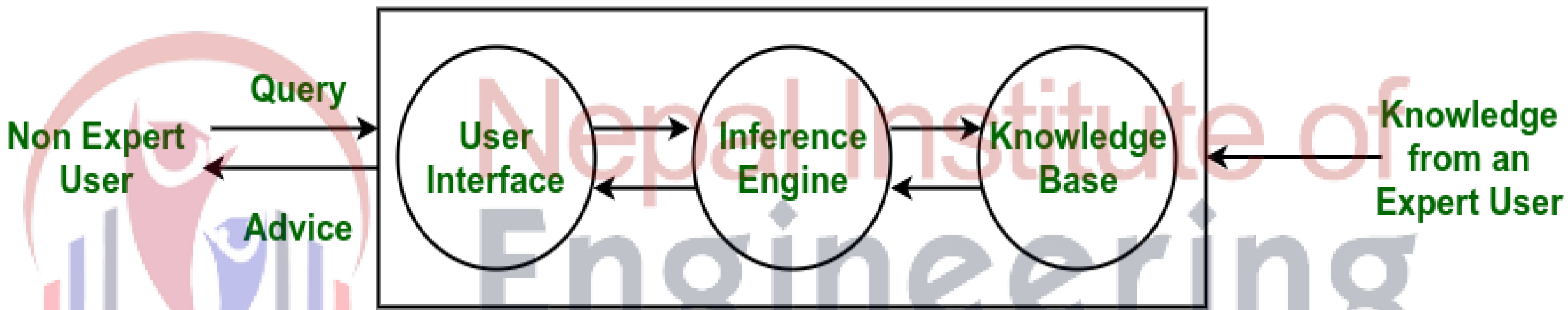
Expert System and Natural Language Processing

Expert System

- The expert system is a part of AI, and the first ES was developed in the year 1970, which was the first successful approach of artificial intelligence.
- It solves the most complex issue as an expert by extracting the knowledge stored in its knowledge base.
- The system helps in decision making for complex problems using **both facts and heuristics like a human expert**.
- It is called so because it contains the expert knowledge of a specific domain and can solve any complex problem of that particular domain.
- These systems are designed for a specific domain, such as **medicine, science,** etc.
- The performance of an expert system is based on the expert's knowledge stored in its knowledge base.

Nepal Institute of Engineering

- The more knowledge stored in the KB, the more that system improves its performance.
- One of the common examples of an ES is a suggestion of spelling errors while typing in the Google search box.
- Components of Expert System



Knowledge Base

- The knowledge base represents facts and rules.
- It consists of knowledge in a particular domain as well as rules to solve a problem, procedures and intrinsic data relevant to the domain.

- **Components of Knowledge Base**
  - **Factual Knowledge:** The knowledge which is based on facts and accepted by knowledge engineers comes under factual knowledge.
  - **Heuristic Knowledge:** This knowledge is based on practice, the ability to guess, evaluation, and experiences.
- **Knowledge Acquisitions:** It is the process of extracting, organizing, and structuring the domain knowledge, specifying the rules to acquire the knowledge from various experts, and store that knowledge into the knowledge base.

Nepal Institute of Engineering

Inference Engine

- The function of the inference engine is to fetch the relevant knowledge from the knowledge base, interpret it and to find a solution relevant to the user's problem.
- The inference engine acquires the rules from its knowledge base and applies them to the known facts to infer new facts.
- Inference engines can also include an explanation and debugging abilities.
- Inference engine uses the below modes to derive the solutions:
  - **Forward Chaining:** It starts from the known facts and rules, and applies the inference rules to add their conclusion to the known facts.
  - **Backward Chaining:** It is a backward reasoning method that starts from the goal and works backward to prove the known facts.

Nepal Institute of Engineering

Knowledge Acquisition and Learning Module

- The function of this component is to allow the expert system to acquire more and more knowledge from various sources and store it in the knowledge base.

User Interface

- This module makes it possible for a non-expert user to interact with the expert system and find a solution to the problem.

Explanation Module

- This module helps the expert system to give the user an explanation about how the expert system reached a particular conclusion.

