ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
UNIVERSITY OF WEST ATTICA

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
UNIVERSITY OF THESSALY

Large Language Models versus Natural Language Understanding and Generation

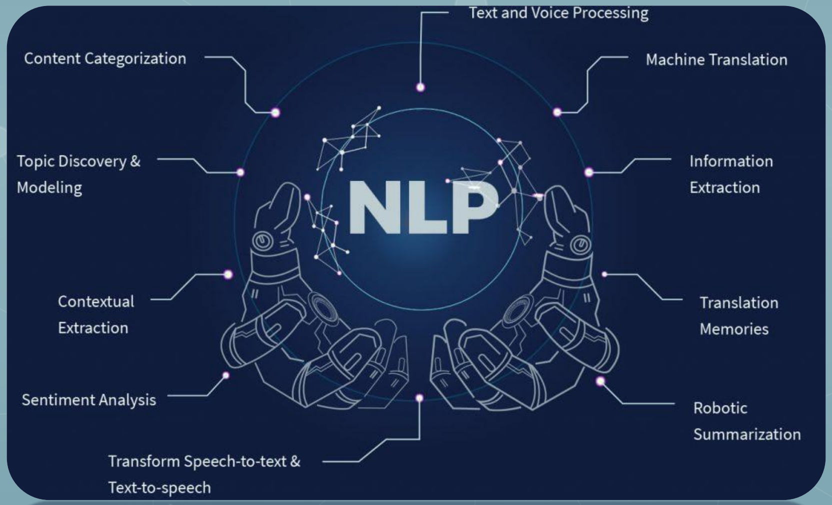
Nikitas Karanikolas
Dept. of Informatics and Computer Eng.
UNIWA

Eirini Manga
Dept. of Informatics and Computer Eng.
UNIWA

Nikoletta Samaridi
Dept. of Informatics and Computer Eng.
UNIWA

Eleni Tousidou
Dept. of Electrical and Computer Eng.
UTH

Michael Vassilakopoulos
Dept. of Electrical and Computer Eng.
UTH



November, 2023

Introduction

- How do people learn a foreign language?
  - "Grammar – Translation" vs "Communicative approach"

- The same trends have been applied to the way machinery can be "educated" to comprehend the unfamiliar, human language

- The "rule based" **Natural Language Generation** algorithms combined with **Natural Language Understanding** ones, on one hand, and the "text based" **Large Language Models**, on the other, are two respective alternatives for human language understanding and generation

Definitions

- **Natural Language Processing** (**NLP**): a branch of Artificial Intelligence (AI) which enables computers to analyse and synthesize natural (human) language and speech
    - **Natural Language Generation** (**NLG**): a software process that produces natural written or spoken text from structured and unstructured data
    - **Natural Language Understanding** (**NLU**): a software process which analyses what natural language means (rather than simply what individual words say)
        - NLG and NLU are usually combined
- **Large Language Model** (**LLM**): a software process that uses deep learning techniques and massively large data sets to learn how to understand and generate natural text

Lecture targets

- Present applications of NLP (supported by LLMs and/or NLU/NLG systems)
- Present the two alternative approaches for such applications, LLMs on the one hand and NLU/NLG systems on the other and their functioning
- Highlight the capabilities, strengths and weaknesses of these approaches
- Discuss arising challenges and possible future directions
- In other words, contribute to a deeper comprehension of the evolving landscape of AI and human-computer interaction

Delving into Natural Language Processing Applications

- **Evolution in AI:** NLP has redefine engagement, understanding, and creation of text-based information.
- **Insightful Adaptability:** Trained on vast textual data, NLP can offer adaptability across diverse applications.
- **Innovation Catalyst:** Fueling innovation across multiple fields by showcasing transformative potential.

Text Generation

- **Revolutionizing Industries:** Automating processes across diverse sectors.
- **Content Creation:** product descriptions, ads, promotions, news articles and blogs.
- **Financial Reporting Automation:** Analyzing market data to automatically generate financial reports.
- **Educational Contribution:**
  - Facilitating e-learning by producing quizzes, study materials, and explanations.
  - Enhancing engagement and personalized learning experiences.
- **Creative Writing Support:** Contributing to narratives, short stories, and poetry.
- **Chatbot & Virtual Assistant Roles:**
  - Empowering chatbots with natural language responses for customer inquiries.
  - Assisting users in tasks, providing information, and troubleshooting issues.

Text Summarization

- **News Organizations:** Automatic generation of brief news article summaries for quick comprehension.
- **Researchers:** Condensing academic papers for rapid identification of key findings.
- **Content Aggregation Platforms and Websites:** Creating digests aiding user article selection based on interests.
- **Finance Sector:** Summarizing financial reports for swift evaluation by investors and analysts.
- **Healthcare Professionals:** Summarizing medical conversations, records, and research papers for streamlined tasks.
- **Social Media Insights:** Summarizing discussions and reviews for understanding public sentiment and reactions.
- **Market Research:** Distilling consumer feedback and reviews for prompt trend identification.
- **Educational Support:** Creating concise educational material summaries for student comprehension.

Question Answering

- **Company Chatbots:** Addressing customer inquiries, providing information, and troubleshooting.
- **Voice-Activated Assistants:** assist users and execute tasks.
- **E-commerce:** Enabling proactive information seeking for informed purchasing decisions.
- **E-Learning Platforms:** Addressing student queries, providing explanations and solutions.
- **Healthcare:** Assisting healthcare professionals with clinical queries, symptoms, and generating reports.
- **Legal Sector:** Providing insights into case law, statutes, and regulations for legal research.
- **Technical Support:** Offering guidance and solutions for software and hardware queries.

Machine Translation

- **Translation Services:** Text, documents, websites, and spoken language.
- **Content Creation and Marketing:** Translating content into multiple languages.
- **Government and Diplomatic Use:** Translating official documents, treaties, and diplomatic communications.
- **News Agencies:** Swiftly translating news articles to provide timely and accurate information in multiple languages.
- **E-commerce and Retail:** Translating product descriptions, reviews, and checkout processes.
- **International Business Communications:** Translating contracts, emails, and documentation for negotiations and collaborations.
- **Travel Industry Integration:** Providing multilingual information for tourists through apps and websites.
- **Customer Support:** Facilitating communication between international customers and support teams.
- **Language Learning Platforms:** Offering language courses and translation assistance to enhance language skills.

