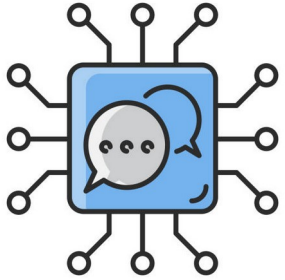
Language Analysis and Computational Linguistics

Tushar B. Kute,
http://tusharkute.com



Language Analysis

- Analysis refers to how the writer conveys meaning through language techniques, such as figures of speech, sentence structure, tone and word choice.
- When analysing language you must show that you are aware of how it is written.
- This means identifying the language features used and explaining their effect.

Language Analysis

- For example, in persuasive speaking or writing what you are primarily examining is the art of rhetoric.
- This means more than the identification of persuasive techniques used by writers; it means that you must identify how writers use these techniques to persuade.
- If you only identify the technique being used you are only halfway there. Although it is a necessary step in the analytical process, of itself, it is not sufficient.

Language Analysis

- Language analysis explores how authors/writers/speakers convey meaning through specific language analysis techniques, such as figures of speech, word choice, sentence structure, and register or tone.

Language Analysis: Levels

- There are a number of levels at which language analysis can be done. These levels are:
  - Phonetics and phonology - the sound system of a language; for example, its phonemes (units of sound) and prosody (the rhythm and intonation of language).
  - Grammar, including morphology - the analysis of the way words, clauses, phrases and sentences are put together in English. This is sometimes called 'syntax'.
  - Discourse - extended communication in a range of texts and contexts.
  - Pragmatics - the way language is used in specific contexts.

Language Analysis: Techniques

- Tools and techniques for language analysis include: consideration of audience, literary purpose, genre, mode, and literary representation.
- These help readers understand language on its own terms, as well as in literary, social, and geographical contexts.

Language Analysis: How?

- Language analysis explores how a text:
  - Is shaped according to the context, audience, and the conventions of genre.
  - Aids the reader to explore the relationship between readers and characters.
  - Constructs meaning, intention, and viewpoints.

Language Analysis: Genre

- Genre groups literary texts into styles, shared conventions, settings, and themes. Genres include poetry, novels, plays, short fiction, blogs, films etc.
- For example, George RR Martin's Game of Thrones (1996 - present) series is classified as fantasy.
- Fantasy novels are always set in a fictional universe but are inspired by real-world myth and folklore. These novels may use magic and mythical creatures, such as dragons, as part of their genre.

Language Analysis: Audience

- The audience is the anticipated, or target, readership. The author must consider how the text suits the intended audience.
- In literary works, characters can either stand-in for the audience or directly interact with them. For example:
  - A chorus is a group of people who sing or comment on the dramatic actions or events in a play.
  - Self-conscious narrators may address their audience directly.
  - An 'audience surrogate' is a figure the audience can identify with because they ask questions the audience wants to be answered.

Language Analysis: Foregrounding

- Foregrounding is a literary device which makes an image, symbol, or word a prominent or important feature in the text. It is an attention-seeking device that repeats content or breaks established patterns and calls the reader's attention to the author's language choices.
- There are two foregrounding techniques to highlight at this stage:
  - Parallelism is a technique that repeats content with unexpected regularity. The text is foregrounded by repeated patterns and figures of speech, such as alliteration, enjambment, and anaphora.
  - Deviation happens when an author sets up and breaks deliberately established patterns of language and sound for effect. This is achieved through external or internal deviation.

Language Analysis: Literariness

- Literariness explores the literary value of texts.
- Literary texts are considered to have aesthetic and moral value, akin to works of art.
- Such texts are written in such a way as to elevate them from other texts.
- They might win a literature prize like the 'Man Booker Prize', or feature in A-Level and University syllabuses.

Language Analysis: Literariness

- General ideas of literariness can be subjective.
- What you find meaningful or beautiful may differ from what another person finds meaningful or beautiful.
- However, literary theorist Roman Jakobson argues that works have a characteristic set of textual properties.
- Such properties include the use of certain poetic and literary devices, such as figures of speech.

Paradigm and Syntagm

- Paradigm and syntagm are two concepts in semiotics that direct how signs relate to one another.
- Both these concepts are used in the textual analysis to the effective communication using signs.
- The key difference between paradigm and syntagm is that paradigms are about substitution whereas syntagms are about the positioning.
- The two terms syntagmatic relationship and paradigmatic relationship also stem from syntagm and paradigm and refers to the relationship with other syntagms and paradigms, respectively.

What is Paradigm?

- A paradigm is a set of linguistic items that create mutually exclusive choices in particular syntactic roles.
- A paradigmatic relationship includes signs that can replace each other. This replacement usually changes the meaning.
- In paradigmatic relationships, a lexical item can be typically substituted by another of the same category.
- For example, a noun is replaced by a noun and a verb by a verb.

What is Paradigm?

- If we take another look at the example sentences, we used for syntagms,

Anne killed a mosquito.

- Here, mosquito can be replaced by a variety of words that would imply similar meanings.
- Anne might have killed a spider, bug, ant or another insect. Each of these words belongs to a paradigm of animals or insects Anne might have killed.

What is Syntagm?

- Syntagm is a linguistic unit which consists of a set of linguist forms such as letters, phonemes or words that are in a sequential relationship to one another.
- A syntagm is also known as a chain of signifiers.
- The relationship within a syntagm is known as a syntagmatic relationship.
- Syntagmatic relationship involves a sequence of signs that create meaning as a whole. They are all about the positioning.

What is Syntagm?

- Words which make up sentences, sentences which would make up paragraphs, paragraphs which make up chapters are some examples of syntagms and syntagmatic relationships.
- To be more clear, words in a sentence can be considered as syntagms, and they form a syntagmatic relationship that gives rise to meaning.
- Changing the sequence of syntagms in the sentence can result in a change of meaning.

What is Syntagm?

- For example,

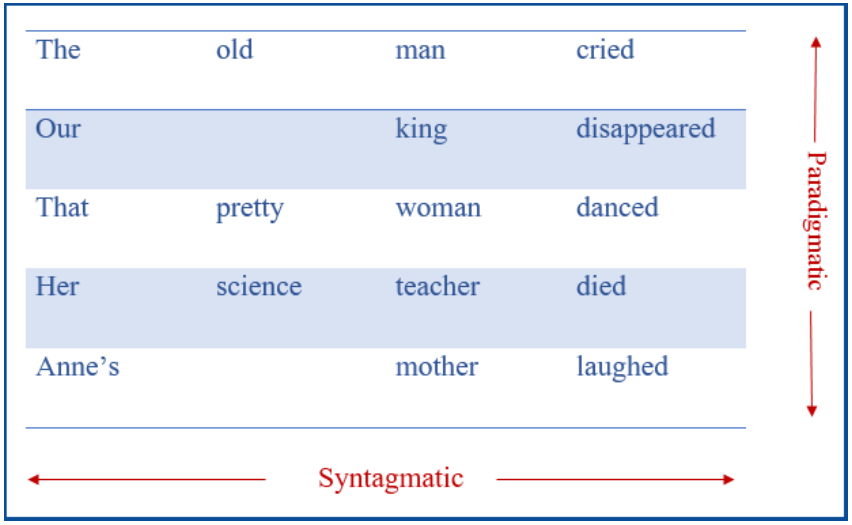
Anne killed a mosquito.

A mosquito killed Anne.

- The above sentences use the same words (syntagms), but the difference in the order (syntagmatic relationship) give two very different meanings.

Paradigm vs. Syntagm

- The following table might help you to understand the difference between paradigm and syntagm and their relationship more clearly.
- The syntagmatic relationship can be observed along the horizontal axis, and the paradigmatic relationship can be observed along the vertical axis.



Paradigm vs. Syntagm

Paradigm vs Syntagm	
Paradigm is a set of linguistic items that form mutually exclusive choices in particular syntactic roles.	Syntagm is a linguistic unit consisting of a set of linguistic forms (phonemes, words, or phrases) that are in a sequential relationship to one another.
Meaning	
In paradigmatic relationships, signs get meaning from their association with other signs.	In syntagmatic relationships, signs get meaning from their sequential order.
Focus	
Paradigmatic relationships are about substitution.	Syntagmatic relationships are about positioning.

Summary

- Paradigms and syntagms guide the relationship between signs.
- Both paradigmatic and syntagmatic relationship play a major role in determining the meaning of a linguistic unit.
- In paradigmatic relationships, the signs get their meaning with their association of other signs whereas in syntagmatic relationships, signs get their meaning from their sequence.

Form and Function

- An understanding of form and function is a variation on the "knowing what vs. knowing how" discussion:
  - Form refers to the name of a thing (along with its definition)
  - Function refers to how a thing is acting or operating in a particular context.
- When analyzing grammar in this textbook, understanding the difference between FORM and FUNCTION is one of the keys to successful analysis of words, phrases, clauses, and sentence in their context.

Form and Function

- To begin, our language can be seen as layers upon layers, and we can analyze our language from a variety of perspectives, for a variety of reasons.
- So the first point to remember is that you can identify, or name, a FORM at the word level, at the phrase level, and at the clause level.
- Similarly, you can also analyze how each FORM is FUCTIONING in the context of a sentence (or passage).
- Once again, you can identify the FUNCTION of a FORM at the word level, at the phrase level, and at the clause level.

Form and Function

- FORM names things at the WORD level (noun, verb, adjective, adverb, preposition, relative, etc.), at the PHRASE level (main verb phrase, prepositional phrase, present participle phrase, past participle phrase, gerund phrase, infinitive phrase, etc.), and at the CLAUSE level (independent clause, dependent clause, subordinate clause, relative clause, nominal clause, etc.).
- There are lots of FORMS, but a finite number of FUNCTIONS.

Functions

- In fact, there are only four (4) FUNCTIONS:
- nominal - a noun or any group of words that can substitute for a noun (word, phrase, or clause) and perform nominal functions. In other words, they can do anything that a noun can do, including filling any of the common nominal slots in a sentence (subject, direct object, subject complement, or object of the preposition most commonly).
- adjectival - an adjective or any group of words that can substitute for an adjective (word, phrase, or clause) and perform adjectival functions. In other words, they can do anything that an adjective can do (modify nouns or nominals most commonly).

Functions

- adverbial - an adverb or any group of words that can substitute for an adverb (word, phrase, or clause) and perform adverbial functions. In other words, they can do anything that an adverb can do.
- main verb phrase (MVP) - the main verb of a clause plus its auxiliary or helping verbs in an unchanging order.

Linguistic Analysis

- Linguistic analysis refers to the scientific analysis of a language sample.
- It involves at least one of the five main branches of linguistics, which are phonology, morphology, syntax, semantics, and pragmatics.
- Linguistic analysis can be used to describe the unconscious rules and processes that speakers of a language use to create spoken or written language, and this can be useful to those who want to learn a language or translate from one language to another.
- Some argue that it can also provide insight into the minds of the speakers of a given language, although this idea is controversial.

