LSC - 3 – Modality

# Introduction

Modality is a broad concept studied in different fields, each with its own approach: philosophy, logic, grammar.

General definition in Baker et al, 2010: *“Modality might be construed broadly to include several types of attitudes that a speaker might have toward an event or state”*

Might indicate:

* **Factivity**: whether an event, state or proposition happened or not (desired/planned/probable)
* **Evidentiality**: scope and reliability of information
* **Sentiment**: speaker’s feelings towards event, state or proposition

Example: express different modality of event with respect to degree of certainty, reliability, sentiment

* + - GM will lay off workers
    - A spokesman for GM said GM will lay off workers
    - The politician claimed that GM will lay off workers
    - GM may lay off workers
    - Some wish GM would lay off workers
    - Will GM lay off workers?
    - Many wonder whether GM will lay off workers

# Approaches: an overview

Here we provide some examples of the different approaches to modality.

## Linguistic approach: Palmer, 2001

Basic classification (each with its own even more detailed subclassification)

* **Propositional** modality: *“concerned with the speaker's attitude to the truth-value or factual status of the proposition”* (e.g. “Charlie must be at home now”)
  + Epistemic: express judgement about factual status
  + Evidential: indicate evidence for factual status
* **Event** modality: *“refers to events that are not actualized, events that have not taken place but are merely potential”* (e.g. “Charlie must come in now”)
  + Deontic: related to obligation and permission à factors external to speaker
  + Dynamic: related to ability or willingness à factors internal to speaker

Further categories:future, negative, interrogative, imperative-jussive, presupposed, conditional, purposive, resultative, wishes, and fears.

Relation to tense and aspect:

* Modality concerned with **status** of proposition describing event
* Tense concerned with **time** of event
* Aspect concerned with “the nature of the event in terms of its **internal temporal constituency**”

## Philosophy approach: Von Fintel, 2006

Modality defined as *“a category of linguistic meaning having to do with the expression of possibility and necessity.”* (e.g. “Charlie might be home” à there is a possibility that C is home VS “Charlie must be home” à in all possibilities C is home)

Expressions that can convey modal meaning, aka modal **cues** or **triggers**:

* **Modal auxiliaries**: Charlie might be home
* **Semimodal verbs**: Charlie has to be home
* **Adverbs**: Perhaps Charlie is home
* **Nouns**: There is a possibility that Charlie is home
* **Adjectives**: It is necessary that Charlie is home
* **Conditionals**: If it is evening, Charlie is home

Types of modal meaning based on the premises for the modality (note that the same trigger could express more than one modality):

* **Epistemic**: given what is known (“It has to be cold”, there are people wearing winter coats and gloves)
* **Deontic**: given set of laws/moral principles (“Students have to book a seat”, library regulation)
* **Bouletic**: given desires (“I have to pass this exam”, student who wants good grades)
* **Circumstantial**: given set of circumstances (“You have to call the doctor”, given your current state of poor health)
* **Teleological**: given a goal and possible means to reach it (“To go to work without getting wet, you have to drive your car”)

## Logic approach: Portner, 2009

Modal forms classified based on which level modality appears:

* **Sub-sentential**: at the level of constituents smaller that a clause. Can be expressed through modal adjectives and nouns, verbs and adjectives expressing the attitude (believe, hope, know, certain, pleased…), infinitives, negative polarity items, verbal mood (indicative, subjunctive)
* **Sentential**: at the level of the clause. Can be expressed through modal auxiliaries and adverbs, conditional sentences.
* **Discourse**: at discourse level, any modal meaning which can’t be assigned to the other two types. “I might go to the library. I should return this book”: if I go to the library, I should return the book

# Modality and negation

Modality and negation have not been studied thoroughly in their interaction. However, they can be found combined, giving rise to many strange phenomena which are difficult to explain systematically.

Example: English verb *may* and negation  
He may not have any cake. [deontic, “not allowed”] à modality is the scope of the negation  
He may not be home. [epistemic, “possible that not”] à negation is the scope of modality  
  
Example: English must vs German müssen  
He must not have any cake. [“obligatory that not”] à negation scope of modality  
Er muss nicht zuhause bleiben [“He doesn’t have to stay home.”] à modality scope of negation

# Annotating Modality

For NLP applications having annotated corpora is fundamental. Even though there are several resources with modality and negation annotations at different levels (expression, event, relation, sentence), there is not a defined annotation standard. Here we give two examples of the possible annotation schemas.