Text Classification

- **Email Sorting:** Distinguishing spam from genuine emails by evaluating content and sender details.
- **E-commerce Taxonomy:** Organizing products and multi-level taxonomy trees.
- **Social Media for Sentiment Analysis:** Assessing user-generated content for sentiments - positive, negative, or neutral.
- **Personalized Recommendations:** Tailored suggestions in e-commerce platforms.
- **Healthcare Data:** Classifying medical records and patient data to enhance organization and support research.
- **Legal Document Classification:** Simplifying document management for legal firms by classifying contracts and case files.
- **Fraud Detection in Finance:** Identifying suspicious transactions and user behavior for fraud detection.

Sentiment Analysis

- **Social Media and Customer Feedback:** Assessing sentiments to manage reputation and enhance product quality.
- **Diverse Data Sources:** Processing surveys, forms, and reviews to gauge customer satisfaction.
- **Financial Sector :** Analyzing news, reports, and social media for market sentiment.
- **Polling and Political Forecasting:** Gauging public sentiment for informed decision-making in politics.
- **Entertainment Industry:** Evaluating audience reactions to movies and TV shows.
- **Product Launches:** Assessing public sentiment to evaluate market reception.
- **Healthcare:** Understanding patient sentiment to drive improvements in patient care.
- **Social Media:** Measuring campaign effectiveness and tracking brand trends.
- **News Content Tailoring:** Assessing reader reactions to tailor preferences.
- **Customer Support Insights:** Identifying dissatisfied customers.
- **Travel Services:** Analyzing reviews to enhance services and increase bookings.

Chatbots - Virtual Assistants

- **E-commerce & Online Services:** Chatbots can assist customers, provide product details, and resolve issues instantly.
- **Banking & Finance:** Virtual assistants aid with inquiries and financial guidance.
- **Travel:** Chatbots facilitate bookings, offer travel information.
- **Healthcare:** Virtual assistants provide medical information and schedule appointments.
- **Language Learning :** Virtual tutors assist in practicing languages.
- **Human Resources:** Chatbots aid with HR-related inquiries and company policies.
- **Online Retail:** Chatbots offer personalized product suggestions.
- **Business Reception:** Virtual receptionists manage calls and provide information.
- **Education:** Chatbots can assist with course-related queries and assignments.
- **News Updates:** Chatbots deliver tailored news content and updates to users.
- **Smart Home Devices:** Voice-activated assistants are used for inquiries, smart device control, and information provision.

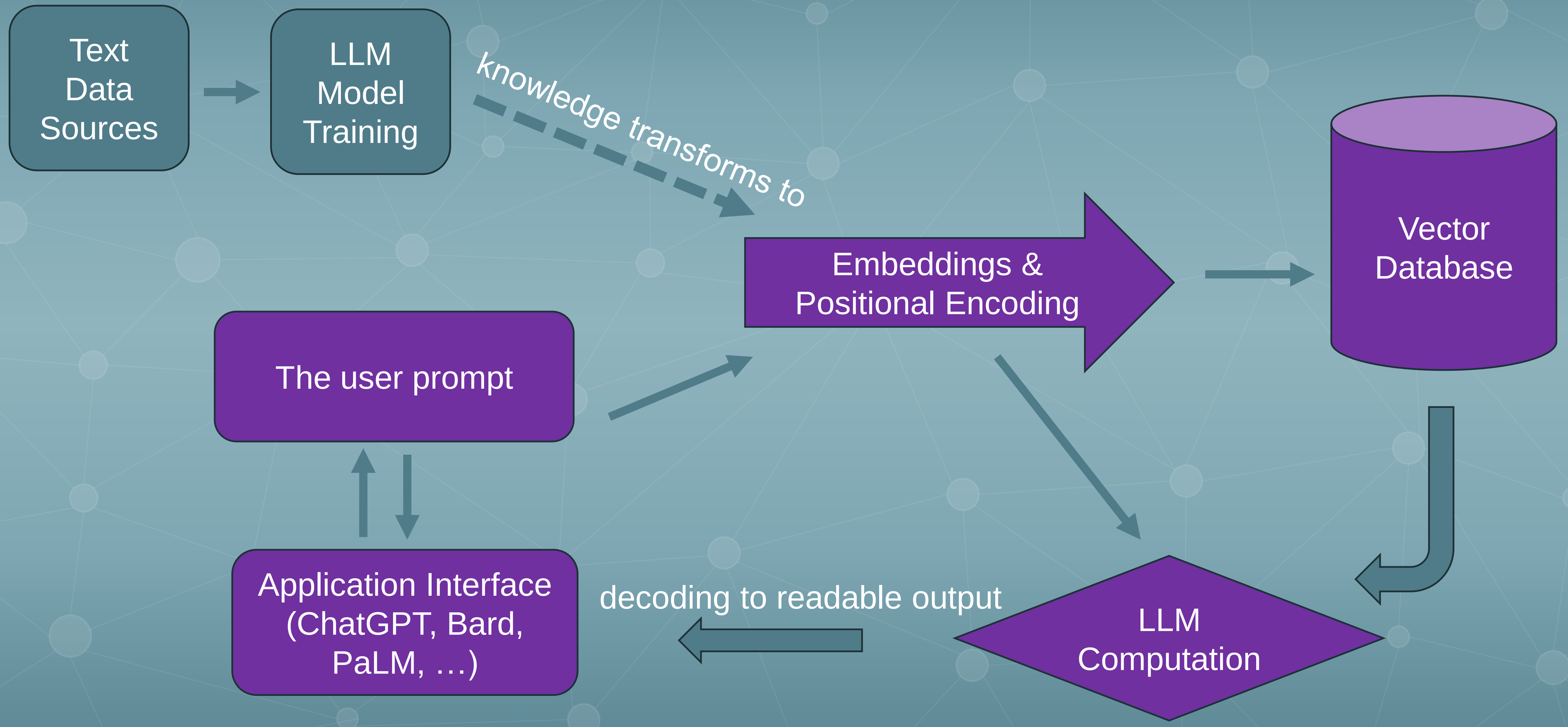
Image Captioning

- **Social Media Platforms:** Use NLP to create image captions, benefiting user engagement and accessibility, particularly for the visually impaired.
- **E-commerce:** Image captioning in catalogs aids product management and improves search experiences.
- **Education:** Educational platforms use image captions to explain visual content, assisting students in understanding complex visuals.
- **Healthcare:** Image captioning assists in describing medical images, aiding in diagnostics and conveying findings among medical professionals.
- **Image Search and Retrieval:** Image captioning in search engines helps users find images by describing them in text queries.
- **Content Moderation:** Social media platforms use image captioning to identify inappropriate or harmful content, ensuring a safer online environment.
- **News Content:** News agencies utilize image captioning to enhance the accessibility of news articles for readers.