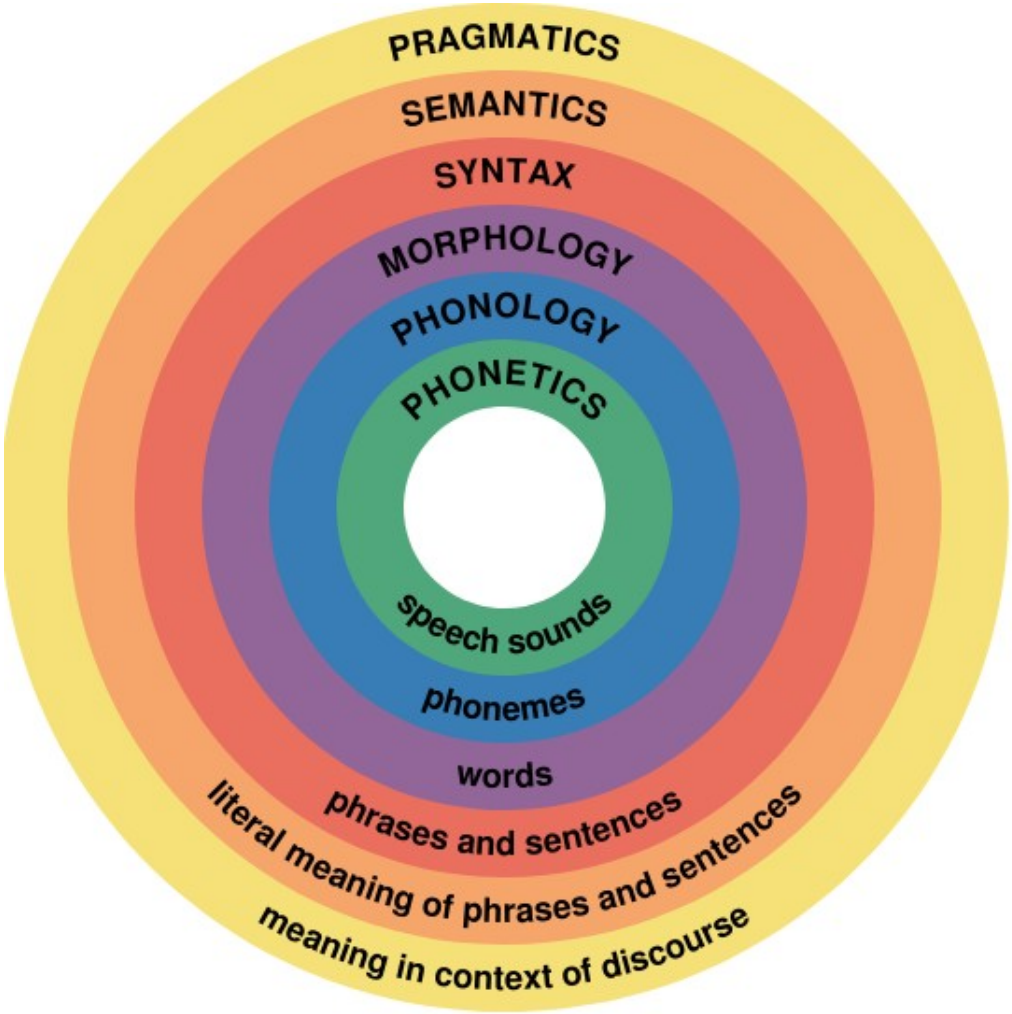
Linguistic Analysis

- The discipline of linguistics is defined as the scientific study of language.
- People who have an education in linguistics and practice linguistic analysis are called linguists.
- The drive behind linguistic analysis is to understand and describe the knowledge that underlies the ability to speak a given language, and to understand how the human mind processes and creates language.

Linguistic Analysis

- The five main branches of linguistics are
  - phonology,
  - morphology,
  - syntax,
  - semantics, and
  - pragmatics.
- An extended language analysis may cover all five of the branches, or it may focus on only one aspect of the language being analyzed.
- Each of the five branches focuses on a single area of language.

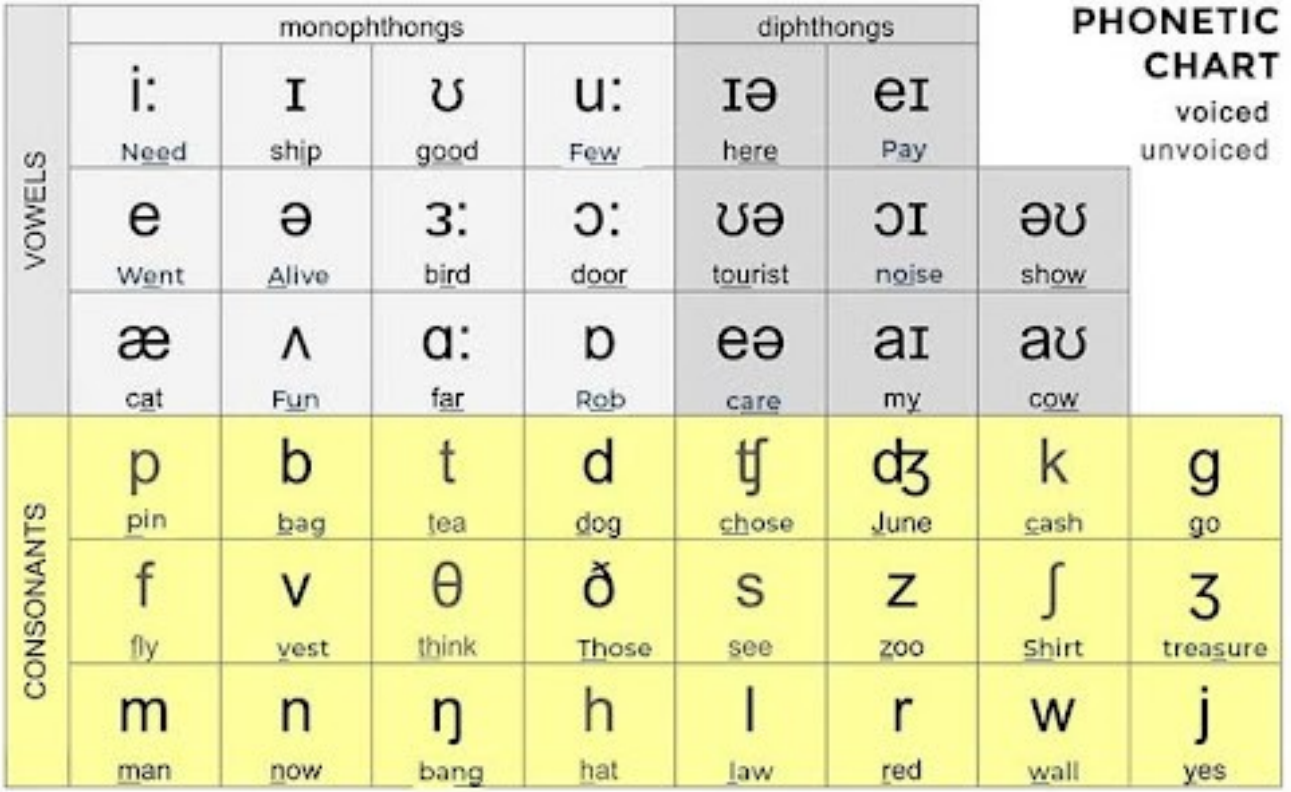
Linguistic Analysis



Phonetics

- Phonetics is a branch of linguistics that studies how humans produce and perceive sounds, or in the case of sign languages, the equivalent aspects of sign.
- Linguists who specialize in studying the physical properties of speech are phoneticians.
- The field of phonetics is traditionally divided into three sub-disciplines based on the research questions involved such as how humans plan and execute movements to produce speech (articulatory phonetics), how various movements affect the properties of the resulting sound (acoustic phonetics), or how humans convert sound waves to linguistic information (auditory phonetics).

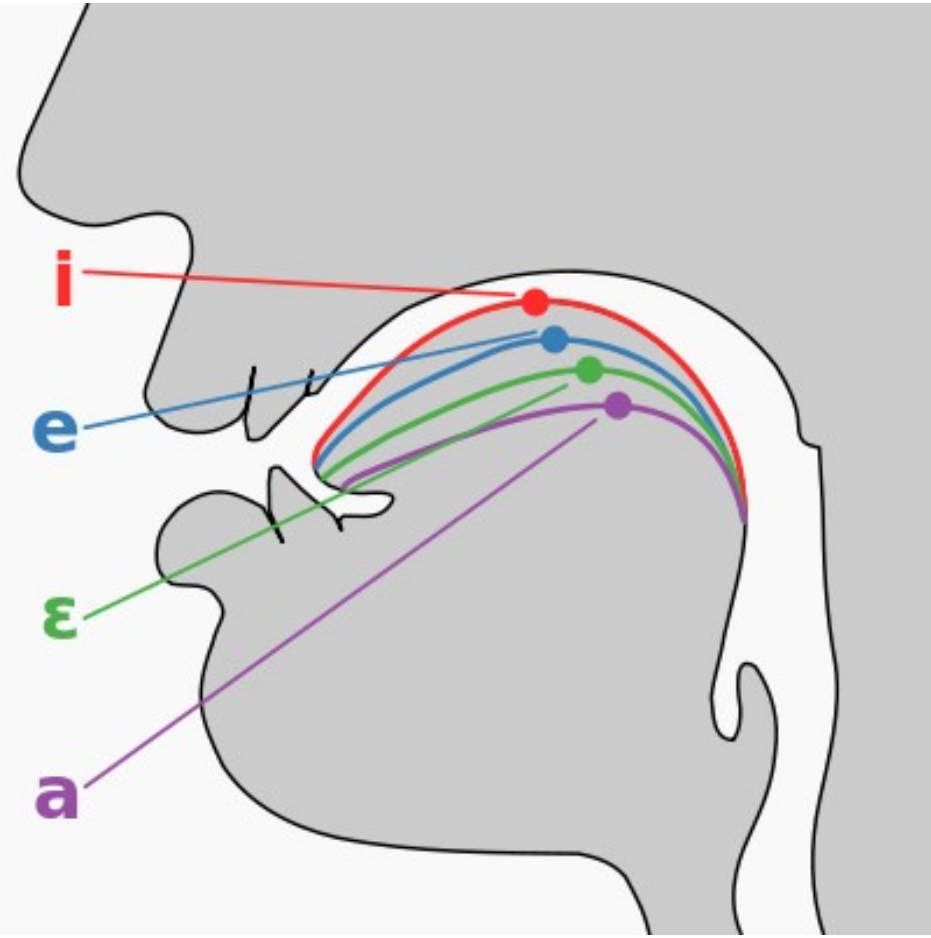
Phonetics


PHONETIC CHART
voiced
unvoiced
VOWELS
CONSONANTS
monophthongs
diphthongs
iː Need
ɪ ship
ʊ good
uː Few
ɪə here
eɪ Pay
e Went
ə Alive
ɜː bird
ɔː door
ʊə tourist
ɔɪ noise
əʊ show
æ cat
ʌ Fun
ɑː far
ɒ Rob
eə care
aɪ my
aʊ cow
p pin
b bag
t tea
d dog
ʧ chose
ʤ June
k cash
g go
f fly
v vest
θ think
ð Those
s see
z zoo
ʃ Shirt
ʒ treasure
m man
n now
ŋ bang
h hat
l law
r red
w wall
j yes

Phonology

- Phonology refers to the study of the sounds of a language.
- Every language has its own inventory of sounds and logical rules for combining those sounds to create words.
- The phonology of a language essentially refers to its sound system and the processes used to combine sounds in spoken language.