## OntoSem project (Nirenburg and Raskin, 2004)

OntoSem is a general-purpose syntactic-semantic analyser: it processes unrestricted raw text carrying out several levels of linguistic analysis. As part of the semantic analysis, it encodes modality information, for which an annotated corpus was created (Nirenburg et al. 2008)

Encoded properties of modality:

* **Type**:
  + Polarity: proposition positive or negated
  + Volition: extent to which someone wants or doesn’t want event/state to occur
  + Obligation: extent to which someone considers event/state necessary
  + Belief: extent to which someone believes content of proposition
  + Potential: extent to which someone believes event/state is possible
  + Permission: extent to which someone believes event/state is permitted
  + Evaluative: extent to which someone believes event/state is a good thing
* **Value**: Scalar between 0 and 1.0
* **Scope**: predicate affected by modality
* **Attributed-to**: to whom the modality is assigned (default is speaker)

Example: “Entrance to the tower **should** be totally **camouflaged**” (meaning: the entrance to the tower is camouflaged)

* **Type**: obligative
* **Value**: 0.8
* **Scope**: camouflaged
* **Attributed-to**: speaker

Note: even if value attribution is arbitrary, it might be useful for relative comparisons (e.g. “disfavour” has a lower value than “adore” on the evaluative scale, no matter the specific values)

## Modality Lexicon (Baker et. Al, 2010)

This is another example of modality annotation strategy. In this paper, authors only focused on those modal words that express ‘factivity’.

Therefore, the modalities taken in consideration were all related to factivity, and they were eight (P indicates a proposition and H the holder (the experiencer) of the modality:

1. **Requirement** (does H require P?)
2. **Permissive** (does H allow P?)
3. **Success** (does H succeed in P?)
4. **Effort** (does H try to do P?)
5. **Intention** (does H intend P?)
6. **Ability** (can H do P?)
7. **Want** (does H want P?)
8. **Belief** (with what strength does H believe P?)

A modality lexicon was then semi-automatically built for the modalities, and this was used for training automatic modality tagging systems.

**The modality lexicon**

The modality lexicon was built for 150 lemmas.

To build it, some assumptions have been made regarding:

* Scope of modality and negation. Same annotation for:
  + **I do not believe that he left**
  + I **believe he didn’t leave**
* Duality of meaning for *require* and permit
  + not require P to be true = Permit P to be false
  + **not permit P to be true = Require P to be false.**
* Entailment between modalities. Annotators were provided a specificity-ordered modality list:
  + **requires → permits**
  + **succeeds → tries → intends →is able → wants**
* Sentences without an overt trigger word are tagged as Firmly Believes
* Nested modalities are not marked, only one modality is marked
* The holder is not marked

Every entry in the lexicon consisted of:

(1) A string of one or more words: for example, should or have need of.

(2) A part of speech for each word: the part of speech helps us avoid irrelevant homophones such as the noun can.

(3) A modality: one of the modalities above.

(4) A head word (or trigger): the primary phrasal constituent to cover cases where an entry is a multiword unit, e.g., the word hope in hope for.

(5) One or more subcategorization codes: patterns that show the structural relationship of targets to triggers for different verb types, where targets define events, states or relations with the scope of the modality

Example:

**String**: Need   
**Pos**: VB   
**Modality**: Require   
**Trigger**: Need   
**Subcat**: **V3-passive-basic** – The government is needed to buy tents.   
**Subcat**: **V3-I3-basic** – The government will need to work continuously for at least a year. We will need them to work continuously.   
**Subcat**: **T1-monotransitive-for-V3-verbs** – We need a Sir Sayyed again to maintain this sentiment.  
**Subcat**: **T1-passive-for-V3-verb** – Tents are needed.   
**Subcat**: **Modal-auxiliary-basic** – He need not go.

The 150 lemmas were identified by using a thesaurus and then a mapping between subcategorization codes and patterns was established.

Here is an example of a sentence tagged through the use of this lexicon:

Input: Americans should know that we can not hand over Dr. Khan to them.

Output: Americans <TrigRequire should> <TargRequire know> that we <TrigAble can> <TrigNegation not> <TargNOTAble hand> over Dr. Khan to them.