LLMs: Short History

- The process of evolution of language models has gone under four main stages:
  - The first stage was in 1990s, where statistical models on n-gram languages were used both in NLP and Information Retrieval tasks
  - The next stage involved the use of neural networks, such as RNNs, to be replaced later by the more efficient bi-LSTM models, leading to the class of Pre-trained Language Models (PLMs)
  - However, those models' performance excelled when the model's size or the data size was significantly increased
  - This was the opening of today's Large Language Models (LLMs)

Flow inside a Large Language Model

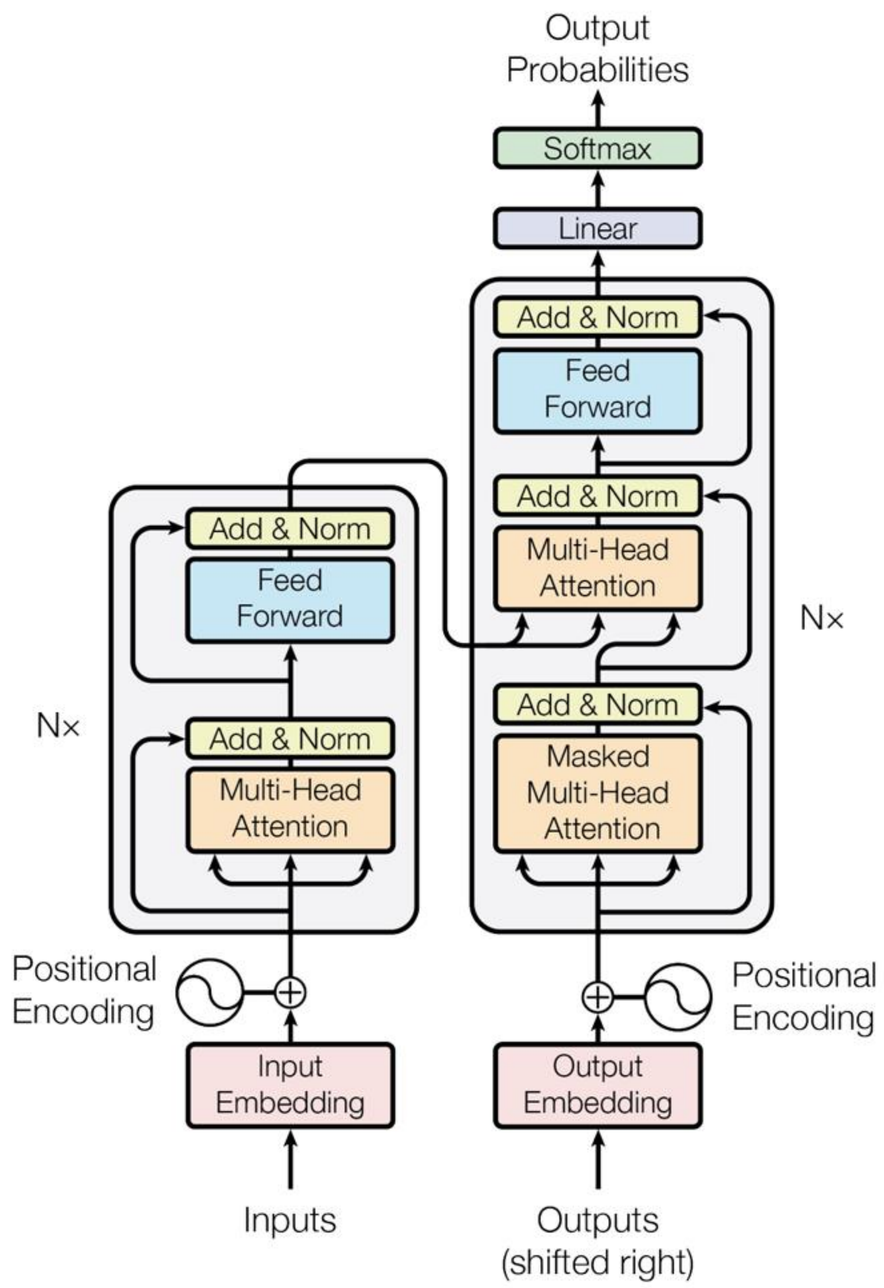


LLMs: Architecture (1)

- Transformers comprise the fundamental building block of LLMs
- They are a specific type of deep neural network designed to handle sequential data effectively, that contains and trains hundreds of billions of parameters
- Their operation is based on a mechanism called self-attention
- Self-attention allows the model to comprehend the role each word has in the input sentence, ignoring its specific position therein
- The produced text achieves superior human-like naturalness with less fine-tuning

Transformer

"Attention is all you need": The transformer architecture, figure taken by the original paper ("Attention Is All You Need" by A. Vaswani et al.), depicting the encoder/decoder architecture, the positional encoding and the attention layer



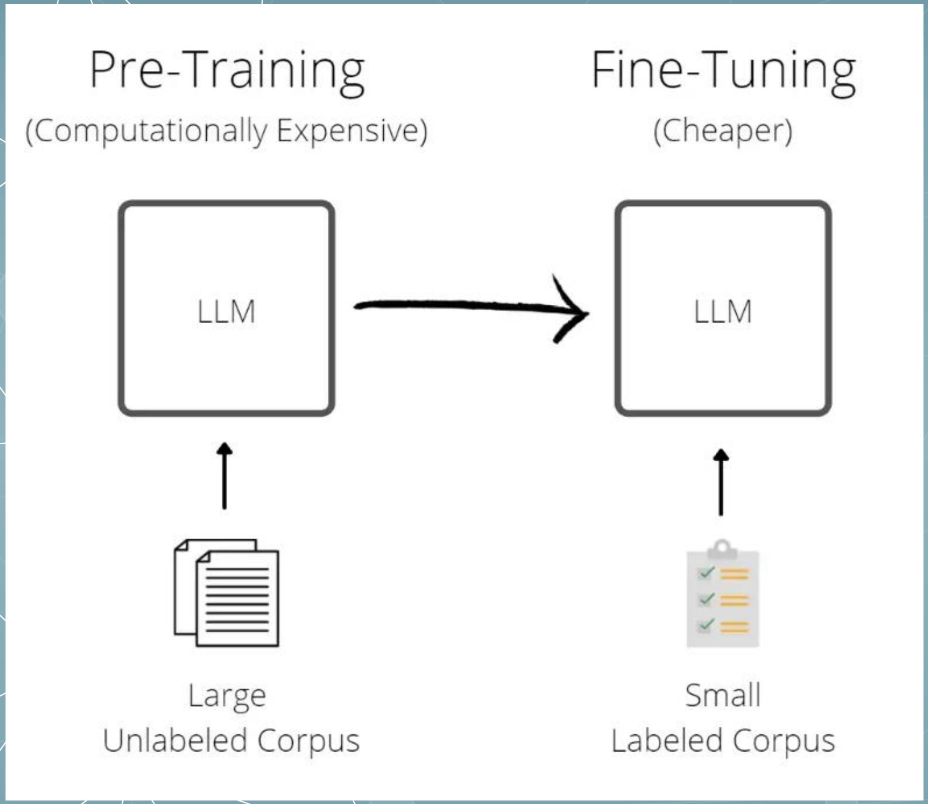
LLMs: Architecture (2)

