow., Use Case burrow., and so forth would be drawn. The exactness of the product is normal up to around 80% with the contribution of the product build gave that he has taken after the pre-necessities of the product to set up the information situation. The given situation ought to be finished and written in straightforward and remedy English. Under the extent of our undertaking, programming will play out a total examination of the situation to discover the classes, their traits and activities.

It will also draw the following diagrams.

1. Class Diagrams
2. Activity diagrams
3. Use-Case Diagrams
4. Sequence Diagrams
5. Component Diagrams

The composed framework for producing UML outlines was begun with the points that there ought to be a product which can read the situation given in English dialect what's more, can draw the a wide range of the UML graphs, for example, Class outline, action graph, succession chart, utilize case outline, part graph, organization chart. Be that as it may, last two of them part graph, organization outline are as yet untouched. There is likewise some edge of changes in the calculations for producing initial four composes Class chart, action chart, arrangement outline, utilize case graph. Current precision of creating outlines is around 80% to 85%. It can be upgraded up to 95% by enhancing the calculations and actuating the capacity of learning.

References

[1]. Thomas C. Rindflesch (1999) Natural language processing.

[2]. Ruth Malan and Dana Bredemeyer, (2001) ―Functional Requirements and Use Cases‖, Bredemeyer Consulting, ruth_malan@bredemeyer.com

[3]. L. R. Tang and R. J. Mooney, (2001), ―Using Multiple Clause Constructors in Inductive Logic Programming for Semantic Parsing‖, Proceedings of the 12th European Conference on Machine Learning (ECML- 2001), Freiburg, Germany, pages 466–477.

[4]. Woo, H., & Robinson, W. (2002). ―A light-weight approach to the reuse of use-cases specifications‖, In Proceedings of the 5th annual conference of the Southern Association for Information Systems (pp. 330- 336) Savannah: SAIS

[5]. Imran Sarwar Bajwa, M. Abbas Choudhary, (2006) ―A Rule Based System for Speech Language Context Understanding‖ Journal of Donghua University, (English Edition) 23 (6), pp. 39-42.

[6]. Mohd Ibrahim, Rodina Ahmad "Class diagram extraction from textual requirements using Natural language processing (NLP) techniques",IEEE journal 2010

[7]. Deva Kumar Deeptimahanti, Muhammad Ali Babar "An Automated Tool for Generating UML Models from Natural Language Requirements" IEEE journal 2009

[8]. Farid Meziane, Nikos Athanasakis, Sophia Ananiadou, 2007, Generating Natural Language specifications from UML class diagrams, Springer-Verlag London Limited 2007

[9]. Gobinda G. Chowdhury , 2001, Natural Language Processing.

[10]. L. Mich, NL-OOPs: "From Natural Language to Object Oriented Using the Natural Language Processing System LOLITA.", Natural Language Engineering, 2(2), 1996, 87.

IOSR Journal of Computer Engineering (IOSR-JCE) is UGC approved Journal with Sl. No. 5019, Journal no. 49102.

Sandip kr. Shrivastav "Generating Uml Diagram Using Natural Language Processing and Use Case Diagram." IOSR Journal of Computer Engineering (IOSR-JCE) 20.3 (2018): 01-05.