International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020
ISSN: 1475-7192

Recent Research Trends and Advancements in Computational Linguistics

[1]R. Ganesan, [2]M. Thangamani, [3]S. Jafar Ali Ibrahim

Abstract--*AI and machine learning change everything about computing. Innovations in the processing of natural languages change everything about interaction between human computers. Some of these developments have become possible due to ground—breaking natural language processing work. The scientific study of language from a machine viewpoint is ascribed as computational linguistics (CL). CL's computer work focuses on natural language modelling and is based on computing and artificial intelligence. The goal of machine linguistics is to build software to interpret the natural language that we use to interact in everyday life. Scientists in the field of Informatics have been faced with interactions between people and equipment. The rapid and exponential growth of regimented knowledge such as records, databases and social media has rendered it even more important to predict the future of digital linguistics. In fact, the capacity of the providers to use whatever information available for tactical insight and decision making can be improved significantly by human and machine connectivity. We also addressed global trends and advances in machine linguistics in this research article.*

Key words--*Computational Linguistics; HPSG; Voice Based CL Systems; AI in Computational Linguistics.*

I. INTRODUCTION

The natural language analysis is otherwise regarded as computational linguistics. Computer systems are smartly sensitive to human and textual content. We concentrate on the meaning of words. When you're asking Google, Siri or Alexa, they are extremely good to know the speech, but it doesn't listen and question half the time. We found human languages in our world with their syntactic structures and their intricacy is still fascinating enthralling. We were involved in the language and language analysis itself, but we were also interested in the machine. This side aspect calls for the kind of people interested more in how words are combined to create meaning and feeling and how they come up with metaphors.

II. METHODOLOGY

HPSG as well as its Computer Linguistic Feature

Head—Driven Phrase Structure Grammar (HPSG) is the single most influential paradigm for semantic grammar definition in computational linguistics. HPSG is a combination of different components with different status within a grammar framework (strategies and components). It illustrates that the various elements and status of the theory are not developed and explicitly explained. Linguistic generalizations can be found in explicitly specified yet clearly reflective concepts, rather than in any other grammatical context.

[1]*Professor, Department of English, Kongu Engineering College, Tamilnadu, India, id ganesankec@gmail.com*
[2]*Assistant Professor, Department of CT-UG, Kongu Engineering College, Tamilnadu, India, manithangamani2@gmai.com*
[3]*Research Scholar, Department of CSE, Kongu Engineering College, Tamilnadu, India, jafartheni@gmail.com*

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

In fact, a standardized formal structure is used to define grammar pattern, generic grammar and unique grammars. If one understands HPSG's integrated sleek design, one also comprehends why the framework can be the basis for theoretical and computer-based research, even if certain specific linguistic or universal language theories and analyses cannot be adopted or subsequently modified. HPSG in our view, can be considered as a package for formal language theories and not just grammars that already contain a collection of important theories and analyses. HPSG's modular design offers a great deal of versatility that envelops the following.

- The overall framework is further developed.
- New grammar model variations are tested.
- Different kinds of implementation approaches are experimented.
- Different grammar data types are added.
- The framework is applied to new languages.
- The components of grammar have evolved.

On other hand, the various aspects of liberty make it tough enough to determine the characteristics that are important to HPSG grammar models, HPSG conceptual rigor application or unique HPSG grammars for certain languages. What are the features, formality or semantics of the initial HPSG that are referred to as HPSG? It is apparent that there cannot be a formalizable answer to this question. Many suggestions now include significant modifications to the original framework and analysis.

HPSG resources deployment

There are countless ways in which HPSG-based computer tools are used.

Documentation of vocabulary and checking of linguistic assumptions

Grammar engineering**,** that is, the production of grammar in software is an important procedure to check linguistic theories at a scale, according to Müller (1999), Bender (2008) and Bender et al. (2011). By "on scale," we mean both large amounts of data and incorporated language models which handle multiple phenomena simultaneously. Here are the few descriptions of strategies of phrase hypotheses [1-7].

- CoreGram
- Grammar Matrix
- Aggregation

Indications of the language

In this segment, we illustrate many different ways of creating a language viewpoint that has not previously been established through non-computational methods for grammar engineering work within HPSG [8].

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

Ambiguity

Now that broad reporting precision grammars are available, the extent of ambiguity has become evident. By taking both coverage and precision seriously, it is possible to investigate it on a large scale, quantifying the sources of ambiguity and the information needed to resolve it.