- Transformers present excellent parallelization ability, leading to excellent scalability features
- Rather than sequentially, the model processes all input text in parallel, significantly reducing its training time
- The different architectural schemes that are used today are Encoder-Decoder, Causal Encoder and prefix Encoder architecture
- These architectures can be further extended and scaled by a Mixture of Experts (MoE), a technique that involves dividing a model into specialized sub-models called experts, activating only one or a few experts for each input token
- The model's performance seems to be increasing significantly either by increasing the number of experts or the number of parameters

LLMs: The different kinds of Training (1)

- A crucial stage in LLMs deployment is their training phase
- The need to get “educated” in the human’s world of knowledge is answered by a long training phase
- During this phase, billions of parameters are being calculated to tune the model’s operation, leading to a Pre-trained model
- After this, tasks such as text generation, language translation and sentiment analysis can be carried out successfully

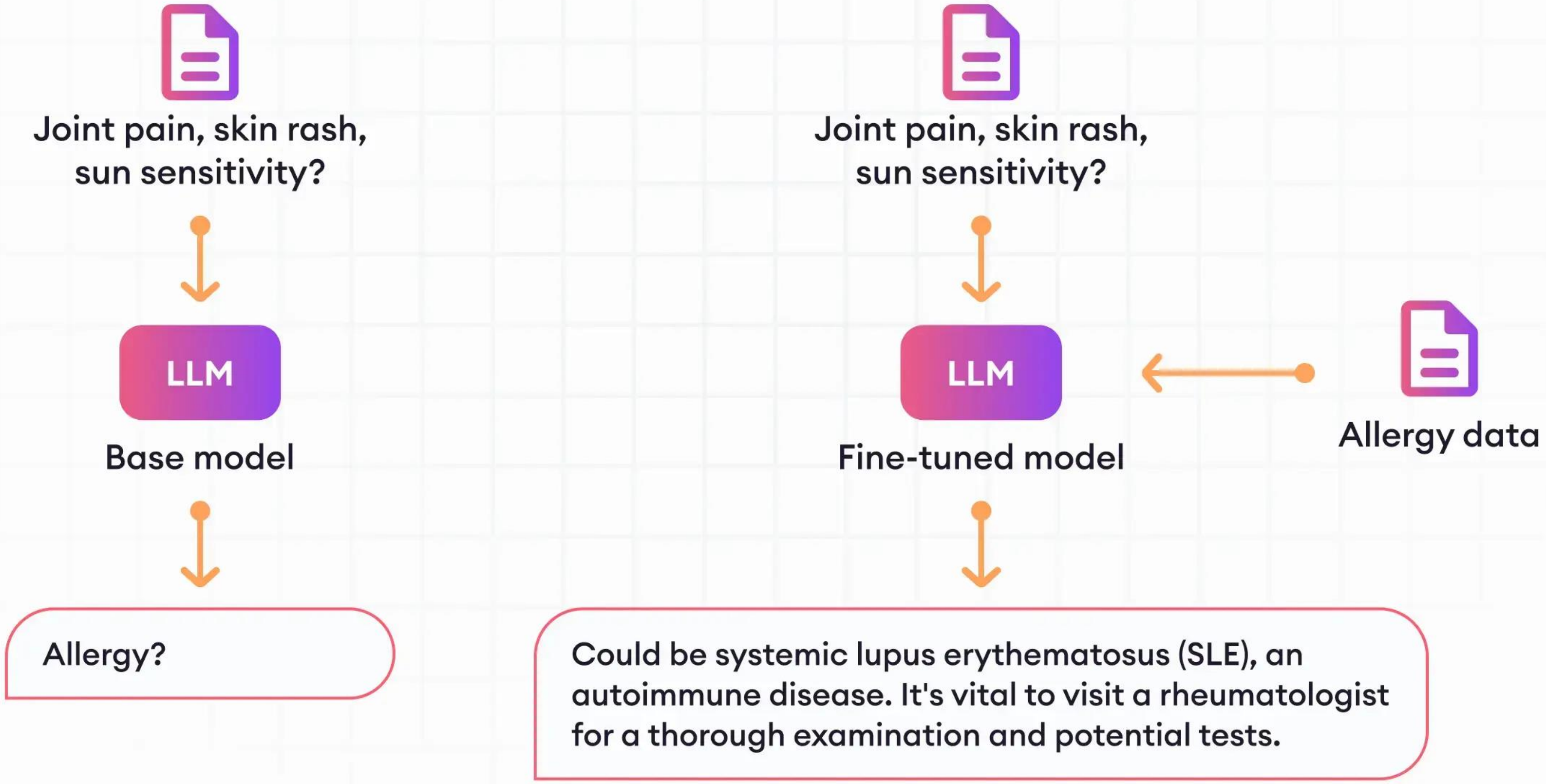
LLMs: The different kinds of Training (2)



- Fine-tuning: a next phase of training which enhances the model’s abilities
  - The pre-trained model gets further “education” on more specific tasks, accommodating its general knowledge with a more refined/specialized knowledge, improving further its performance
  - The produced model excels in tasks like question answering, fraud detection, text classification and others

What does fine-tuning do for the model ?

