



4. Language based research in Computer Science: A Introduction

Dr. Ambrish Tripathi

Associate Professor

MUMT, Bilaspur (C.G.)

Dr. Indra Kumar Pandey

Assistant Professor

MUMT, Bilaspur(C.G.)

Abstract

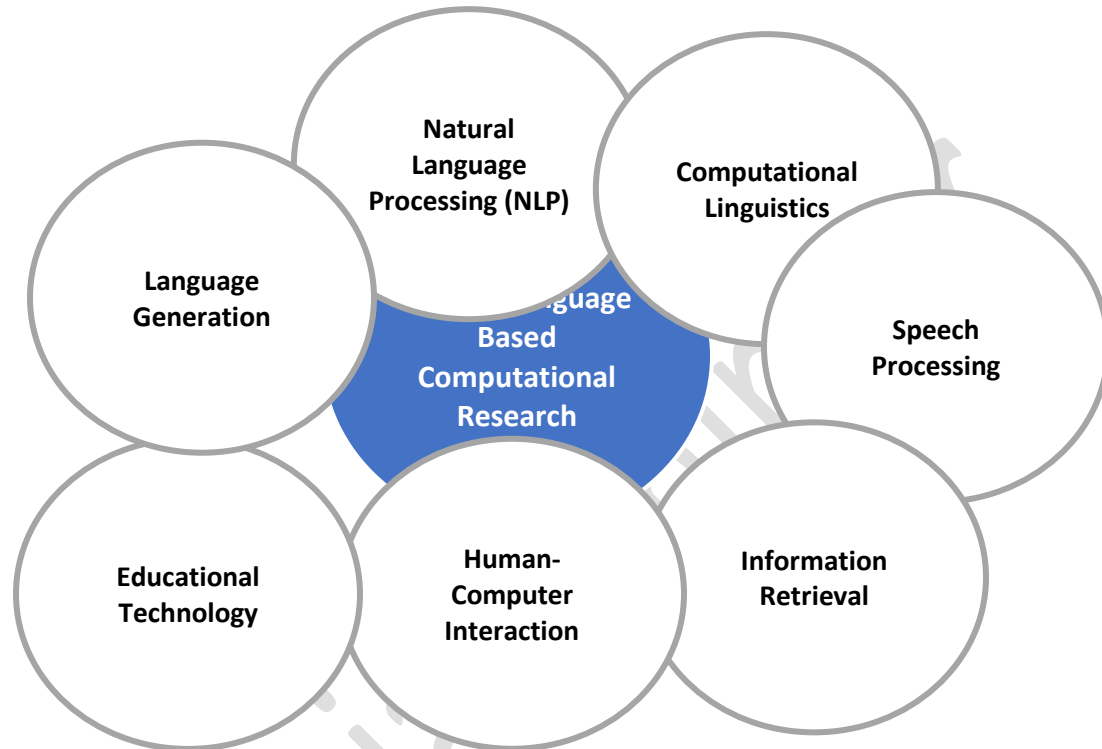
Language is a system of random speech symbols. Through which man expresses his feelings and thoughts. Every language has its own symbol system, in which feelings and thoughts are expressed through sound, word, phrase and sentence etc. Due to this difference in symbols, one language differs from another; that is to say that any language is completely independent from any other language and has its own independent grammatical rules, which we call linguistic structure. This internal system is called the grammar of that language.

Keyword- Language based research, Computer Science, Natural Language

Introduction

Machines are very important and beneficial in human life. The use of machines has become useful not only for needs but also for making entertainment easier. Humans want to understand and talk in different languages or translate any text in their own language without knowing any other language, There are some automatic translation applications to communicate in different languages that overcome language barriers like Google Translate, Microsoft Translator, etc. which is now becoming possible through research and development. Humans use machines like their everyday objects. Making machines understand natural language is also a matter of challenge. The machine is also useful for people with disabilities as technologies such as voice assistants enable hands-free interaction and accessibility, making technology more inclusive.

omputers can also analyze large amounts of text, revealing sentiment, themes, and trends in the text, which benefits researchers, marketers, and writers—such as when one item's review is like large(length of piece of cloth), it indicates positive, but another item's review is larger(length of shirt), which indicates negativity. There are many dimensions of research on human language in the machine perspective-



NLP(Natural Language Processing): Natural language processing is a field of artificial intelligence in which machines are made aware of human-like language (natural language). Its goal is to enable machines to understand, interpret and respond to human language in a valuable way.