Long—tail phenomena

HPSG's value as a theoretical framework enables both 'central and peripheral' anomalies to be studied within a single cohesive model. In fact, the extent to which a phenomenon should be considered as “core, peripheral” or something else between them can be examined through the implementation of large-scale grammars across a range of languages (Müller, 2014). In addition, it becomes clear how long the "peripheral" anomalies will last when dealing with actual data and large-scale grammars. In fact, the ongoing development of widely distributed linguistic tools allows more and more low-frequency (or comparatively low-frequency) anomalies to be established.

Analysis—order effects

Grammatical engineering means making and building on unique analyses. This has both advantages and disadvantages: On the one hand, it means that the results of earlier work can directly construct new grammar engineering. It also ensures that the work on which it operates is limited to any new grammar engineering work.

Voice based systems

Current examples involve too many voice-based systems. Some examples include Siri, Alexa and voice from Google. We call these conversation agents applications, which understand and accomplish things depending on what we say in our spoken language. It is easy to make a conversation that reproduces random things, which are half-connected to what we asked. But how can we build dialog officers who can understand and do the tasks they ask for? One way of achieving it is to set up this type of contextualisation that we spoke about before. It is based on data accessible with its organized information from outlets like Freebase and Wiki data. Ironically, these systems must be constructed by hand right away. The vision calls for a system that dynamically creates them after reading blogs or online articles.

The artificial intelligence as the basis for CL

The science and engineering of making smart machines, especially a smart computer program, involves developing artificial intelligentsia. So it is useful to know how Bolshakov and Gelbukh [10] claim' . . . neither the pace of computers nor programming tools can increase, nor many toy systems for the 'comprehension' of languages are further built in smaller fields will solve one of the most difficult problems of the evolutionary theory-automatic text understanding. 'The quotation created by these machine linguists shows that linguists and developers work together to achieve a difficult task in modern times. In order to understand human intelligence, i. e. using devices, however, there must be no incorporation of biologically viable methods in human intelligence. Computer intelligence includes strategies.

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

When it makes sense for the mechanics to do a job and obey their actions, we often feel proud of our being a little wise. AI seeks to ensure that computers track and perceive the systems in different fields of work. Computer programs integrated into this AI ensure dramatic results and efficiency.

The math algorithm is the quintessence of AI and it determines the performance of a computer program for the consistency of logical steps in any algorithm. The great logicians of mathematics, Kurt Godel and Alan Turing, proved in their studies in the 1930s that algorithms could not solve any problem in a mathematical area. For example, whether a sentence of first-order logic is a theorem or not is one issue, and whether a polynomial equation is an integer solution in several variables, which depends on the given domains of AI. Human beings have always been able to resolve problems within these areas and this has been brought forward as a claim that robots are simply unable to do that quickly. This is because they need to be very wise and if the elements are planned according to mathematical logic, then we will be able to infuse it with knowledge in a computer program.

Nonetheless, it is important to consider a brief picture of the divisions or the essential sections of AI before thinking about the help that linguistic scientists got. McCarthy [11] addressed these branches of AI in his numerous works and many online definitions of these categories of AI are essential. The subsequent writing is an attempt to present certain sections to readers in the simplest way without distorting the main topic described by McCarthy.

a) Logical AI:

There is a certain scenario where a machine learns about the world in general and its aims can be represented in statistical or logical terms. The plan, by applying such stored data, determines what activities are required and what is sufficient for the achievement of its objectives. In fact, this is the case on an average PC-the computer works based on the task. But AI is far more complicated, ambiguous and voluminous to an ordinary machine running any user friendly software in this context.

b) Search:

Each AI-based computer system also needs to examine several choices e.g. motions in a chess game or inferences of a theorem proving method, etc. The ongoing improvements in methods and new discoveries are made in many other areas of how this can be achieved more effectively. It is named "Quest" by AI scientists.

Pattern recognition:

The program often allows a system to match what it sees by pattern if a computer program has some form of discovery. For instance, a vision system, in order to find your face, can try to match an eye and a nose design. More nuanced phenomena, such as voice acceptance in natural-language text, in chess roles and historical events, can also be explored and examined by this project. Such complicated patterns require different approaches than the most researched simple patterns. With the aid of new research, the aforesaid list of voice recognition systems and vision recognition systems have reasonably made their advent into the commercial world and with the help of new research more complex task-based techniques which require more complex logic and vigorous programming are in the making now.