# Tasks related to processing modality

* **Detecting Speculated Sentences**: Identifies instances within text where the author expresses uncertainty or conjecture, rather than definitive facts.
  + Example: A medical study stating "The drug could potentially reduce symptoms" would be marked as speculative.
* **Scope Resolution**: Determines the boundaries within a sentence or text that a particular modality or negation affects, clarifying which parts of the text are under its influence.
  + Example: In the sentence “I might call you tomorrow”, the scope of the modality can either be the whole sentence (I generally want to call you, but I’m not sure it will happen tomorrow) or just the verb call (I might call you or I might text you tomorrow)
* **Finding Negated and Speculated Events**: Searches for and identifies events in a text that are described as either not happening or occurring under conditions of uncertainty.
  + Example: "The launch was not successful" indicates a negated event, while "The launch might be delayed" indicates a speculated event.
* **Modality Tagging**: Labels words or phrases within a text that express modality, categorizing them according to their type such as necessity, possibility, or permission.
  + Example: "Are allowed to enter" would be tagged as permission, and "Should consider" as a suggestion.
* **Belief Categorisation**: Classifies statements according to the level of belief or certainty expressed by the speaker or writer.
  + Example: "I believe the results are accurate" would be categorized as expressing a personal belief, while "The results are confirmed" expresses certainty.
* **Processing Contradiction and Contrast**: Identifies and processes instances where two or more elements in a text are in opposition or where a contrast is drawn, which may involve negation or contrasting modality.
  + Example 1: "He claimed to be at work, but his colleague said he was absent" presents a contradiction that needs to be resolved.
  + Example 2: "The software is easy to use, unlike the previous version" highlights a contrast that needs to be understood in context.

# Modality in applications

* **Sentiment Analysis**: Determines the emotional tone behind a body of text, identifying positive or negative sentiments, their intensity, and the specific entities they refer to, while also distinguishing between subjective opinions and objective statements.
  + Example: A review stating, "I absolutely love the new features of this app, but the customer service can be unhelpful," would be analysed for the positive sentiment towards the app's features and the negative sentiment and modal expression of possibility towards customer service.
* **Recognizing Textual Entailment**: Assesses whether the truth of one text fragment follows from another text, considering the context and the modality expressed within the content.
  + Example: Given the statement "John must submit the report by Friday," a system recognizes that "John is required to submit the report this week" is entailed.
* **Machine Translation**: Translates text from one language to another, ensuring that nuances of modality and negation are preserved to maintain the original meaning and intent.
  + Example: Translating "You shouldn't park here" from English to French must convey the prohibition modality accurately.
* **Text Mining**: Extracts meaningful information from text, identifying patterns, trends, and relationships, and interpreting the modality to understand the context and implications of the content.
  + Example: Mining customer feedback to find common issues mentioned, such as "often crashes" or "usually excellent service," considering the frequency modality.
* **Identifying the Structure of Scientific Articles**: Analyses the composition of scientific texts to categorize and summarize the various sections, such as abstract, methodology, results, and conclusion, for better information retrieval and understanding.
  + Example: Detecting the hypothesis in the introduction, "This study aims to prove..." and the results in a separate section, "The data confirms..."
* **Trustworthiness Detection**: Evaluates the reliability of information by analysing the language used for certainty, evidence, and credibility indicators, often considering modality markers that may signal uncertainty or subjective claims.
  + Example: Determining the reliability of a rumour on social media, such as "It is rumoured that..." by looking for corroborating evidence or lack thereof.

# Symbolic representation

## Review

Like for negation, systems for modality detection are **Rule-Based.**

Different rule-based systems may use different language-based tools for identifying and dealing with modality, sometimes utilising more of them together:

* **Trigger Words:** identify modality by looking up for modality words on a list (such as need, can, want). Whenever a word from this list is encountered, then we have a hint that a modality statement may be present.
* **POS tags:** to identify patterns which may indicate the presence of a modality statement and help to disambiguate. For example, if the word ‘need’ is a noun like in the sentence: “The medicines were in strong need”, then this is not a modality statement.
* **Dependency parsing:** to identify modality by looking at more complex relationships between the sentence’s elements. For example, the modality word is expected to be the head of a sentence.
* **Regular expressions:** identify patterns in text that could be associate with modality statements.
* **Logical languages:** use the constructs of formal logical languages to represent modality statements.