Google Translate



The main components of natural language processing are text processing, syntax and parsing, semantics, emotional analysis, named entity recognizer, machine translation etc. It is used in customer support as chatbot's and virtual assistants, as content recommendation, for getting health related information, for information retrieval. The biggest challenge in this is ambiguity.

Computational Linguistics: Computational linguistics is an emerging discipline that combines linguistics and computer science to understand human language using computational techniques and to develop models and algorithms for it. It covers many topics including natural language processing (NLP), machine translation, speech recognition, text recognition, and text analysis. Computational linguistics is moving at a rapid pace, has already achieved some important milestones and is laying the foundation for a new future, some of which are as Advances in linguistic and model conversational agents such as GPT-4, BERT and T5, CLIP, DALL-E



Speech Processing: Speech processing is an applied field of computer science that analyzes, interprets, and processes written and spoken human language. It includes a variety of techniques and applications, including: Speech recognition: Converts spoken language into text. It is used in virtual assistants, transcription services, and voice commands, determining who is speaking based on vocal characteristics related to authentication. It is often used for security and personalization, and analyzing speech to detect emotional states, which can enhance human-computer interactions.

Information Retrieval: Information retrieval (IR) in the context of language technology involves techniques and systems that help users find relevant information from large datasets, such as text documents, databases, or the web. This involves Search Engines, Document Management Systems, and Recommendation Systems etc. Information retrieval (IR) faces several challenges, especially as data volumes and user expectations grow, as Scalability from Large Datasets, Ambiguity from Words with multiple meanings can lead to confusion in search results and Different words or phrases that mean the same thing can result in missed relevant documents and User Intent Understanding Contextual Interpretation: Accurately determining what users mean based on their queries, including understanding implied intent and context, Data Quality, Dynamic Content Security and Privacy (Safeguarding user data while providing personalized search experiences is increasingly important), Natural Language Processing Limitations Understanding Context and Nuance: Despite advances, NLP models still struggle with deeper semantic understanding, sarcasm, and idiomatic expressions.

Educational Technology: Language-based educational technology (EdTech) works on enhancing language learning and teaching through digital apps and platforms. Some useful work in this field are-

1. Language based tools(apps): Mobile Application(Duolingo, Babbel) and Speech based
2. Online Courses (MOOCs, Live Tutorial)
3. Storyline and Multimedia
4. language Exchange plate form (HelloTalk)
5. Virtual Reality (VR) and Augmented Reality (AR)
6. Cultural Integration: Incorporating cultural aspects into language learning to enhance understanding and appreciation of the language context.



Language generation: It is an important area within natural language processing (NLP) and has a wide number of applications. It refers to the process of producing organized and relevant content based on given sentences, words or symbols. Here is an overview of the key concepts, technologies and applications related to language generation:

1. **Natural Language Generation (NLG):** The subfield of NLP focused on to creating people based -like text from structured data or unstructured inputs.
2. **Models:**
 - Rule-Based Systems: Early systems used predefined rules to generate text but lacked flexibility.
 - Statistical Models: Leveraged probabilistic approaches for better fluency and coherence.
 - Neural Networks: Recent advancements use deep learning models (e.g., RNNs, LSTMs, Transformers) to produce more nuanced text.

Techniques use in Language Generation is –

Techniques Text completion: Used in chat bots and writing assistants to generate text that can complete a given prompt. , Paraphrasing: Recapitulating text and creating content without changing the meaning., Dialogue systems: Replacing human language with an application's Dialogue. Story generation: Creating narratives or creative content, often using prompts or themes provided by users. Controlled generation: Creating text based on specific constraints such as tone, style, or subject matter.

Reference:

1. <https://www.javatpoint.com/artificial-intelligence-ai>
2. [www.google .co.in](http://www.google.co.in)



3. Prof. Neha Salini, Asst. in SDAM College Dinagar, "Research paper on artificial intelligence & Its Application", (ISSN:2456-3315)
4. मल्होत्रा, विजय कुमार, कंप्यूटर के भाषिक अनुप्रयोग, वाणी प्रकाशन, नई दिल्ली।
5. James Allen (Author) "Natural Language Understanding" 2nd Edition.
6. <https://www.qualtrics.com/en-au/experience-management/customer/natural-language-generation/>
7. [https://www.aidataanalytics.network/data-science-ai/articles/what-is-natural-language-generation#:~:text=Natural%20language%20generation%20\(NLG\)%20is,\(NLU\)%20to%20generate%20NLP.](https://www.aidataanalytics.network/data-science-ai/articles/what-is-natural-language-generation#:~:text=Natural%20language%20generation%20(NLG)%20is,(NLU)%20to%20generate%20NLP.)