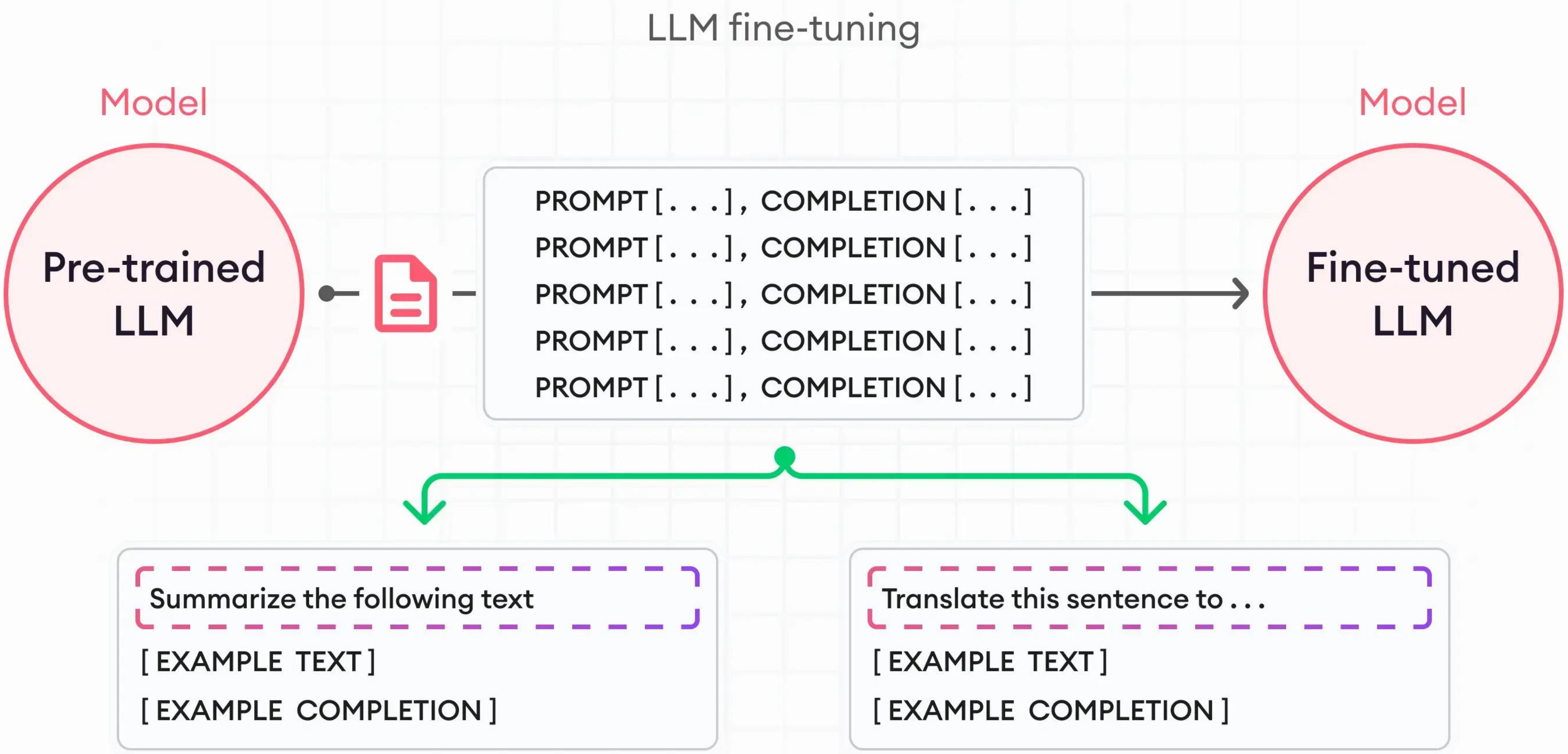
- Lets you add more data into the model than what fits into the prompt
- Gets the model to learn the data rather than just get access to it



LLMs: The different kinds of Training (3)

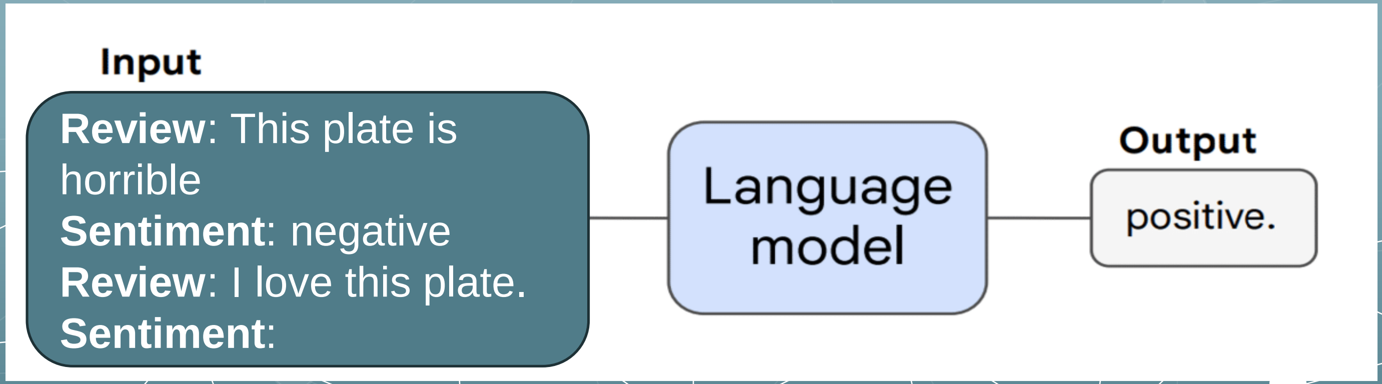
- Instruction-tuning: a kind of fine-tuning operation
  - Models are refined by being given explicit instructions and guidance to generate text that aligns to the users' specific requirements
  - The idea is to train the language model to answer a wide variety of prompts to learn how to respond to new, unseen prompts
  - Involves less processing time because there is no need for training
  - Accompanied with fine tuning, infuses the model with more specialized knowledge and flexibility to respond to a given challenge

Using prompts to fine-tune LLMs with instruction



LLMs: The different kinds of Training (4)

- In-context learning: (also known as few-shot prompting) the model can answer questions without the need to get trained, based only on the specific context given at the prompt
  - The user includes one or more examples of inputs and outputs for a given task before prompting the model to complete the same task
  - Text completion is such a use case



LLMs: The different kinds of Training (5)

- RLHF (Reinforcement Learning with Human Feedback): a technique that instruction-tunes a model based on human evaluation
  - The model provides multiple responses to a given prompt and the responses get ranked by people, from best to worst
  - This results to a reward function for the model to optimize itself
  - Reward functions are difficult to define for generative tasks

LLMs: The different kinds of Training (6)

- Multi-step reasoning: Chain of Thought (CoT) prompt, infuses reasoning by presenting the model examples of a problem and its solution based on a reasoning process
  - When being introduced a new problem, the trained LLM will be able to produce an answer after following a similar reasoning process

LLMs: Abilities emerging by Scalability

- Some of the afore-mentioned processes can be exploited only when the size of models goes beyond a certain limit
- In-context learning and instruction-following emerge in large language models, excelling their performance
- Moreover, the lack of reasoning in LLMs inspired researchers to enhance language models by step-by-step reasoning

NLG

- Systems, models and applications that can produce discourse (written text or oral speech) with meaning not recognized as artificial

- Stages: Planning, Microplanning, Realizing, Presenting (optional)

- Planning - decide what is:
  - the most interesting parts of input data (structured data),
  - the order of communicated ideas/facts,
  - the hidden rhetorical relations [cause, sequence, etc]

- Microplanning:
  - aggregation,
  - anaphora generation,
  - selection of lexical items,
  - decide the syntactic structure of each sentence/phrase

NLG

- Realizing morphologically and orthographically correct discourse:
  - inflection,
  - orthography,
  - ordering of adjectives.