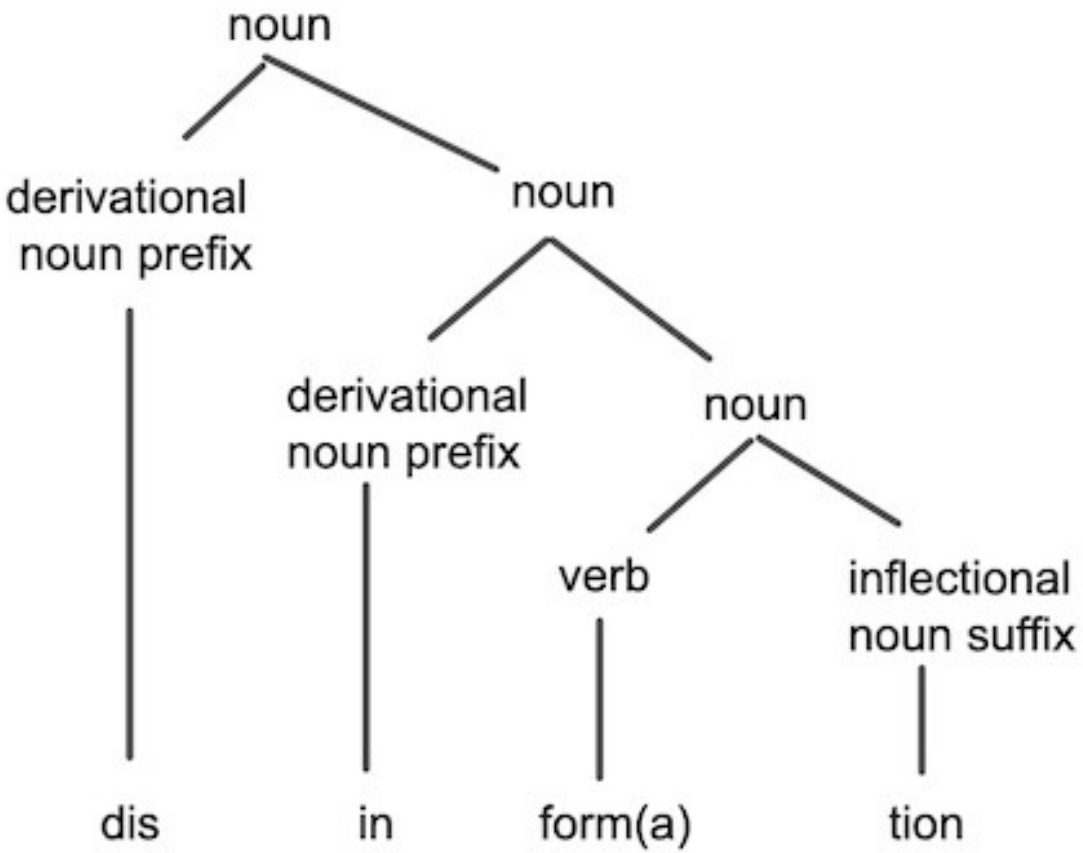
Phonology



Morphology

- Morphology refers to the study of the internal structure of the words of a language.
- In any given language, there are many words to which a speaker can add a suffix, prefix, or infix to create a new word. In some languages, these processes are more productive than others.
- The morphology of a language refers to the word-building rules speakers use to create new words or alter the meaning of existing words in their language.

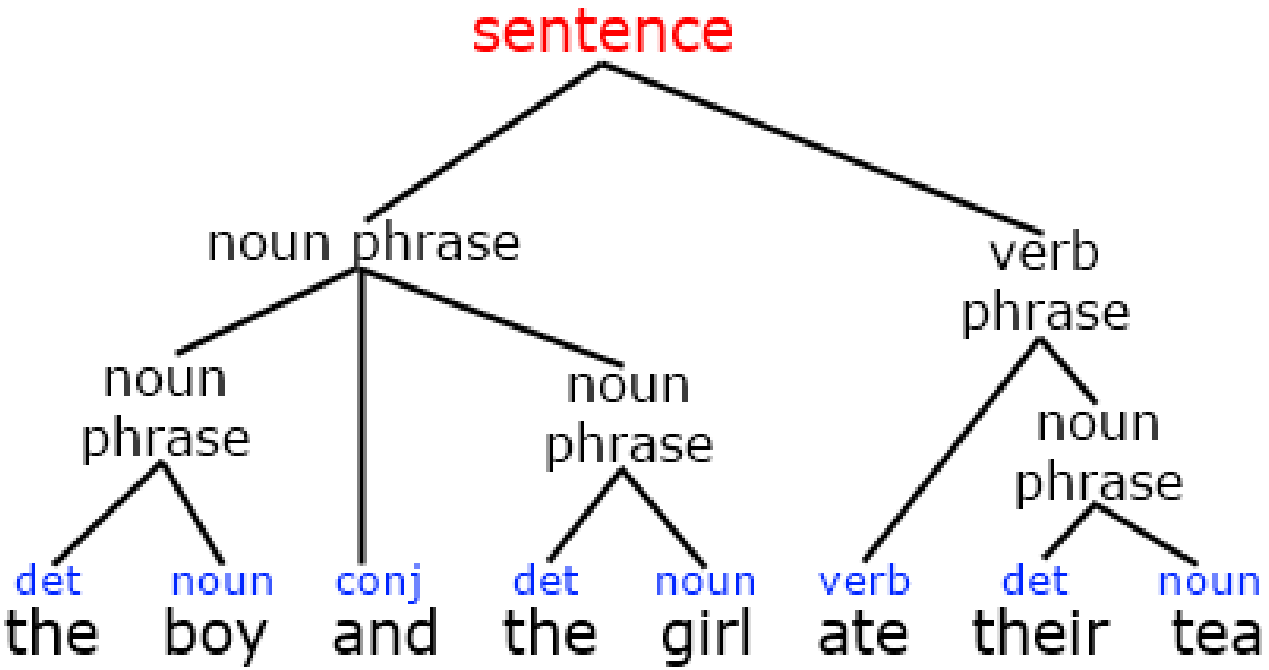
Morphology


noun
derivational
noun prefix
noun
derivational
noun prefix
noun
verb
inflectional
noun suffix
dis
in
form(a)
tion

Syntax

- Syntax is the study of sentence structure.
- Every language has its own rules for combining words to create sentences.
- Syntactic analysis attempts to define and describe the rules that speakers use to put words together to create meaningful phrases and sentences.

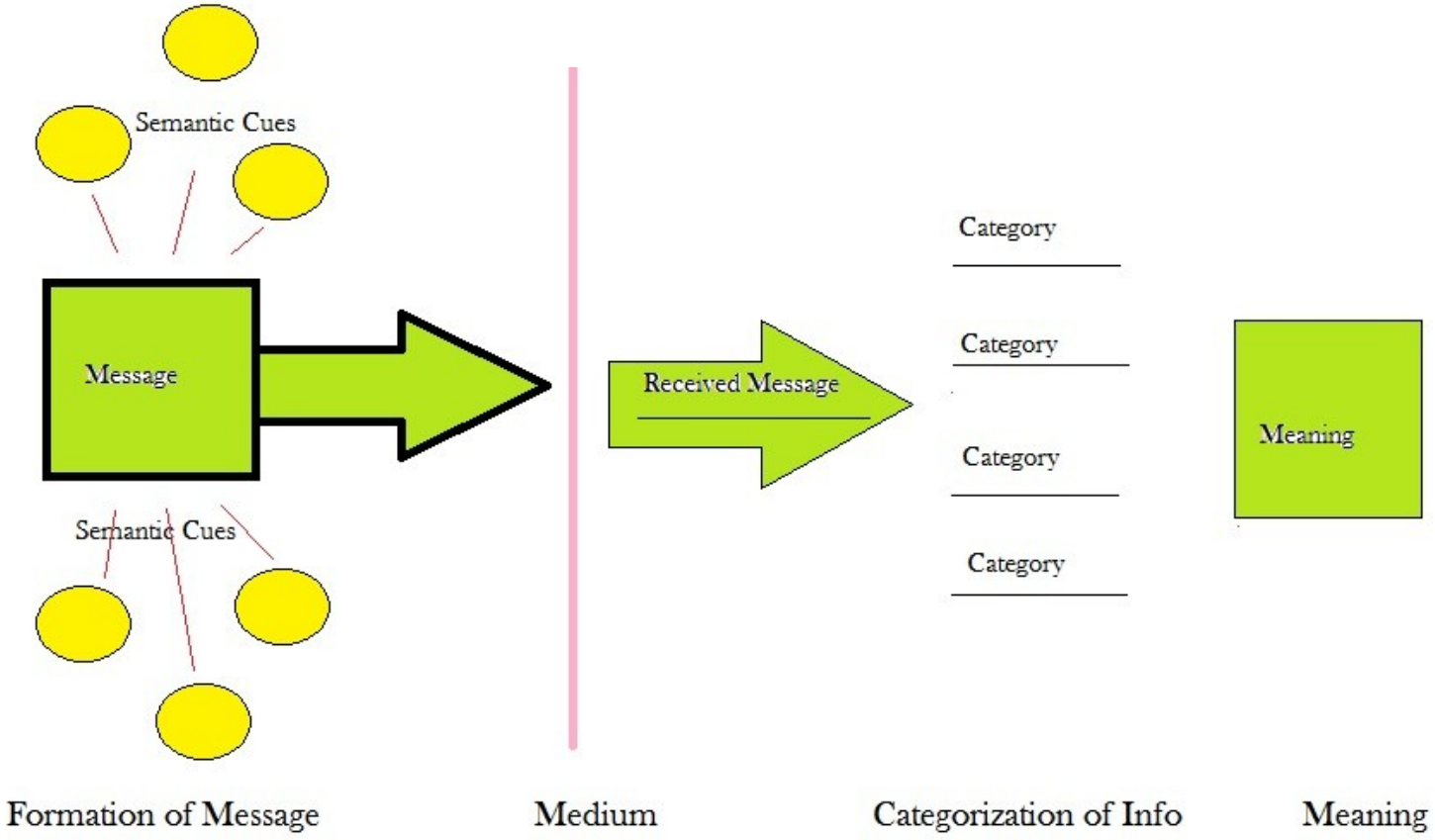
Syntax


sentence
noun phrase
verb phrase
noun phrase
noun phrase
noun phrase
det
noun
conj
det
noun
verb
det
noun
the
boy
and
the
girl
ate
their
tea

Symantics

- Semantics is the study of meaning in language.
- Linguists attempt to identify not only how speakers of a language discern the meanings of words in their language, but also how the logical rules speakers apply to determine the meaning of phrases, sentences, and entire paragraphs.
- The meaning of a given word can depend on the context in which it is used, and the definition of a word may vary slightly from speaker to speaker.