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

c) Representation:

In order to articulate the ideas in the computer programs in some manner, information about the universe must be provided. Fortunately, natural mathematical reasoning languages are highly helpful and useful in carrying out this research in AI. Nevertheless, it is often said that the AI also stands early in its evolutionary stages in terms of portraying information for the purpose of building awareness. The representation approach must be so simple to use for logic and can be quickly tested and modified because the information is accessible in many ways. It can be readily evaluated as important or invalid as applicable to certain issues. Researchers have not been able to do justice to the issue of representation of facts and so the creation of AI remains elusive. The voluminous details or documents constantly changing are the explanation for this vulnerability. Many approaches currently in use include constraints, frames, reasoning, link repositories, texts, semántic networks, and description charts which are not adequate to apply AI efficiently.

d) Inference:

The' inferior' approach is a big conceptual tool. The assumption is that explicit information can be derived from implied representations. It is known in AI as a system responsible for information development. Deductive knowledge uses a series of axioms to generate new assumptions. A series of facts or observations begins with inductive inference, generalizing, explaining and legislation. The computer program demonstrates inferences that some information can be separated from others. Mathematical and logical deduction is appropriate for such purposes. But the logical inference was added in the early 1970s with new methods of non-monotonic inference. The simplest form of non-monotonic reasoning is the default reasoning in which a conclusion must be deduced by default, but if the evidence goes against its intended inference the conclusion may be withdrawn. It helped the AI save time for taking the right indexing decisions. When we learn from a duck, for example, we may assume that it can float, but this statement can be revoked when we think of a penguin. It is likely that an assumption which represents the non-monotonic essence of thought may have to be removed.

e) Common sense (knowledge and reasoning):

This is one area which could help people fight smart computers. This is where AI is the farthest forward. Although significant advances have been made, for example in the development of nonmonotonic thinking systems and theories of action, many more ideas are needed to improve the functioning of the system. A broad yet spotty set of common sense information is found in the Cyc-system [12] applied to the system.

f) Knowledge education:

This stage is very critical in the development of natural languages. In this process of discovering from practice, AI methods focused on connection and the neural networking should become advanced. The rules contained in reasoning are also known. Programmes can only work out what their formalisms will mean, so, sadly, almost everyone is dependent on very limited information capabilities. However, recent efforts are made to generate log references in AI programs that can record the steps if the desired value is not computed and this log reference can be later used to prevent such wrong steps in computing.

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

g) Planning

The planning phase in the AI programmes, especially in terms of action results, begins with a general fact of the planet. It is also concerned with any specific situation and the determination of a target. From these, a plan is developed to achieve the objective. The technique is only a series of acts in the most common cases.

h) Epistemology:

In AI, epistemology is concerned with resolving problems during the computing process.

i) Ontology:

As we learn, ontology is an analysis of the existing things. Through AI, the systems and phrases deal with different artefacts, and we analyse the basic properties of these kinds. Focus on ontology began in the early stage of the computation in the 90s and today, particularly NLP is deemed to be an integral part of any calculation method.

j) Heuristic:

approach is a way to try to explore something integrated in a programme. The word is used in AI in several ways. In some approaches, heuristic functionality is used to assess how far a parameter can be located in a search box from one objective. Heuristic approach forecasts two nodes in a search tree to look at whether one node is greater than the other.

III. RESULTS AND DISCUSSION

The domains of Computational Linguistics

We shall now try to explore other areas in which the use of computational linguistics is focused on the aforesaid AI concepts. They have rendered our existence easier all around us.

a) Search engines:

Google et al, the first significant area related to machine linguistics, is on-line search engine. The details were digitized and shown on the protocols that are posted on the network. The data are primarily accessible in the free-floating format for digital data. In natural languages, they are indexed and digitized. Search engines are AI-compatible and powerful for marking of keywords or phrases in the search engines and for matching data stored on the "World Wide Web" (www from now on) to exploit the intention. The exponential growth in computational linguistics has provided for the gradual development of precision and improvement in the quest method.

b) Web dictionaries:

The world is full of online dictionaries. The job of computer programmers and language experts makes this easier. This is completed. The restoration of classical languages such as Greek, Hebrew, Sanskrit, and Old Tamil was possible with the help of these on-line dictionaries. For these ancient languages, there are articles and online tutorials. Without the cooperation of computer programmers and language experts, this could not have been possible. The exciting and challenging work in machine linguistics has helped us to reinvent these classical