- Presenting (optional stage):
  - Written text: Titles, Emphasis (e.g. bold), punctuation marks
  - Oral speaches: Intonation, Sentence Type (Affirmative, Negative, Imperative, Exclamatory, Interrogative)

- Prerequisite for NLG is NLU

NLU

- NLU is the translation of human/natural language text to some structured form (Predicate Calculus, FRL, KL-ONE, Case Grammars, etc)

- Natural Language Understanding follows (usually) the following tasks:
  - Tokenization,
  - Part of Speech Tagging,
  - Syntactic Analysis,
  - Structural Disambiguation (Resolution of Syntactic Ambiguity),
  - Word Sense Disambiguation,
  - Semantic Representation,
  - Anaphora Resolution,
  - Optional Tasks: Affective Computing, Discourse Analysis, Pragmatics.

Tokenization & POS

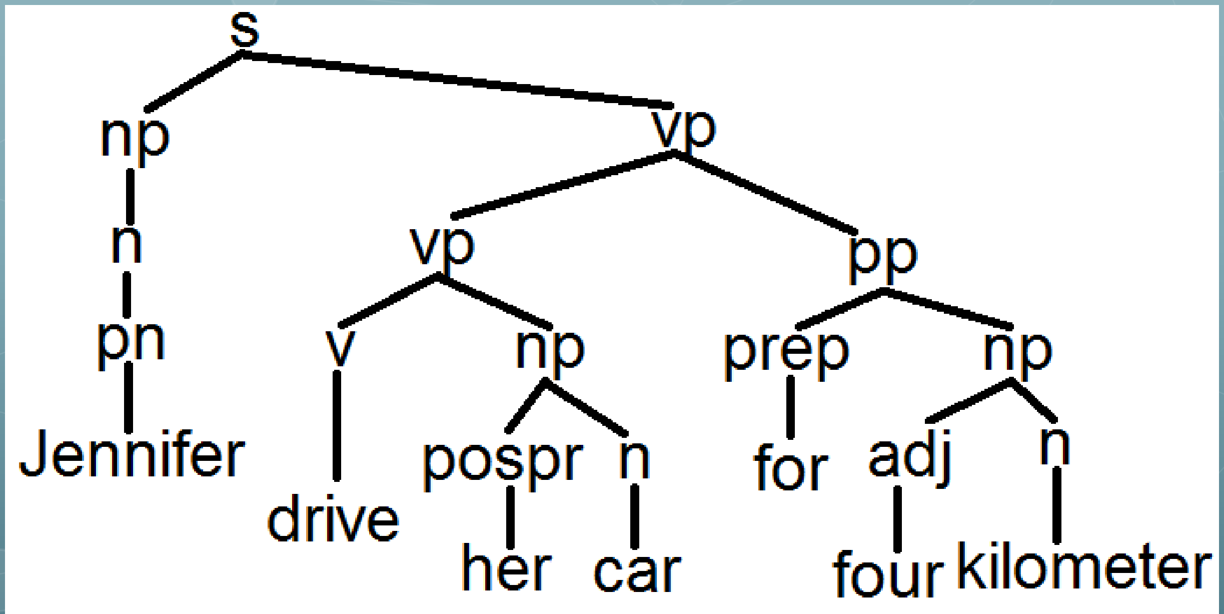
- Tokenization
  - The boundaries between words are identified and the punctuation marks are removed
  - The root form of a word (the lexeme) is identified
- Tokenization example:
  - input: "Jenifer drove her car, for four kilometers."
  - output: {[Jennifer, Jennifer], [drove, drive], ..., [four, four], [kilometers, kilometer]}.

- Part of Speech Tagging (POS)
  - Labeling the tokens with their syntactic category (N, V, ADJ, ADV, PRON, etc).
  - Optionally assign features (mood, tense, voice, number, person for Verbs; sex, number, case for Nouns)

Syntactic Analysis

- Check if a sentence is well formed
- Return the derivation (syntactic / parse tree)

```
s( np(n(pn(Jennifer))),
   vp( vp( v(drive),
           np(pospr(her), n(car)))
       pp( prep(for),
           np(adj(four), n(kilometer)))
   )
)
```



Where:
pn="proper noun",
pospr ="possessive pronoun",
perpr="personal pronoun",
det="determiner",
prep="preposition",
pp="prepositional phrase",
np="noun phrase",
vp="verb phrase",
s="sentence".

Syntactic Ambiguity

- Occurs when the grammar assigns more than one possible syntactic/parse trees to the input sentence.
- A well known attachment ambiguity is the prepositional phrase (PP) attachment ambiguity.

- Example: "Look at the dog with one eye."

```
s( np(n(you))
   vp( v(look),
       pp ( p(at),
            np( det(the),
                n(dog),
                pp(p(with),
                   np(adj(one),n(eye)))
            )
       )
   )
)
```

"Look at the dog using only one of your eyes"

```
s( np(n(you))
   vp( v(look),
       pp (p(at),np(det(the),n(dog))),
       pp (p(with),np(adj(one),n(eye)))
   )
)
```

"Look at the dog that has only one eye"

Word Sense Disambiguation

- There are cases of polysemy words (words having two or more meanings). In such cases it is necessary to have some method for selecting the appropriate meaning.

- For example the Greek word "ουρά" can has the following meanings:

  - the tail of an animal ("η αλεπού έχει όμορφη ουρά"),

  - the queue of people ("Περίμενα στην ουρά 2 ώρες για να πάρω τη βεβαίωση"),

  - overextended ending ("πιάνο με ουρά", "νυφικό με ουρά", etc).

- The selectional restrictions of the thematic roles/slots of the given verb is one way to resolve the problem.

Semantic Representation 1/3

- One of the possible semantic representation of (the meaning of) a sentence is the case grammars.
- They analyze the surface syntactic structure of sentences by studying the combination of deep cases (i.e. semantic roles) required by a specific verb.
- Deep cases can be: Agent, Object, Beneficiary, Location, Instrument, etc.
- Each verb selects a certain number of deep cases (not all) which form its case frame.
  - The verb "give" in English requires Agent (A), Object (O), Beneficiary (B).
  - Example: "Jones (A) gave money (O) to the school (B)".