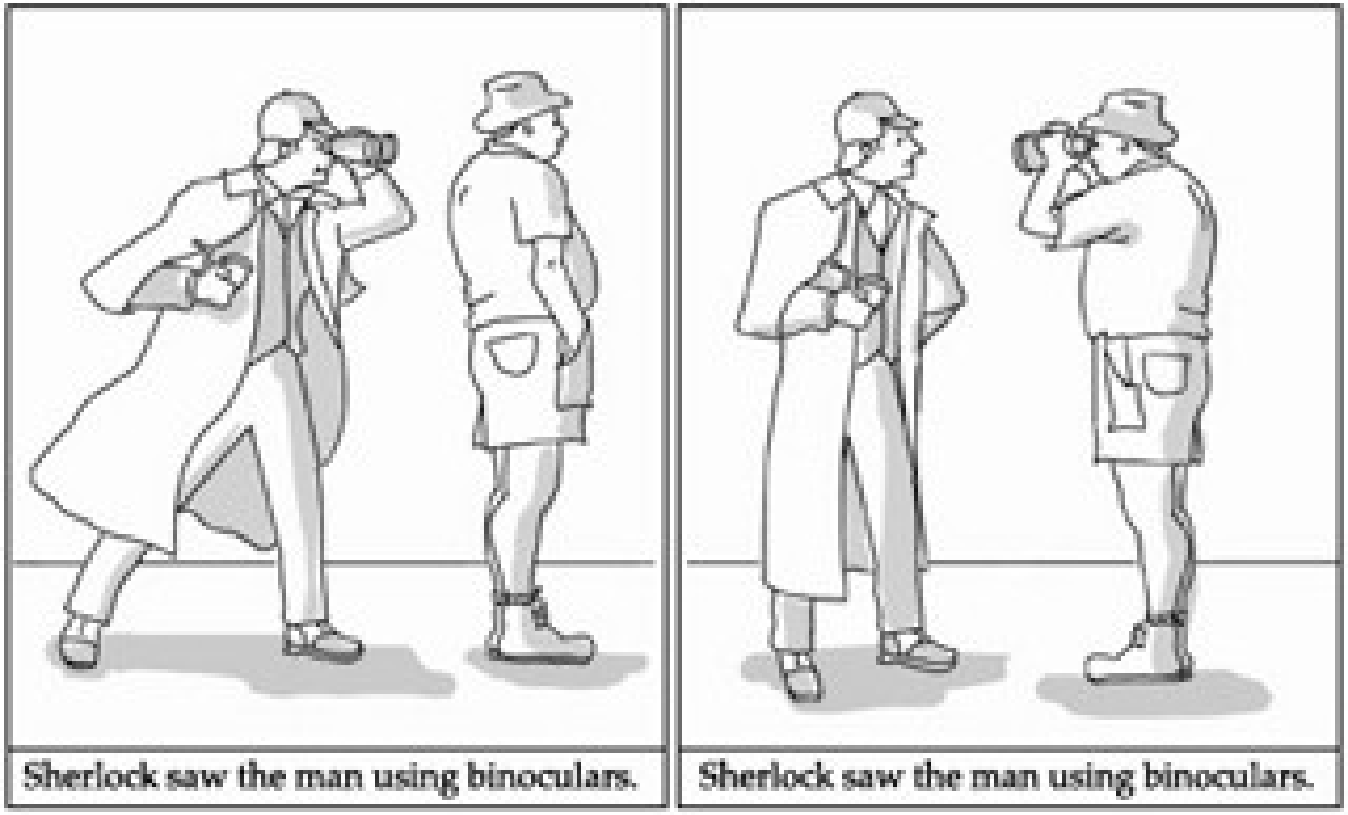
Symantics


Semantic Cues
Message
Semantic Cues
Received Message
Category
Category
Category
Category
Meaning
Formation of Message
Medium
Categorization of Info
Meaning

Pragmatics

- Pragmatics is the study of the social use of language.
- All speakers of a language use different registers, or different conversational styles, depending on the company in which they find themselves.
- A linguistic analysis that focuses on pragmatics may describe the social aspects of the language sample being analyzed, such as how the status of the individuals involved in the speech act could affect the meaning of a given utterance.

Pragmatics


Sherlock saw the man using binoculars.
Sherlock saw the man using binoculars.

Lexicology

- Lexicology is the branch of linguistics that analyzes the lexicon of a specific language.
- A word is the smallest meaningful unit of a language that can stand on its own, and is made up of small components called morphemes and even smaller elements known as phonemes, or distinguishing sounds.
- Lexicology examines every feature of a word – including formation, spelling, origin, usage, and definition

Lexicology

- Lexicology also considers the relationships that exist between words.
- In linguistics, the lexicon of a language is composed of lexemes, which are abstract units of meaning that correspond to a set of related forms of a word.
- Lexicology looks at how words can be broken down as well as identifies common patterns they follow.
- Lexicology is associated with lexicography, which is the practice of compiling dictionaries.

Linguistic Analysis: Summary

- Linguistic analysis has been used to determine historical relationships between languages and people from different regions of the world.
- Some governmental agencies have used linguistic analysis to confirm or deny individuals' claims of citizenship.
- This use of linguistic analysis remains controversial, because language use can vary greatly across geographical regions and social class, which makes it difficult to accurately define and describe the language spoken by the citizens of a particular country.

Artificial Intelligence

- Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions.
- The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

Artificial Intelligence

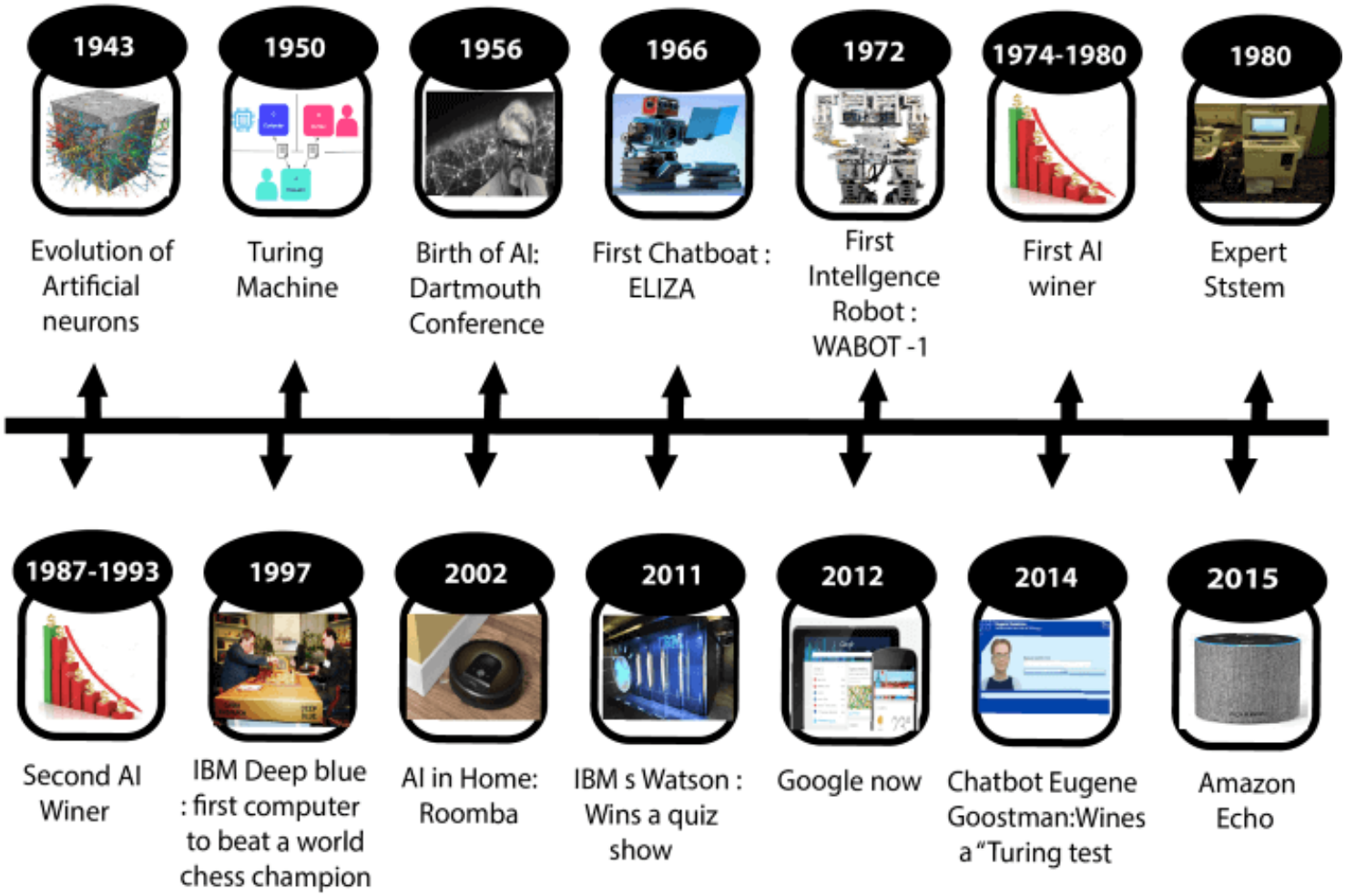
- Artificial intelligence is based on the principle that human intelligence can be defined in a way that a machine can easily mimic it and execute tasks, from the most simple to those that are even more complex.
- The goals of artificial intelligence include mimicking human cognitive activity.

Artificial Intelligence

Humans + Computers > Humans Alone

Humans + Computers > Computers Alone

Artificial Intelligence : Evolution

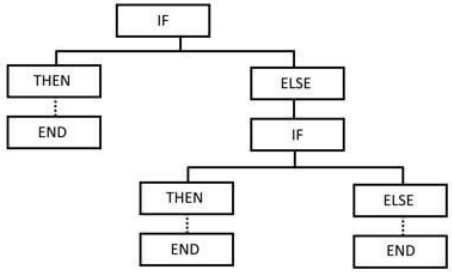


Types

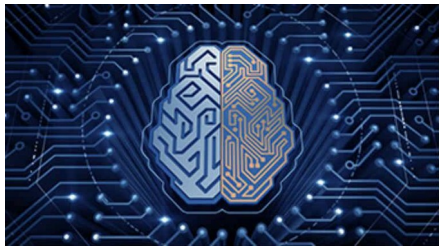
- “Strong” Artificial Intelligence
  - Computers thinking at a level that meets or surpasses people
  - Computers engaging in abstract reasoning & thinking
  - This is not what we have today
  - There is no evidence that we are close to Strong AI
- “Weak” Pattern-Based Artificial Intelligence
  - Computers solve problems by detecting useful patterns
  - Pattern-based AI is an Extremely powerful tool
  - Has been used to automate many processes today
  - Driving, language translation
  - This is the dominant mode of AI today

Major AI Approaches

- Two Major AI Techniques
  - Logic and Rules-Based Approach



  - Machine Learning (Pattern-Based Approach)

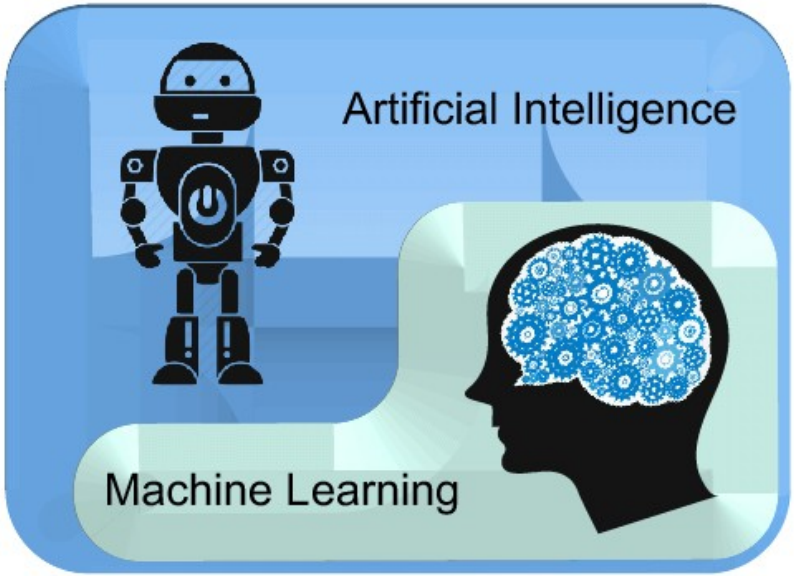


Rules based approach

- Logic and Rules-Based Approach
  - Representing processes or systems using logical rules
  - Top-down rules are created for computer
  - Computers reason about those rules
  - Can be used to automate processes

Learning vs. Designing

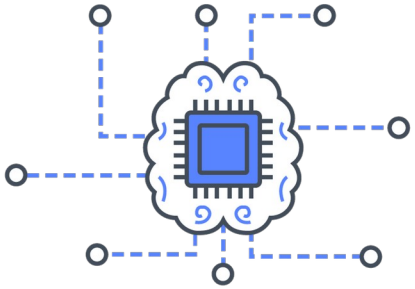
- AI is a bigger concept to design intelligent machines that can simulate human thinking capability and behavior, whereas, machine learning is an application or subset of AI that allows machines to learn from data without being programmed explicitly.