Nepal Institute of Engineering

S.NO	Procedural Knowledge	Declarative Knowledge
1.	It is also known as Interpretive knowledge.	It is also known as Descriptive knowledge.
2.	Procedural Knowledge means how a particular thing can be accomplished.	While Declarative Knowledge means basic knowledge about something.
3.	Procedural Knowledge is generally not used means it is not more popular.	Declarative Knowledge is more popular.
4.	Procedural Knowledge can’t be easily communicate.	Declarative Knowledge can be easily communicate.

Nepal Institute of Engineering

5.	Procedural Knowledge is generally process oriented in nature.	Declarative Knowledge is data oriented in nature.
6.	In Procedural Knowledge debugging and validation is not easy.	In Declarative Knowledge debugging and validation is easy.
7.	Procedural Knowledge is less effective in competitive programming.	Declarative Knowledge is more effective in competitive programming.

Nepal Institute of Engineering

Development of Expert System

- Here, we will explain the working of an expert system by taking an example of MYCIN ES.

Below are some steps to build an MYCIN

- Firstly, ES should be fed with expert knowledge.
- In the case of MYCIN, human experts specialized in the medical field of bacterial infection, provide information about the causes, symptoms, and other knowledge in that domain.
- The KB of the MYCIN is updated successfully. In order to test it, the doctor provides a new problem to it.
- The problem is to identify the presence of the bacteria by inputting the details of a patient, including the symptoms, current condition, and medical history.

- The ES will need a questionnaire to be filled by the patient to know the general information about the patient, such as gender, age, etc.
- Now the system has collected all the information, so it will find the solution for the problem by applying if-then rules using the inference engine and using the facts stored within the KB.
- In the end, it will provide a response to the patient by using the user interface.

Participants in the development of Expert System

- There are three primary participants in the building of Expert System:
- **Expert:** The success of an ES much depends on the knowledge provided by human experts. These experts are those persons who are specialized in that specific domain.

- **Knowledge Engineer:** Knowledge engineer is the person who gathers the knowledge from the domain experts and then codifies that knowledge to the system according to the formalism.
- **End-User:** This is a particular person or a group of people who may not be experts, and working on the expert system needs the solution or advice for his queries, which are complex.

Nepal Institute of Engineering

Natural Language Processing

- Natural Language Processing (NLP) is a branch of computer science and artificial intelligence dedicated to enabling computers to comprehend human language.
- Leveraging computational linguistics alongside statistical, machine learning, and deep learning models, NLP empowers computers to analyze and interpret text or voice data, capturing nuances such as intentions and emotions.
- This technology underpins various applications like translation, voice recognition, summarization, and chatbots, enhancing efficiency and productivity in businesses.
- However, NLP faces challenges due to the complexities of human language, including homonyms, sarcasm, idioms, and variations in structure, necessitating meticulous training of algorithms to accurately understand and process linguistic nuances.

Nepal Institute of Engineering

Some common techniques used in NLP include

- Tokenization: the process of breaking text into individual words or phrases.
- Part-of-speech tagging: the process of labeling each word in a sentence with its grammatical part of speech.
- Named entity recognition: the process of identifying and categorizing named entities, such as people, places, and organizations, in text.
- Sentiment analysis: the process of determining the sentiment of a piece of text, such as whether it is positive, negative, or neutral.
- Machine translation: the process of automatically translating text from one language to another.
- Text classification: the process of categorizing text into predefined categories or topics.

Nepal Institute of Engineering

Working of Natural Language Processing (NLP)

- This can include tasks such as language understanding, language generation, and language interaction.

Speech Recognition:

- First, the computer must take natural language and convert it into machine-readable language. This is what speech recognition or speech-to-text does. This is the first step of NLU.
- Hidden Markov Models (HMMs) are used in the majority of voice recognition systems nowadays. These are statistical models that use mathematical calculations to determine what you said in order to convert your speech to text.
- HMMs do this by listening to you talk, breaking it down into small units (typically 10-20 milliseconds), and comparing it to pre-recorded speech to figure out which phoneme you uttered in each unit (a phoneme is the smallest unit of speech). The program then examines the sequence of phonemes and uses statistical analysis to determine the most likely words and sentences you were speaking.