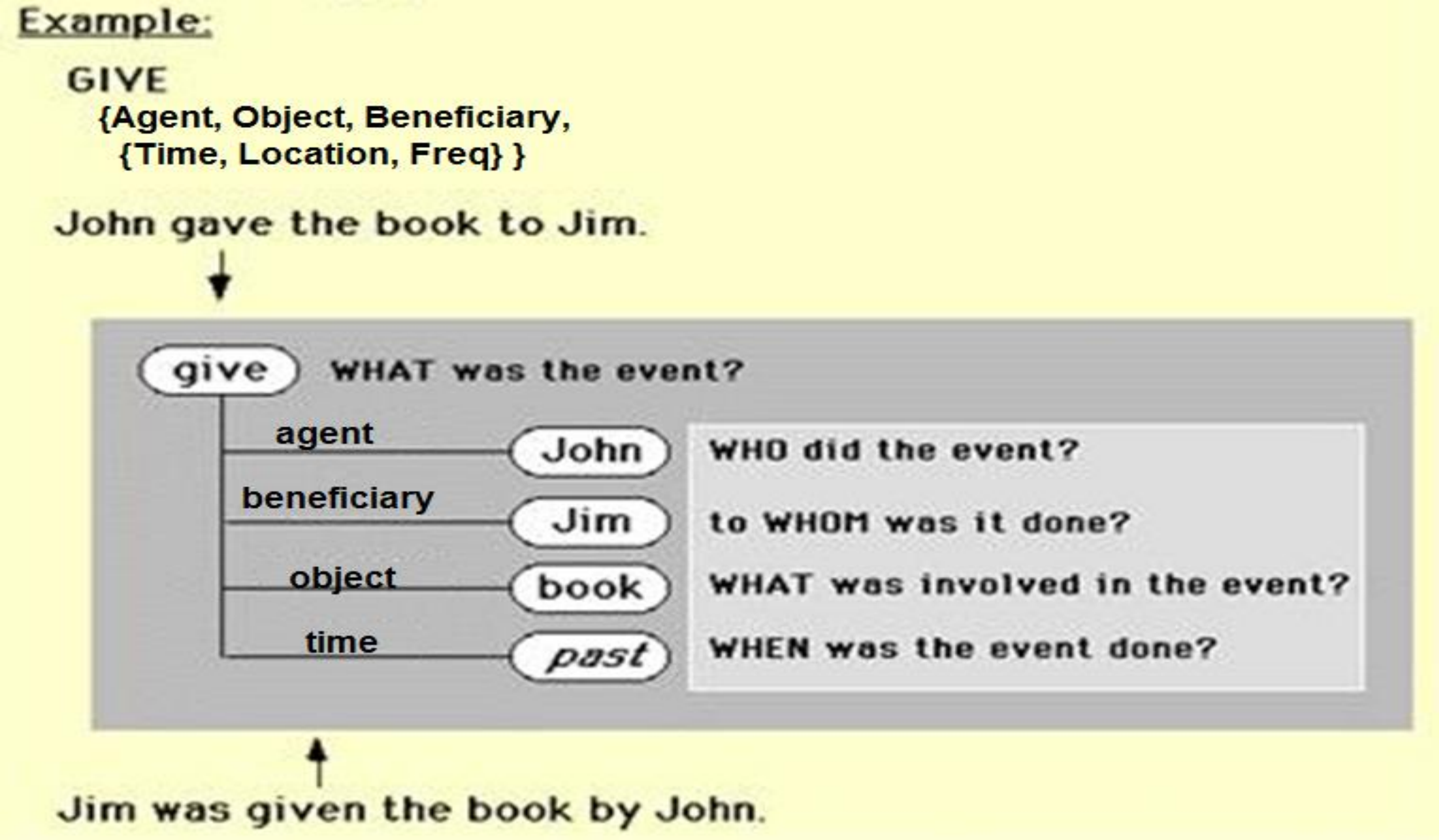
Semantic Representation 2/3

- There are constraints for the deep cases:
  - a deep case can occur only once per sentence
  - Some of the cases are obligatory and others are optional.
  - Obligatory cases may not be deleted.

- Case frames can also impose constraints in their slots (deep cases). These constraints are named “Selectional Restrictions”.
  - For example, the "Objective" slot of the Greek verb “τρώω/έφαγα” ("eat“) should be filled with some kind of food.
  - Consequently, this "selectional restriction" can't allow a mental object (e.g. the language) to fill the "Objective" slot.

- This way in Greek sentence "Εχθές, έφαγα ψητή γλώσσα και πατατοσαλάτα" (translated as "Yesterday, I ate grilled tongue and potato salad") resolve or decrease the ambiguity of Greek word “γλώσσα” (select the meaning "tongue" or "fish" but not the meaning "language").

Semantic Representation 3/3



Anaphora Resolution 1/2

- The problem of resolving what a pronoun or a noun phrase refers to is named Anaphora Resolution.

- "John is going to visit Nick. He is a good man." can have two possible meanings:
  - John (who is a good man) is going to visit Nick,
  - John is going to visit Nick (who is a good man).

- Various Anaphora Types:
  - Pronominal, as in "John works hard. He wants to buy a new car.",
  - Possessive pronoun, as in "Jennifer drove her car in the desert for four kilometers.",
  - Reflexive pronoun, as in "Mary and John had dinner together. Mary cooked a wonderful roast beef by herself.",
  - Reciprocal pronoun, as in "Mary and Aleksandra are friends of each other.",
  - Lexical anaphora, as in "Engineers from many companies attended the conference. The participants found the topics very attractive.",
  - One anaphora, as in "If you can not attend a tutorial in the morning, you can go for an afternoon one.".

Anaphora Resolution 2/2

- Regarding the distance between the anaphor and the referenced (antecedent)
  - Intrasentensial: referenced and anaphor are in the same sentence,
  - Intersentensial: referenced and anaphor in different sentences.

- Factors used in the resolution include:
  - gender and number agreement,
  - “c-command” constraints,
  - semantic consistency,
  - syntactic parallelism,
  - semantic parallelism,
  - salience,
  - proximity etc.

- An anaphora resolution system based on c-command is discussed by [Karanikolas93].

- Every other anaphora resolution solution based on c-command requires detailed syntactic/parse trees

Affective Computing & Discourse Analysis

- Affect is an attitude or emotion that a speaker brings to an utterance.
- Affects are: admiration, anger, contentment, disbelief, disgust, excitement, …

- Affective Computing is the expanding capability of smart devices for detecting and responding to human emotions.

- Discourse Analysis will enable an NLU system to reveal the hidden motivations behind a text.

- Uncover:
  - rhetorical relations between sentences (elaboration, contrast, generalization, violated expectation, …),
  - rhetorical relations of causality (why something was said),
  - rhetorical relations of time (sequence, parallel).