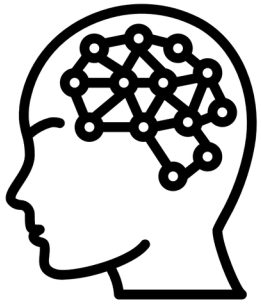
Branches Of Artifical Intelligence

- Artificial Intelligence can be used to solve real-world problems by implementing the following processes/ techniques:
  - Machine Learning
  - Deep Learning
  - Natural Language Processing
  - Robotics
  - Expert Systems
  - Fuzzy Logic

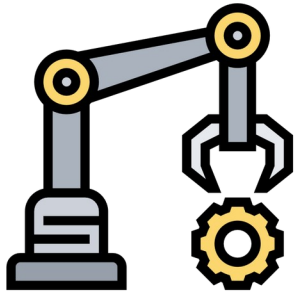
Branches Of Artifical Intelligence



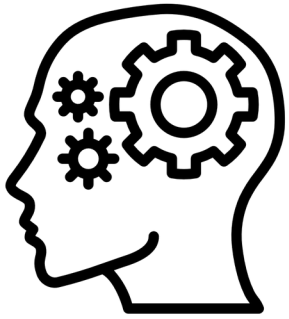
Machine Learning



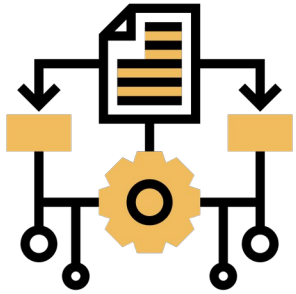
Neural Networks



Robotics



Expert Systems



Fuzzy Logic

Natural Language Processing

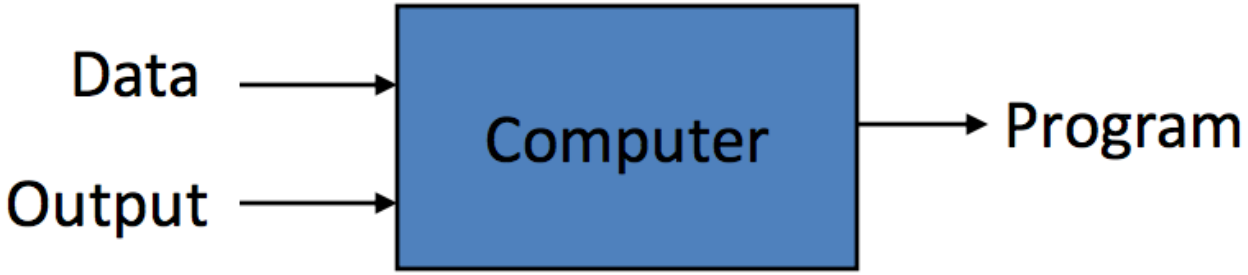
Machine Learning

- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.
- The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide.
- The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Machine Learning

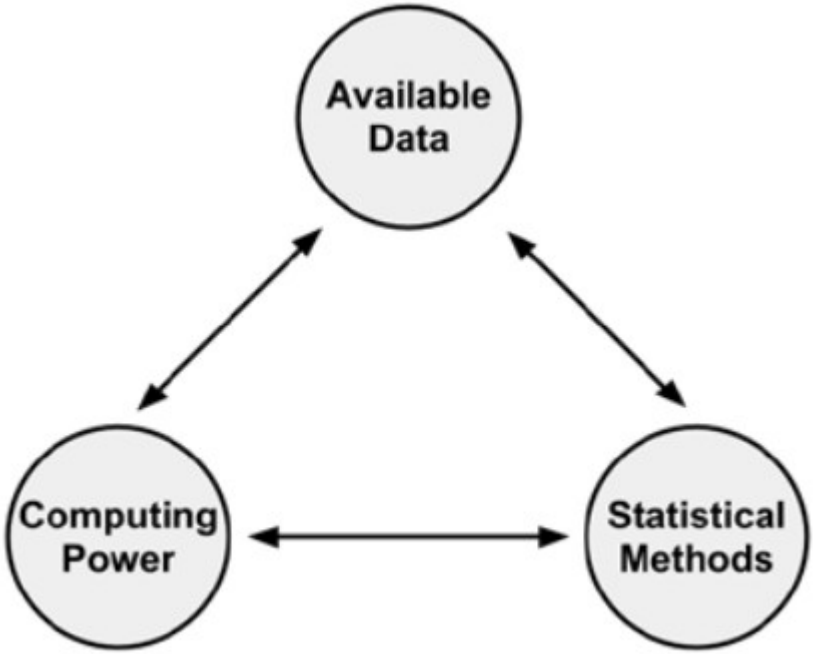
Traditional Programming

Machine Learning



Machine Learning

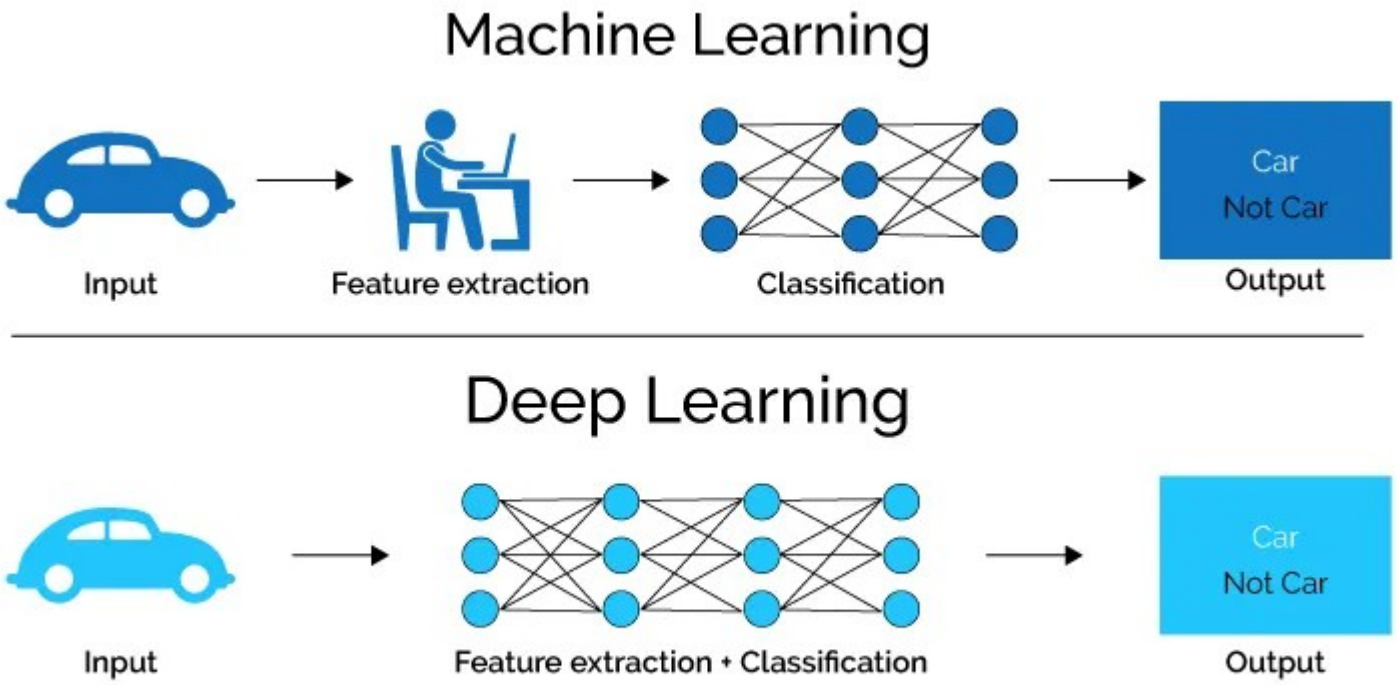
- The field of study interested in the development of computer algorithms for transforming data into intelligent action is known as machine learning.



Deep Learning

- Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers.
- These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to “learn” from large amounts of data.
- While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy.

Deep Learning



Robotics

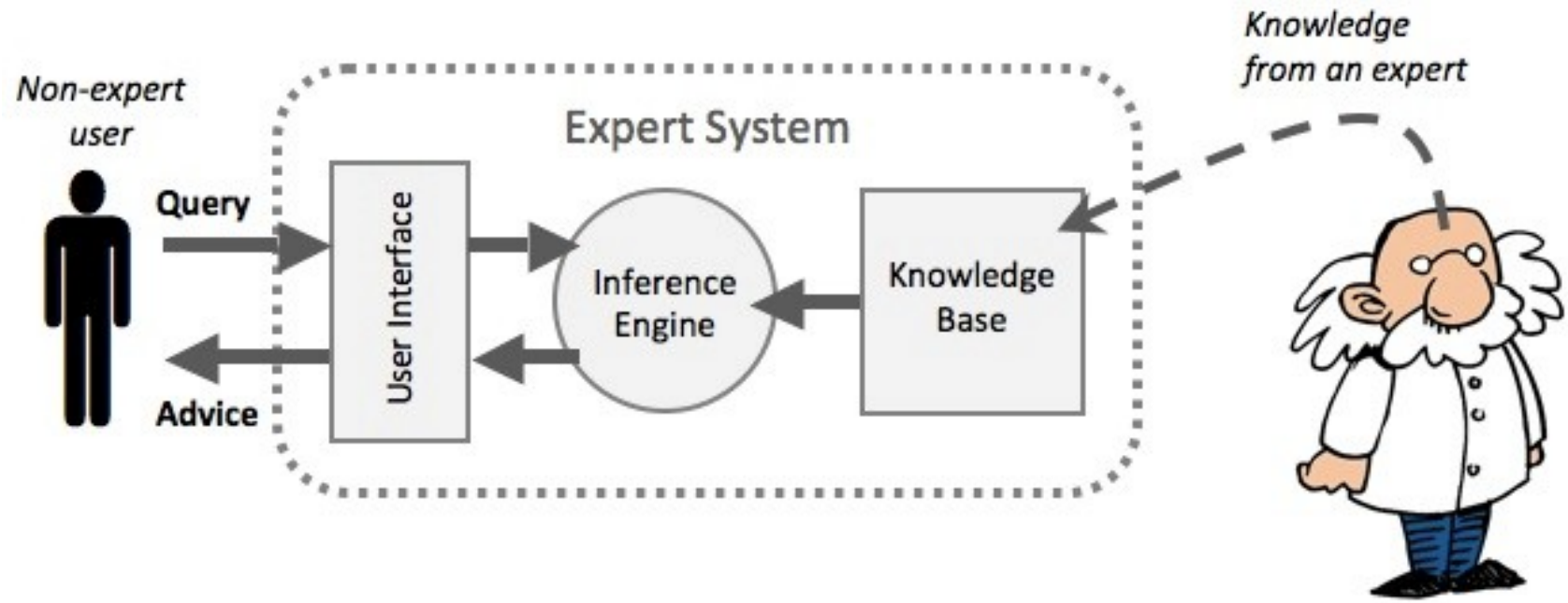
- Robotics is a branch of engineering that involves the conception, design, manufacture and operation of robots.
- The objective of the robotics field is to create intelligent machines that can assist humans in a variety of ways.
- Robotics can take on a number of forms. A robot may resemble a human, or it may be in the form of a robotic application, such as robotic process automation (RPA), which simulates how humans engage with software to perform repetitive, rules-based tasks.