Nepal Institute of Engineering

Natural Language Understanding (NLU):

- The next and hardest step of NLP is the understanding part.
- First, the computer must comprehend the meaning of each word. It tries to figure out whether the word is a noun or a verb, whether it's in the past or present tense, and so on. This is called Part-of-Speech tagging (POS).
- A lexicon (a vocabulary) and a set of grammatical rules are also built into NLP systems. The most difficult part of NLP is understanding.
- The machine should be able to grasp what you said by the conclusion of the process. There are several challenges in accomplishing this when considering problems such as words having several meanings (polysemy) or different words having similar meanings (synonymy), but developers encode rules into their NLU systems and train them to learn to apply the rules correctly.

Nepal Institute of Engineering

Natural Language Generation (NLG)

- NLG is much simpler to accomplish. NLG converts a computer's machine-readable language into text and can also convert that text into audible speech using text-to-speech technology.
- First, the NLP system identifies what data should be converted to text. If you asked the computer a question about the weather, it most likely did an online search to find your answer, and from there it decides that the temperature, wind, and humidity are the factors that should be read aloud to you.
- Then, it organizes the structure of how it's going to say it. This is similar to NLU except backward. NLG system can construct full sentences using a lexicon and a set of grammar rules.
- Finally, text-to-speech takes over. The text-to-speech engine uses a prosody model to evaluate the text and identify breaks, duration, and pitch. The engine then combines all the recorded phonemes into one cohesive string of speech using a speech database.

Nepal Institute of Engineering

Steps of NLP

Lexical Analysis and Morphological

- The first phase of NLP is the Lexical Analysis. This phase scans the source code as a stream of characters and converts it into meaningful **lexemes**. It divides the whole text into paragraphs, sentences, and words.

Syntactic Analysis (Parsing)

- Syntactic Analysis is used to check grammar, word arrangements, and shows the relationship among the words.
- **Example:** Agra goes to the Poonam
- In the real world, Agra goes to the Poonam, does not make any sense, so this sentence is rejected by the Syntactic analyzer.

Semantic Analysis

- Semantic analysis is concerned with the meaning representation. It mainly focuses on the literal meaning of words, phrases, and sentences.

Nepal Institute of Engineering

Discourse Integration

- Discourse Integration depends upon the sentences that proceeds it and also invokes the meaning of the sentences that follow it.

Pragmatic Analysis

- Pragmatic is the fifth and last phase of NLP. It helps you to discover the intended effect by applying a set of rules that characterize cooperative dialogues.

Applications of Natural Language Processing (NLP)

- **Spam Filters**
  - One of the most irritating things about email is spam. Gmail uses natural language processing (NLP) to discern which emails are legitimate and which are spam.
  - These spam filters look at the text in all the emails you receive and try to figure out what it means to see if it's spam or not.

Nepal Institute of Engineering

- **Algorithmic Trading**
  - Algorithmic trading is used for predicting stock market conditions.
  - Using NLP, this technology examines news headlines about companies and stocks and attempts to comprehend their meaning in order to determine if you should buy, sell, or hold certain stocks.
- **Questions Answering**
  - NLP can be seen in action by using Google Search or Siri Services.
  - A major use of NLP is to make search engines understand the meaning of what we are asking and generate natural language in return to give us the answers.
- **Summarizing Information**
  - On the internet, there is a lot of information, and a lot of it comes in the form of long documents or articles.

Nepal Institute of Engineering

- NLP is used to decipher the meaning of the data and then provides shorter summaries of the data so that humans can comprehend it more quickly.

NLP Challenges

- **Ambiguity**: Human language is inherently ambiguous, with words and phrases often having multiple meanings or interpretations.
- **Context Sensitivity**: The meaning of a word or phrase can vary depending on the context in which it is used, making it challenging for NLP systems to accurately interpret.
- **Syntax and Grammar**: Variations in sentence structure, grammar rules, and syntactic nuances pose challenges for parsing and understanding natural language.
- **Named Entity Recognition (NER)**: Identifying and categorizing named entities (e.g., people, organizations, locations) within text accurately is crucial but challenging due to variations and ambiguity.