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

languages, linking us with our traditional and cultural heritage. Thanks to this phase of digitalizing document content, most of the texts of old times in many countries are readily available on the Web.

c) Internet Transliteration:

The latest' Google India Transliteration ' tool has almost proved to be a blessing for academics and people working in Indian languages. The targeted texts can be read as much as you can by using the "Romanian edition of these Indian languages," in Indian languages such as Bengali, Gujarati, Hindi, Kannada, Malayalam, Marathi, Nepali, Tamil, Telugu, etc. We assume that this is a fantastic gift for humanity from machine linguistics. On the Internet, there are many websites where one language can be quickly converted into another. The hard work of the digital linguistics renders this too likely. The AI strategies are used for labelling the strings evaluated from one language to the next and provide nearly the exact results. It uses the AI-based strategies to classify the strings evaluated in each language and tests nearly correctly. The choice of languages is at present limited. However, further work is required and the collection must be expanded by including more languages to the list and attempts have been made in this direction.

d) Text Messages of Cell phones:

Cell phones in every area of the world have become smarter daily, as the most recent machine linguistic research has shown. The text messages in place / national languages have a significant impact, at least among cultured young people. Mobile companies including Verizon, Vodafone, Sony, LG, Samsung and others have developed their own devices and IT facilities today. Thanks to the research work in digital linguistics, the predictive text that you can type so quickly on your mobile phone is made possible. The three or four phonetic symbols of a language are positioned in one main button of a telephone in the predictive text message process. So consider that on the keyboard for the text message we have to click "before"; there are two ways to do that. Second, we could often press buttons with "abc," "def," "mno," "pqrs" and "def" (15 times) in order to find a simple word like "before" because we cannot use the smart messaging strategy. However, we can enter this word only five times by pressing the buttons if we use the intelligent predictive text messaging. This is the simplest use of CL.

e) Voice and speech recognition system:

Microsoft introduced a new tool during 1999-2000, enabling OFFICE customers to type documents hands-free on a device i.e. without the use of a keyboard. But it was not so attention to machine users because everyone had to train the PC in the method of voice recognition before they can really use this device. It is a wonderful tool. But for computational linguistics it was a great achievement. We must put in a lot of efforts and make contribution in this area of computation. This will be a great help for the upcoming generation of human society.

f) Braille computer:

In the shape of Braille machines, electronic linguistics has given the people who are physically challenged with a wonderful gift. This contribution to society will always support the joint venture between software engineers and language experts or linguists. In this field, a lot of new work is under way each day to enhance and increase operability of the system.

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

g) GPS with voice activation:

Because our mobile phones and add-in hardware are included in the so-called Global Positioning System, the system has proven extremely beneficial. It lets us navigate and enter a goal without any complication as we drive our car when our hands are not free to keep the route accessible. Thanks to current research in computational linguistics, this has also become possible.

h) Automated banking and phone machines:

Most banks in the world are completely voice-enabled and operate very well and have no safety hazard because the voice-enabled system is very advanced in terms of protection. The acoustic token on the spectrogram can never be mapped to a similar level by two men, even when the imitation is finished. This has again been mae possible by computer linguistics to create such a safe bank system. The so-called' Pin no,' which is normally a four-digit security code, is generated by the voice activation device in every branch. This is also an advantage of machine linguistics.

i) Legal records mining:

Legal documents have all been digitized in Europe and America and have been made available for cross-reference in the public domain for use in litigations of any kind. By retrieving these papers, analytical linguistics assist the jury to save time and make judgment by analysing the related cases in the past.

j) Medical records mining:

Medical records are digitized using electronic linguistics in many countries. Such records are accessible for review by medical practitioners in the context of an index of results. This app can be used by physicians to figure out the prior medical condition of patients and use this information to treat them, as well as for homeopathy. The program can be used for the homeopathy medicine also.

k) Product promotion market statistics:

This is a very new venture in computer language technology. This is also recognised in the area of B to B firms as' sentiment-analysis.' Technology collects data on different product categories from the market using webpages and forums. This mathematical method is used by the businesses to promote their products. This method is extremely popular among academics. This software works by running the keywords of any literary text and collects information that research has already been carried out by the world's scholars or not. This tool helps developers and publishers in the publishing house to discover plagiarism scenarios. We must thank the computational linguists for their effort and work in providing such a good method.