Pragmatics 1/2

- With Pragmatics we actually refer to Speech Acts & Implicatures.

- Speech acts are embedded in discourse and they are distinct from physical acts (performing some act) and mental acts (thinking to do or how to do some act).

- Three basic types of direct speech acts:
  - Assertion (Declarative sentence with the verb in Indicative mood),
  - Question (Interrogative sentence with a WH word or auxiliary verb and a “?”
  - Order (Imperative sentence with the verb in Imperative mood).

- There are also indirect speech acts (Greek Language examples):
  - “Ρωτάω αν υπάρχουν περιθώρια βελτίωσης.” (I am asking if there is room for improvement.), there are no signs that it is an interrogative sentence (a Yes/No)
  - “Αν ξεπεράσεις τα όρια, θα απολυθείς.” (If you cross the line, you will be fired.). It is a threat but there is no special sentence form for threats.

Pragmatics 2/2

- Implicatures is another dimension in pragmatics.

- Example:
  - “Mary and Helen are mothers.”,
  - “Tina and Flora are sisters.”.
  - Only in the second sentence we infer that persons have a (reflexive) relationship (sisters of each other).
  - Our beliefs say that it is not possible a mother being child of her daughter.

- In general, with the term “implicatures” we determine the sophisticated endeavor of some advanced NLU systems to identify the full range of inferences that a reader or a hearer would make when confronting/encountering the locutions of an author or a speaker, considered in context.

Shallow Parsing 1/2

- Deep (complete) parsing delivers (almost) always syntactic ambiguity. The ambiguity can be resolved later by next steps. This complicates understanding.

- Partial parsing (parsing of textual chunks) can be useful alternative.

- An example:
  - Sentence: Look at the dog with one eye
  - POS-Tagger output: Look/vb, at/in, the/det, dog/n, with/p, one/det, eye/n
  - Chunker output: [v Look] at [np [det the] [n dog]] [pp [p with] [np [det one][n eye]]]

```
s( np(n(you))
   vp( v(look),
       pp (p(at),np(det(the),n(dog))),
       pp (p(with),np(adj(one),n(eye)))
   )
)
```

where:
vb = verb base form;
in = preposition or subordinate conjunction;
det = determiner;
n = noun;
p = preposition;
np = noun phrase;
pp = prepositional phrase.

Shallow Parsing 2/2

- A Shallow parser is usually working in three step process. The steps are:
  - Word Identification (POS-Tagging),
  - Chunk Identification (using regular expressions or Context Free Grammars or Phrase Structure Grammars with simple rules),
  - Merging / Splitting of Chunks (via rules).

- The later (especially Merging) can be done by:
  - Combining adjacent chunks into a single chunk,
  - Define regular expressions that permit to merge sequences of adjacent chunks to a longer one.

- The final chunks, after the third step, are valuable information to feed the Semantic Analysis and Semantic Representation module of an NLU system. So deep parsing can be possibly replaced by the cheaper Shallow parsing.

Strengths of LLMs

LLMs

- produce text which is (in many cases) indistinguishable from human text
- can be used for a wide range of applications and efficiently automate different types of tasks
- can process vast amounts of text data and learn from it
- produce responses fast
- can be tailored to specific use cases through additional training and fine-tuning
- can work with multiple languages
- enhance user experience providing meaningful and context-aware responses

Weaknesses of LLMs (1)

LLMs

- ➢ are prone to produce misinformation, bias, and impolite language resulting from their training data
- ➢ can generate fake, harmful, misleading, content, or propaganda, raising ethical concerns
- ➢ tend to ‘hallucinate’ and produce, with full conviction, high quality text which contains factually incorrect information
- ➢ since they rely on training data, use potentially obsolete information and may generate responses that are no longer accurate or relevant

Weaknesses of LLMs (2)

LLMs

- operate like black boxes, without the possibility to interpret / explain why they produced specific output
- are not easily controllable and when mistaken answers are identified, it is hard, if not impossible to diagnose and fix the error
- may handle sensitive and personal data and measures for data protection are needed
- are costly to train, since massive amounts of data, powerful computing systems and significant energy consumption are needed

Weaknesses of LLMs (3)

LLMs

- perform well in simple linguistic tasks, but given unexpected prompts, they exhibit weird thought processes and clearly non-human understanding
- exhibit limited “creativity”, since they mimic patterns in their training data
- cannot solve particular tasks, such as arithmetic operations and calculations

- The future training data are expected to be increasingly contaminated by LLM-generated contents themselves
- The process of adapting them to a specific / new task can be complex and the results may not always be optimized

Strengths of NLU / NLG

NLU/NLG systems

- do not require a massive training corpus and are most appropriate when extended data, resources and computing power are not available
- are able to adapt to new or rare cases, when these are covered by their rules
- due to using rules engineered by humans, exhibit transparent and interpretable functioning

Weaknesses of NLU / NLG

- NLU/NLG systems require skilled experts to create rules
- There exist numerous human languages with varying grammar and syntax
- Rules are numerous, complex, overlapping, or contradicting and significant effort is needed for their maintenance
- R&D groups working on NLU/NLG worldwide are isolated and not sharing their techniques and accomplishments, due to interests / competition
- Since languages evolve, NLU/NLG systems may become outdated
- Handling ambiguity and variability of natural languages is hard
- NLU/NLG systems may suffer from bias, originating on opinions of their developers

Discussion and Conclusion (1)

- LLMs and NLU / NLG systems have not been adequately compared and their strengths and weaknesses have not been extensively investigated
- Combining LLMs and NLU/NLG algorithms in hybrid systems is also open. In such systems, NLU / NLG algorithms could be used to
  - filter out irrelevant or inappropriate text
  - produce semantically correct information (NLU) based on recorded events in some knowledge base and syntactically correct sentences (NLG)
- The output will be converted into more understandable texts (LLMs)
- A hybrid system might also use NLU / NLG algorithms and LLMs in parallel
- LLMs could discover potential or test existing rules for NLU/NLG systems

Discussion and Conclusion (2)

- NLU / NLG results are often being more consistent and predictable, than the ones produced by LLMs
- NLU / NLG systems can better mimic human creativity and imagination
- More constructive and reliable results could be obtained if NLU / NLG systems are extended with the consideration of big existing data by LLMs

Thank you for your attention

Nikitas Karanikolas
nnk@uniwa.gr

Eirini Manga
emanga@uniwa.gr

Nikoletta Samaridi
nsamaridi@uniwa.gr

Eleni Tousidou
etousido@uth.gr

Michael Vassilakopoulos
mvasilako@uth.gr