Robotics

Expert Systems

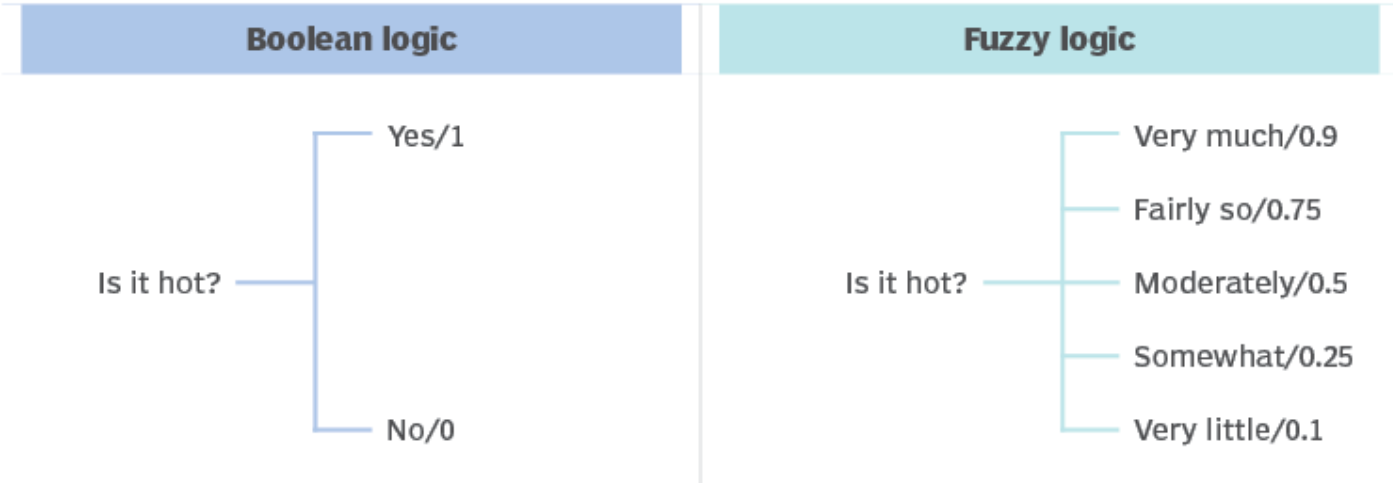
- An expert system is a computer program that uses artificial intelligence (AI) technologies to simulate the judgment and behavior of a human or an organization that has expertise and experience in a particular field.
- Expert systems are usually intended to complement, not replace, human experts.
- Expert systems accumulate experience and facts in a knowledge base and integrate them with an inference or rules engine -- a set of rules for applying the knowledge base to situations provided to the program.

Expert Systems



Fuzzy Logic

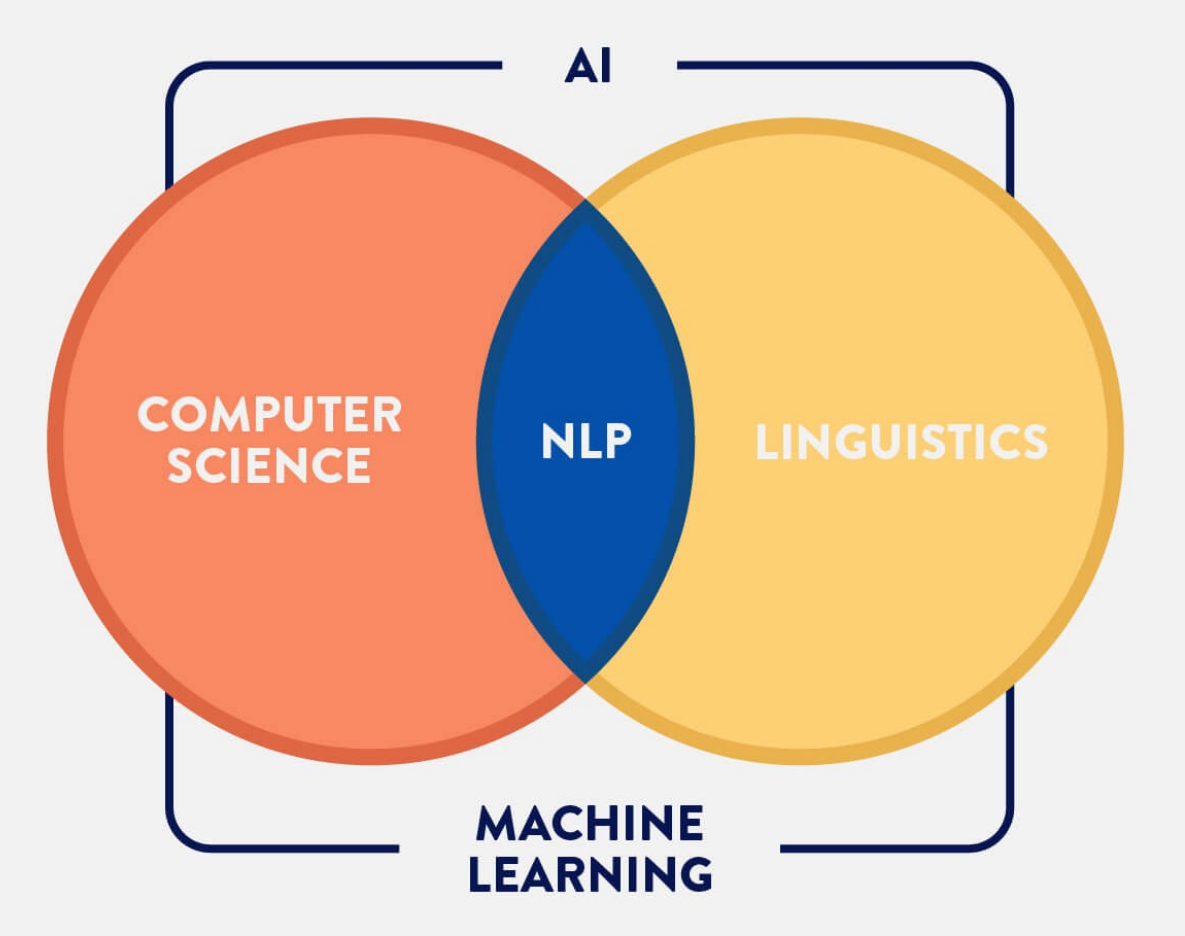
- Fuzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which the modern computer is based.



Natural Language Processing

- Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written -- referred to as natural language. It is a component of artificial intelligence (AI).
- NLP has existed for more than 50 years and has roots in the field of linguistics.
- It has a variety of real-world applications in a number of fields, including medical research, search engines and business intelligence.

Natural Language Processing



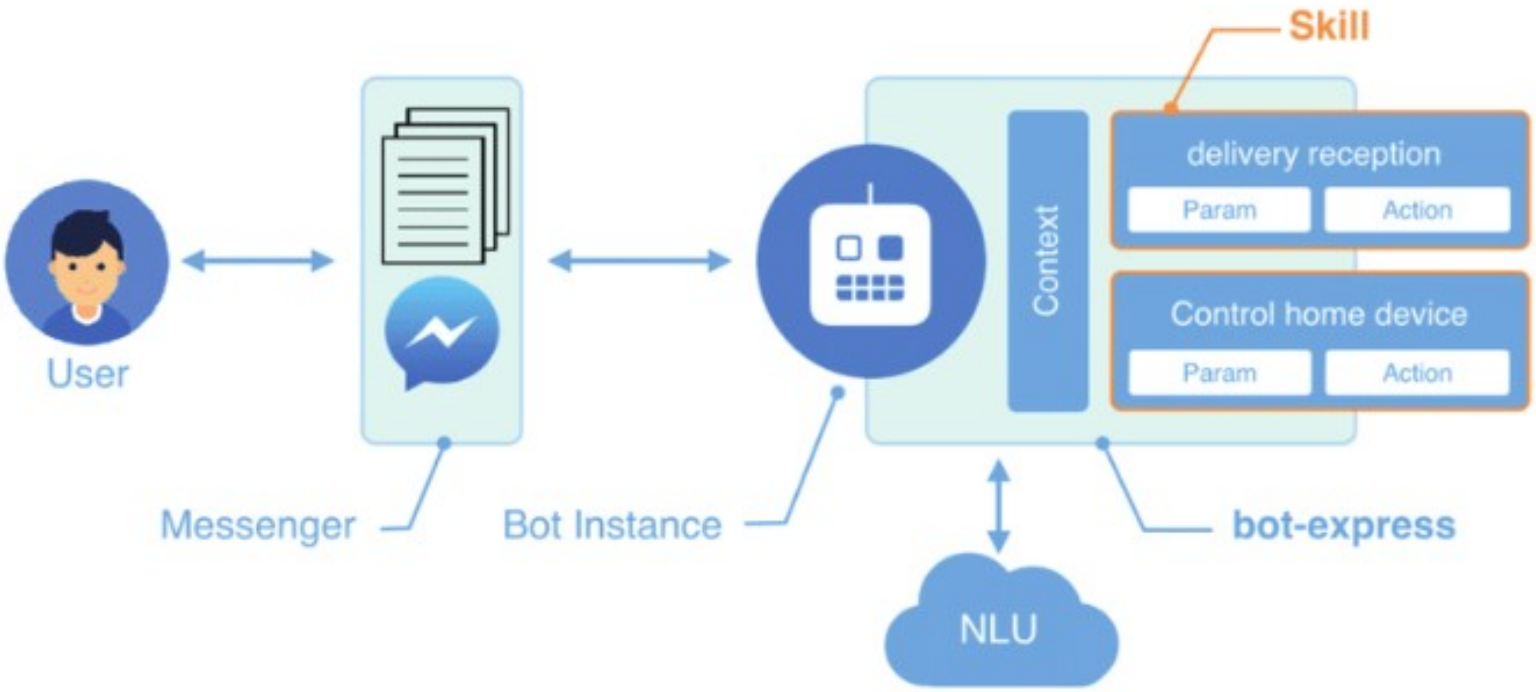
Natural Language Understanding

- Natural language understanding is a branch of artificial intelligence that uses computer software to understand input in the form of sentences using text or speech.
- NLU enables human-computer interaction. It is the comprehension of human language such as English, Spanish and French, for example, that allows computers to understand commands without the formalized syntax of computer languages.
- NLU also enables computers to communicate back to humans in their own languages.

Natural Language Understanding

- The main purpose of NLU is to create chat- and voice-enabled bots that can interact with the public without supervision.
- Many major IT companies, such as Amazon, Apple, Google and Microsoft, and startups have NLU projects underway.