IV. CONCLUSION

We hope for a lot of improvement in CL because of the increasing relevance of artificial intelligentsia. Nonetheless, natural language processing is sophisticated: we will develop cognitive computing systems which can really come into contact with living beings in order to obtain a thorough analysis. Such technologies may mimic human understanding and are therefore used to meet various business needs, such as medical diagnosis, maximizing

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

the production, allocating money and assessing M & A, product growth approaches and consumer behaviour. Regardless of the potential of machine linguistics, its success will depend upon the right blend of language comprehension and contextual reasoning for effective text interpretation.

In the last 20 years, researchers working on artificial intelligence and computational linguistics have been trying to find principles that would enable the processing of natural languages like English by computer systems. This research is intended both to further the theoretical goals by offering a basis for a computing philosophy of communication in the natural language. It proposes to support the technological aims of designing computer-based systems that can interact intelligently with their human users. Natural language processing, in the scientific and business world, has begun in right earnest, especially after twenty years of cultivation in the academic nest. By the end of this decade, natural language interfaces would certainly be commonly implemented. They will pave the way for the genesis of a large number of programming experts with proven computer resources. They will soon provide the research community with fresh insights and a new set of powerful tools for enhancing machine—language interface.

REFERENCES

1. Kay, Martin. 1973. The MIND system. In R. Rustin (ed.), Courant Computer Science Symposium 8: natural language processing. Algorithmics Press.
2. Zwicky, Arnold M., Joyce Friedman, Barbara C. Hall & Donald E. Walker. 1965. The MITRE syntactic analysis procedure for Transformational Grammars. In Proceedings – FALL Joint Computer Conference, 317–326.
3. Petrick, Stanley Roy. 1965. A recognition procedure for Transformational Grammars. Massachusetts Institute of Technology. Dept. of Modern Languages dissertation.
4. Friedman, Joyce, Thomas H. Bredt, Robert W. Doran, Bary W. Pollack & Theodore S. Martner. 1971. A computer model of Transformational Grammar (Mathematical Linguistics and Automatic Language Processing 9). New York: Elsevier.
5. Butt, Miriam, Tracy Holloway King, María-Eugenia Niño & Frédérique Segond. 1999. A grammar writer's cookbook (CSLI Lecture Notes 95). Stanford, CA: CSLI Publications.
6. Baldridge, Jason, Sudipta Chatterjee, Alexis Palmer & Ben Wing. 2007. Dot CCG and Vis CCG: Wiki and programming paradigms for improved grammar engineering with Open CCG. In Tracy Holloway King & Emily M. Bender (eds.), Grammar Engineering across Frameworks 2007 (Studies in Computational Linguistics ONLINE), 5–25. Stanford, CA: CSLI Publications.
7. Ranta, Aarne. 2009. The GF resource grammar library. Linguistic Issues in Language Technology (LiLT) 2. 1–62. http : / / journals . linguisticsociety . org /elanguage/lilt/article/download/214/214-501-1-PB.pdf.
8. King, Tracy Holloway. 2016. Theoretical linguistics and grammar engineering as mutually constraining disciplines. In Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar, 339–359. CSLI Publications.
9. Müller, Stefan. 2014. Kernigkeit: Anmerkungen zur Kern-Peripherie-Unterscheidung.In Antonio Machicao y Priemer, Andreas Nolda & Athina Sioupi(eds.), Zwischen Kern und Peripherie (studia grammatica 76), 25–39. Berlin: de Gruyter.
10. Bolshakov and Gelbukh (2004) Computational Linguistics: Models, Resources and Applications.
11. McCarthy 1990, 2000 and elsewhere, an useful link is http://www- formal.stanford.edu/jmc/mcc59.html
12. http://www.cyc.com/cyc/technology/whitepapers_dir/Cyc_Architecture_and_API.pdf
13. Jargin, Sergei V, and . 2019. Drugs and dietary supplements with unproven effects in research and practice. Journal of Complementary Medicine Research, 10 (1), 27-37. doi:10.5455/jcmr.20181223075028
14. Buch, N., Goyal, J., Kumar, N., Parmar, I., Shah, V., Charan, J.Prevalence of hypertension in school going children of Surat city, Western India(2011) Journal of Cardiovascular Disease Research, 2 (4), pp. 228-232. DOI: 10.4103/0975-3583.89807

DOI: 10.37200/IJPR/V24I8/PR280128
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020