- **Coreference Resolution**: Resolving references to entities mentioned earlier in the text (e.g., pronouns) requires understanding the context and relationships between different parts of the text.
- **Lack of Training Data**: Building robust NLP models often requires large amounts of labeled training data, which may be scarce or costly to obtain for certain languages or domains.
- **Domain Adaptation**: NLP models trained on generic data may not perform well in specialized domains (e.g., medical or legal texts) due to differences in terminology and language usage.
- **Sarcasm, Irony, and Figurative Language**: Recognizing and interpreting sarcasm, irony, metaphors, and other forms of figurative language requires understanding nuanced linguistic cues.
- **Multilingualism and Code-Switching**: Handling multilingual text and code-switching between languages presents challenges for NLP systems, especially in regions where multiple languages are commonly used.

Nepal Institute of Engineering

- **Bias and Fairness**: NLP models can perpetuate biases present in the training data, leading to unfair or discriminatory outcomes, highlighting the importance of addressing bias and promoting fairness in NLP applications.

Machine vision

- also known as computer vision, is a branch of artificial intelligence and computer science focused on enabling machines to interpret and understand visual information from digital images or video.
- It involves developing algorithms and techniques that allow computers to analyze and extract meaningful insights from visual data, similar to how humans perceive and interpret the world through vision.

Key components of machine vision include:

- **Image Acquisition**: Obtaining digital images or video streams using cameras or other sensors.

Nepal Institute of Engineering

- **Pre-processing**: Enhancing the quality of raw image data through techniques such as noise reduction, image enhancement, and geometric corrections.
- **Feature Extraction**: Identifying and extracting relevant features from the pre-processed images, such as shapes, textures, colors, or patterns.
- **Object Detection and Recognition**: Locating objects within images and recognizing them based on predefined categories.
- **Decision Making**: Using the extracted information to make decisions or take actions, such as automated quality control in manufacturing or autonomous navigation for robots.

Robotics

- Robotics in machine learning involves the integration of artificial intelligence and machine learning techniques into robotic systems.
  - **Data Acquisition**: Robots collect data through sensors such as cameras, LIDAR, or joint encoders, capturing information about their environment and actions.
  - **Feature Extraction**: Relevant features are extracted from the raw sensor data, such as object positions, distances, or shapes, to represent the environment and objects within it.
  - **Learning Algorithms**: Machine learning algorithms, including supervised, unsupervised, and reinforcement learning, are employed to process the extracted features and learn patterns or behaviors from the data.

- **Decision Making**: Based on the learned models and environmental inputs, robots make decisions autonomously, such as navigating obstacles, manipulating objects, or performing tasks efficiently and safely.
- **Feedback Loop**: Robots continuously adapt and improve their behavior through feedback loops, where they learn from their experiences and refine their models over time, enhancing their capabilities and performance.

Nepal Institute of Engineering

What is the name for information sent from robot sensors to robot controllers?

a) temperature
b) pressure
c) feedback
d) signal

Decision support programs are designed to help managers make __________

a) budget projections
b) visual presentations
c) business decisions
d) vacation schedules

What is the field of Natural Language Processing (NLP)?

a) Computer Science
b) Artificial Intelligence
c) Linguistics
d) All of the mentioned

Nepal Institute of Engineering

- What is the main challenge/s of NLP?
  a) Handling Ambiguity of Sentences
  b) Handling Tokenization
  c) Handling POS-Tagging
  d) All of the mentioned
- Choose form the following areas where NLP can be useful.
  a) Automatic Text Summarization
  b) Automatic Question-Answering Systems
  c) Information Retrieval
  d) All of the mentioned
- What is Coreference Resolution?
  a) Anaphora Resolution
  b) Given a sentence or larger chunk of text, determine which words (“mentions”) refer to the same objects (“entities”)
  c) All of the mentioned
  d) None of the mentioned

Nepal Institute of Engineering

- What is Machine Translation?
  a) Converts one human language to another
  b) Converts human language to machine language
  c) Converts any human language to English
  d) Converts Machine language to human language
- What is Morphological Segmentation?
  a) Does Discourse Analysis
  b) Separate words into individual morphemes and identify the class of the morphemes
  c) Is an extension of propositional logic
  d) None of the mentioned
- Which of the following is an advantage of using an expert system development tool?
  a) imposed structure
  b) knowledge engineering assistance
  c) rapid prototyping
  d) all of the mentioned

Nepal Institute of Engineering

- In LISP, the function returns t if <integer> is even and nil otherwise ____________

  a) (evenp <integer>)

  b) (even <integer>)

  c) (numeven <integer>)

  d) (numnevenp <integer>)