Natural Language Understanding



Natural Language Understanding

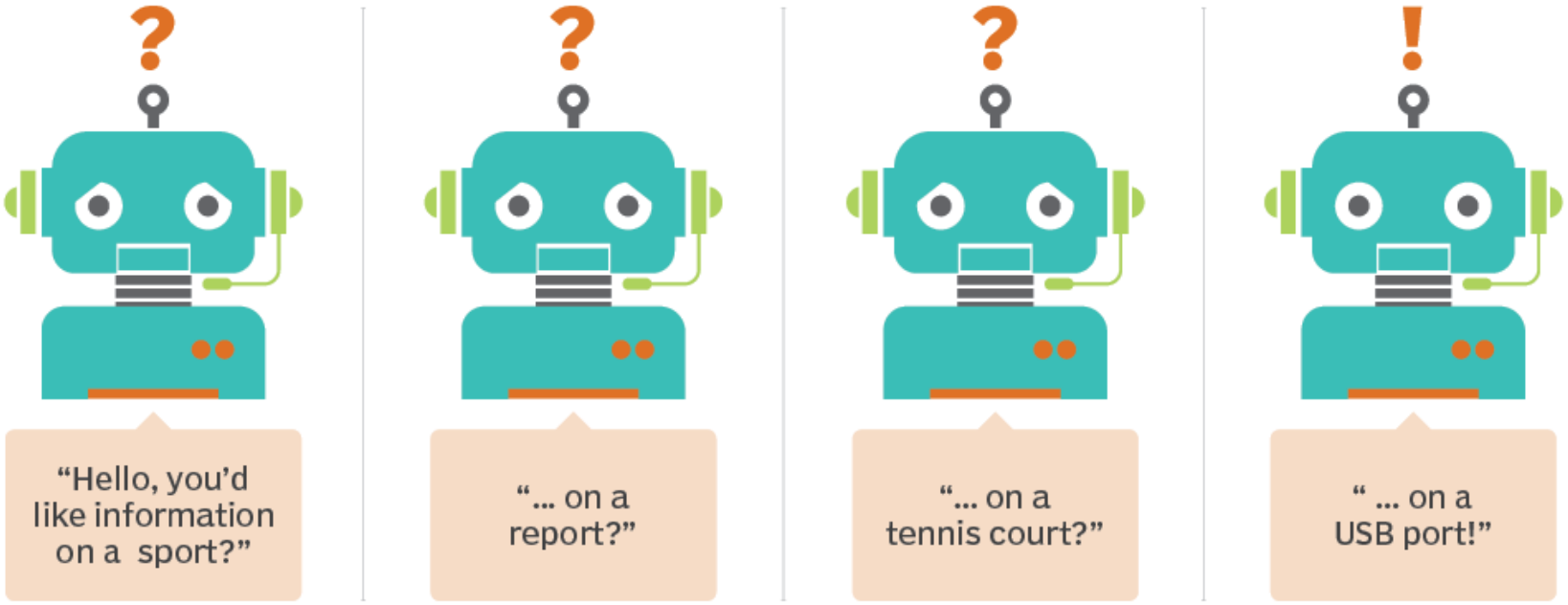
- NLU analyzes data to determine its meaning by using algorithms to reduce human speech into a structured ontology -- a data model consisting of semantics and pragmatics definitions.
- Two fundamental concepts of NLU are intent and entity recognition.
- Intent recognition is the process of identifying the user's sentiment in input text and determining their objective.
- It is the first and most important part of NLU because it establishes the meaning of the text.

Natural Language Understanding

- Entity recognition is a specific type of NLU that focuses on identifying the entities in a message, then extracting the most important information about those entities.
- There are two types of entities: named entities and numeric entities. Named entities are grouped into categories -- such as people, companies and locations. Numeric entities are recognized as numbers, currencies and percentages.
- For example, a request for an island camping trip on Vancouver Island on Aug. 18 might be broken down like this: ferry tickets [intent] / need: camping lot reservation [intent] / Vancouver Island [location] / Aug. 18 [date].

Overcoming Language Barriers

Using natural language development and analytics, chatbots will be able to learn words and phrases outside their usual repertoire as they gain more experience in conversations.



NLU Applications

- IVR and message routing. Interactive Voice Response (IVR) is used for self-service and call routing.
- Early iterations were strictly touchtone and did not involve AI.
- However, as IVR technology advanced, features such as NLP and NLU have broadened its capabilities and users can interact with the phone system via voice.

NLU Applications

- Chatbots
  - Customer support and service through intelligent personal assistants.
  - NLU is the technology behind chatbots, which is a computer program that converses with a human in natural language via text or voice.
  - Chatbots follow a script and can only answer questions in that script.

NLU Applications

- Machine translation
  - translation, a machine learning algorithm analyzes millions of pages of text -- say, contracts or financial documents -- to learn how to translate them into another language.

NLU Applications

- Data capture.
  - Data capture is the process of gathering and recording information about an object, person or event.
  - For example, if an e-commerce company used NLU, it could ask customers to enter their shipping and billing information verbally.
  - The software would understand what the customer meant and enter the information automatically.

NLU Applications

- Conversational interfaces.
  - Many voice-activated devices -- including Amazon Alexa and Google Home -- allow users to speak naturally.
  - By using NLU, conversational interfaces can understand and respond to human language by segmenting words and sentences, recognizing grammar, and using semantic knowledge to infer intent.

Natural Language Generation

- Natural language generation (NLG) is the use of artificial intelligence (AI) programming to produce written or spoken narratives from a data set.
- NLG is related to human-to-machine and machine-to-human interaction, including computational linguistics, natural language processing (NLP) and natural language understanding (NLU).

Natural Language Generation

- Research about NLG often focuses on building computer programs that provide data points with context.
- Sophisticated NLG software can mine large quantities of numerical data, identify patterns and share that information in a way that is easy for humans to understand.
- The speed of NLG software is especially useful for producing news and other time-sensitive stories on the internet.
- At its best, NLG output can be published verbatim as web content.

How NLG works?

- NLG is a multi-stage process, with each step further refining the data being used to produce content with natural-sounding language. The six stages of NLG are as follows:
  - Content analysis. Data is filtered to determine what should be included in the content produced at the end of the process. This stage includes identifying the main topics in the source document and the relationships between them.
  - Data understanding. The data is interpreted, patterns are identified and it's put into context. Machine learning is often used at this stage.

How NLG works?

- Document structuring.
  - A document plan is created and a narrative structure chosen based on the type of data being interpreted.
- Sentence aggregation.
  - Relevant sentences or parts of sentences are combined in ways that accurately summarize the topic.

How NLG works?

- Grammatical structuring.
  - Grammatical rules are applied to generate natural-sounding text. The program deduces the syntactical structure of the sentence. It then uses this information to rewrite the sentence in a grammatically correct manner.
- Language presentation.
  - The final output is generated based on a template or format the user or programmer has selected.

How NLG works: Summary

1 Content analysis	2 Data understanding	3 Document structuring	4 Sentence aggregation	5 Grammar structuring	6 Language presentation
Filtering data and identifying the important content	Interpreting data, identifying patterns and putting it into context	Creating a document plan and narrative structure	Combining relevant sentences to summarize the topic	Applying grammatical rules to create natural-sounding text	Formatting and presenting the final output

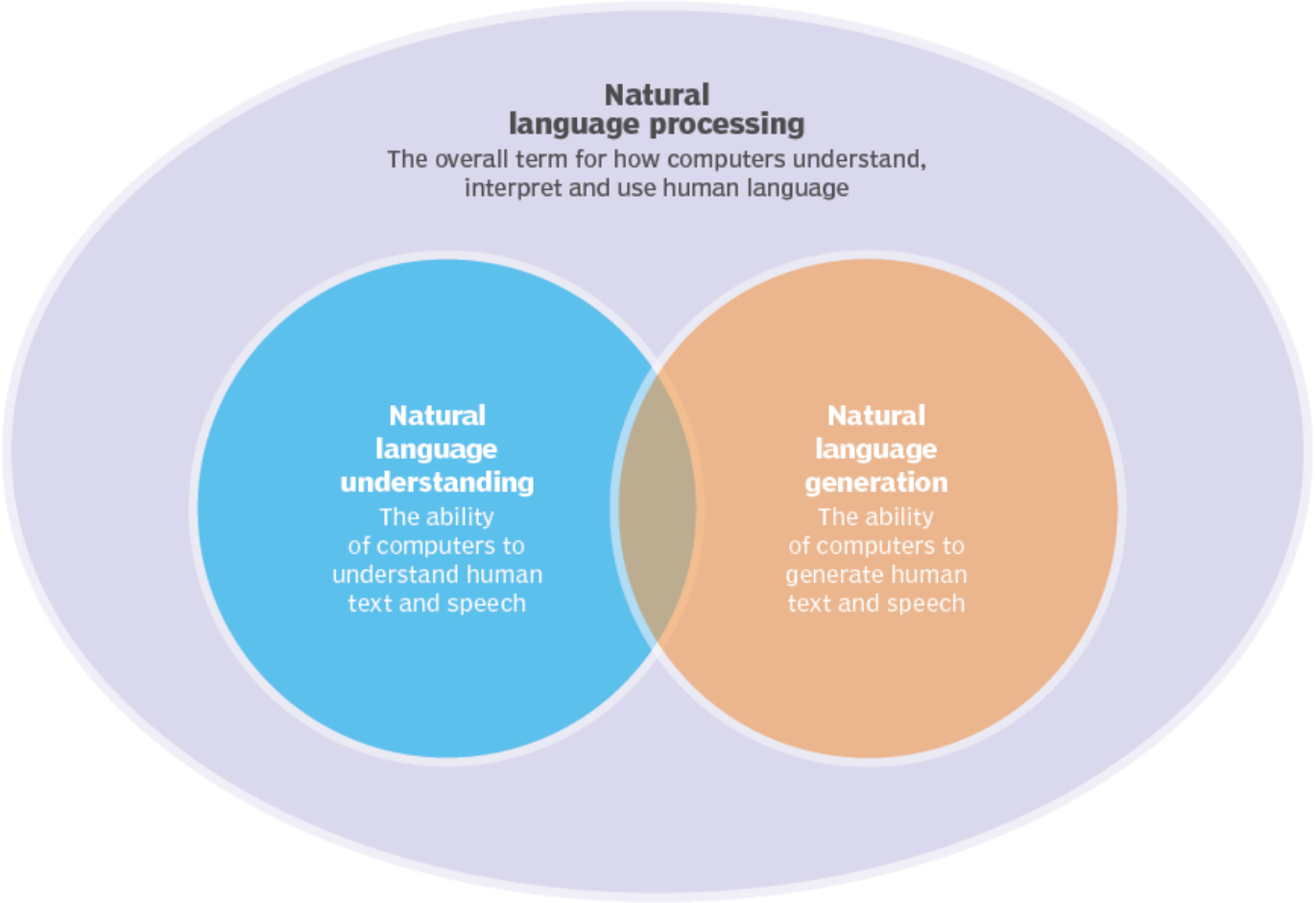
How NLG used?

- Natural language generation is being used in an array of ways. Some of the many uses include the following:
  - generating the responses of chatbots and voice assistants such as Google's Alexa and Apple's Siri;
  - converting financial reports and other types of business data into easily understood content for employees and customers;
  - automating lead nurturing email, messaging and chat responses;
  - personalizing responses to customer emails and messages;

How NLG used?

- generating and personalizing scripts used by customer service representatives;
- aggregating and summarizing news reports;
- reporting on the status of internet of things devices; and
- creating product descriptions for e-commerce webpages and customer messaging.

Relations


Natural
language processing
The overall term for how computers understand, interpret and use human language
Natural
language
understanding
The ability
of computers to
understand human
text and speech
Natural
language
generation
The ability
of computers to
generate human
text and speech

NLG Models and Methodologies

- NLG relies on machine learning algorithms and other approaches to create machine-generated text in response to user inputs. Some of the methodologies used include the following:
- Markov chain.
  - The Markov model is a mathematical method used in statistics and machine learning to model and analyze systems that are able to make random choices, such as language generation.
  - Markov chains start with an initial state and then randomly generate subsequent states based on the prior one.

NLG Models and Methodologies

- Recurrent neural network (RNN).
  - These AI systems are used to process sequential data in different ways.
  - RNNs can be used to transfer information from one system to another, such as translating sentences written in one language to another.
  - RNNs are also used to identify patterns in data which can help in identifying images.
  - An RNN can be trained to recognize different objects in an image or to identify the various parts of speech in a sentence.

NLG Models and Methodologies

- Long short-term memory (LSTM).
  - This type of RNN is used in deep learning where a system needs to learn from experience.
  - LSTM networks are commonly used in NLP tasks because they can learn the context required for processing sequences of data.
  - To learn long-term dependencies, LSTM networks use a gating mechanism to limit the number of previous steps that can affect the current step.

NLG Models and Methodologies

- Transformer.
  - This neural network architecture is able to learn long-range dependencies in language and can create sentences from the meanings of words.
  - Transformer is related to AI. It was developed by OpenAI, a nonprofit AI research company in San Francisco.
  - Transformer includes two encoders: one for processing inputs of any length and another to output the generated sentences.

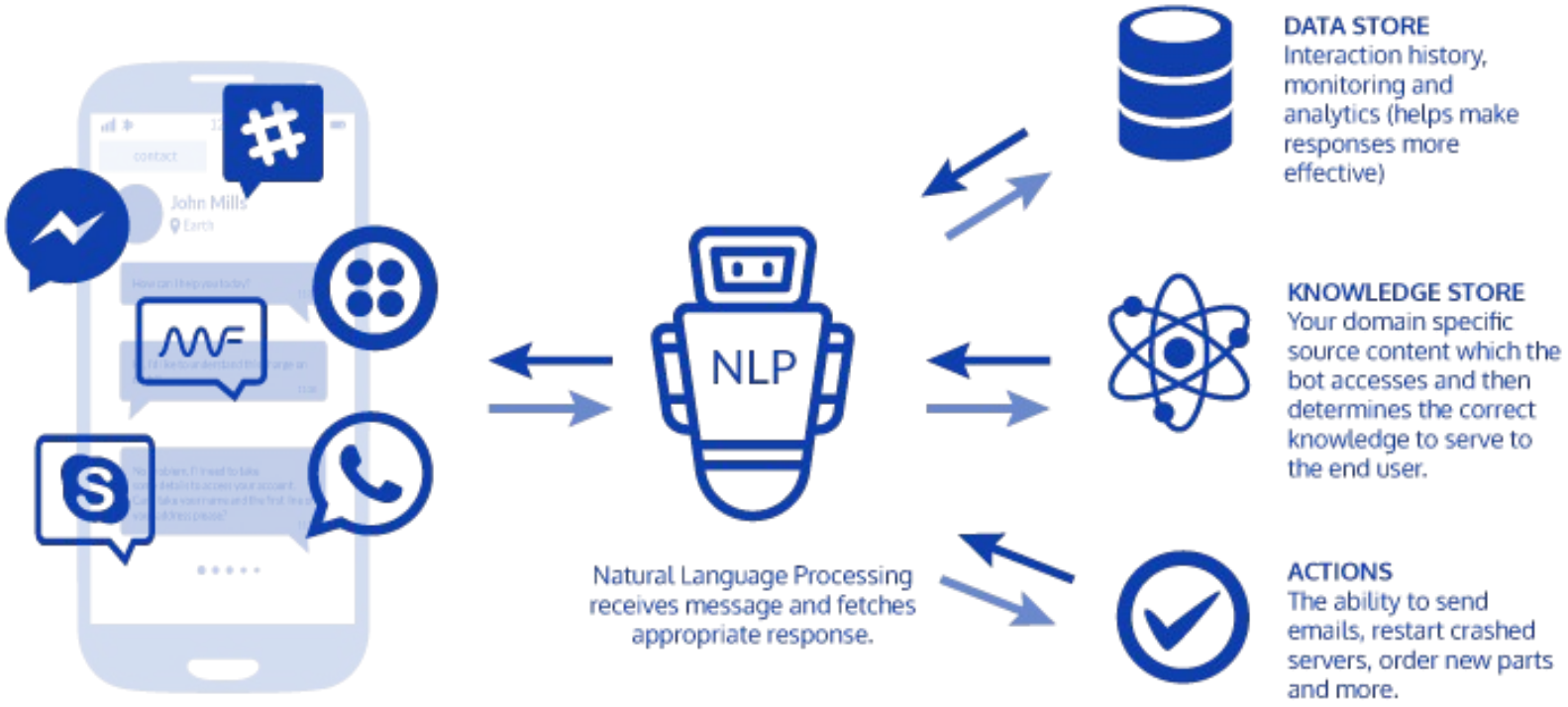
Natural Language Interaction

- Natural Language Interaction (NLI) is the convergence of a diverse set of natural language principles that enables people to interact with any connected device or service in a humanlike way.
- Increasingly known as conversational AI, NLI allows technology to understand complex sentences, containing multiple pieces of information and more than one request.
- It can then react accordingly, creating value and enhancing the user experience.

Natural Language Interaction

- Natural language interaction technology takes natural language processing (NLP) and natural language understanding (NLU) to the next level.
- It allows enterprises to create advanced dialogue systems that utilise memory, personal preferences and contextual understanding to deliver a proactive natural language interface.

Summary


NLP
Natural Language Processing receives message and fetches appropriate response.
DATA STORE
Interaction history, monitoring and analytics (helps make responses more effective)
KNOWLEDGE STORE
Your domain specific source content which the bot accesses and then determines the correct knowledge to serve to the end user.
ACTIONS
The ability to send emails, restart crashed servers, order new parts and more.

Computational Linguistics

- Computational linguistics (CL) is the application of computer science to the analysis and comprehension of written and spoken language.
- As an interdisciplinary field, CL combines linguistics with computer science and artificial intelligence (AI) and is concerned with understanding language from a computational perspective.
- Computers that are linguistically competent help facilitate human interaction with machines and software.

Computational Linguistics

- Computational linguistics is used in tools like instant machine translation, speech recognition systems, text-to-speech synthesizers, interactive voice response systems, search engines, text editors and language instruction materials.
- Typically, computational linguists are employed in universities, governmental research labs or large enterprises.

Computational Linguistics

- A computational linguist is required to have expertise in machine learning (ML), deep learning, AI, cognitive computing and neuroscience.
- Individuals pursing a job as a linguist generally need a master's or doctoral degree in a computer science-related field or a bachelor's degree with work experience developing natural language software.
- The term computational linguistics is also very closely linked to NLP, and these two terms are often used interchangeably.

Computational Linguistics: Applns

Machine translation

Application clustering

Sentiment analysis

Chatbots

Knowledge extraction

Natural language interfaces

Content filtering

Computational Linguistics: Goals

- Create grammatical and semantic frameworks for characterizing languages.
- Translate text from one language to another.
- Retrieve text that relates to a specific topic.
- Analyze text or spoken language for context, sentiment or other affective qualities.
- Answer questions, including those that require inference and descriptive or discursive answers.
- Summarize text.
- Build dialogue agents capable of completing complex tasks such as making a purchase, planning a trip or scheduling maintenance.
- Create chatbots capable of passing the Turing Test.

Computational Linguistics: Approaches

- The corpus-based approach, which is based on the language as it is practically used.
- The comprehension approach, which enables the NLP engine to interpret naturally written commands in a simple rule-governed environment.
- The developmental approach, which adopts the language acquisition strategy of a child -- acquiring language over time. The developmental process has a statistical approach to studying language and does not take grammatical structure into account.

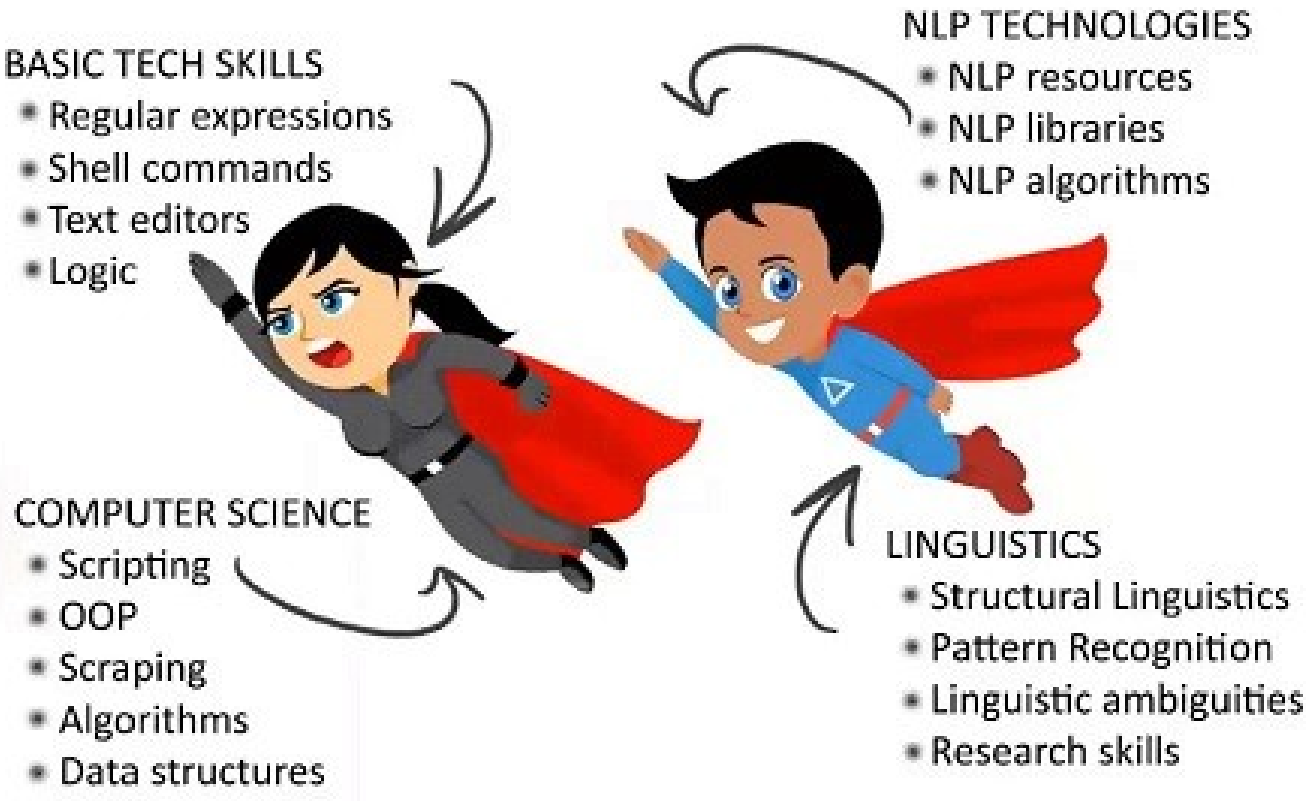
Computational Linguistics: Approaches

- The structural approach, which takes a theoretical approach to the structure of a language.
- This approach uses large samples of a language run through CL models so it can gain a better understanding of the underlying language structures.
- The production approach, which focuses on a CL model to produce text. This has been done in a number of ways, including the construction of algorithms that produce text based on example texts from humans.

Computational Linguistics: Approaches

- The text-based interactive approach, in which text from a human is used to generate a response by an algorithm.
- A computer is able to recognize different patterns and reply based on user input and specified keywords.
- The speech-based interactive approach, which works similarly to the text-based approach, but the user input is made through speech recognition.
- The user's speech input is recognized as sound waves and is interpreted as patterns by the CL system.

Computational Linguists



Thank you

This presentation is created using LibreOffice Impress 7.4.1.2, can be used freely as per GNU General Public License

@mitu_skillologies

@mITuSkillologies

@mitu_group

@mitu-skillologies

@MITUSkillologies

kaggle

@mituskillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

@mituskillologies

contact@mitu.co.in

tushar@tusharkute.com