## Modal logic

A logical language explicitly designed to represent modality statements. Classical modal logic only focused on representing **necessity** and **possibility,** but it has evolved to represent also **obligation, permission, belief** and to provide a temporal dimension to such statements.

Classical modal logic is built upon propositional logic, but it adds two operators and new rules to deal with them: □ to represent **necessity**, and ◇ to represent **possibility.**

**Some examples:**

* “It is necessary that John cooks” 🡪 □ Cooks(John)
* “It is possible that John cooks” 🡪 ◇Cooks(John)

(“John cooks” is represented by Cooks(John))

# Statistical representation

## Review

Statistical NLP models like **Naive Bayes**, **Support Vector Machines (SVMs)**, and **Neural Networks (NNs)** traditionally rely on bag-of-words and TF-IDF vectors to represent text data. In the context of mood detection:

* **Naive Bayes** classifiers assume feature independence and are often used for their simplicity and speed in baseline mood detection models.
* **SVMs** are effective in high-dimensional spaces, making them suitable for datasets with a wide array of mood-indicative features.
* **NNs**, especially shallow ones, can capture non-linear relationships between features and moods, offering a more nuanced mood classification.

### Need for Tailored Datasets

* Training and evaluating these models necessitate datasets annotated with mood labels.
* The quality of mood detection is heavily dependent on the representativeness and granularity of these datasets.
* Feature engineering is critical as it directly influences the model's ability to generalize from the training data to real-world applications.

### Why Feature Engineering Matters:

* It allows the model to focus on relevant aspects of the data, improving mood detection accuracy.
* Tailored features can capture subtleties in language that express mood beyond mere word presence (e.g., intensity of adjectives, negation handling).
* Proper feature selection can significantly reduce the dimensionality of the data, improving computational efficiency.

## Statistical: Feature engineering for mood

* **Sentiment Scores**: Quantitative measures that reflect the positive, negative, or neutral sentiment of words or phrases.

Usage: Classify the overall mood of the text.

* **Affective Lexicons**: Collections of words associated with different emotions or moods, often annotated with their strength or polarity.

Usage: Help in identifying specific moods such as happiness, anger, or sadness from the text.

* **Psycholinguistic Features**: Features derived from psychological research on language processing, which may include variables like concreteness, valence, arousal, and dominance.

Usage: Predict the emotional impact of words beyond simple positive or negative sentiment.

* **Syntactic Features**: Features that capture the structure of sentences, such as part-of-speech tags, phrase structure, and grammatical relations.

Usage: Indicate modality using modal verbs, adverbs, and other parts of speech that convey the speaker's attitude.

* **Semantic Features**: Features that capture the meaning of text, such as semantic role labelling, word-sense disambiguation, and entity recognition.

Usage: Understanding the context in which mood expressions occur, which can affect the interpretation of modality.

* **Pragmatic Features**: Features that consider the context of communication, including speaker intent, politeness, and formality.

Usage: Indicative of the speaker's mood and the degree of belief or certainty expressed in the text.

* **N-Grams and Collocations:** Features that capture sequences of words (bigrams, trigrams, etc.) and their frequency of co-occurrence.

Usage: Useful for detecting fixed expressions that convey mood or modality, such as idioms or phrases like "couldn't be happier."

* **Negation and Intensifiers**: Words that modify the intensity of emotions or reverse their polarity, such as "not" for negation or "very" for intensification.

Usage: Critical for accurate mood detection, as they can change the sentiment of an entire phrase or sentence.

# LLMs representation

## Review

* **Transformers**: The foundation of modern LLMs, Transformers revolutionized NLP with their ability to process words in relation to all other words in a sentence, simultaneously. This architecture is pivotal for understanding the complex nuances of human language, including mood and modality.
* **GPT (Generative Pre-trained Transformer)**: A series of models culminating in GPT-3, known for their deep learning capabilities in generating human-like text. GPT models excel in capturing the subtleties of mood through extensive pre-training on diverse internet text.
* **BERT (Bidirectional Encoder Representations from Transformers)**: BERT's bidirectional training allows for a deeper understanding of context, improving the detection and representation of mood in text. Unlike traditional unidirectional models, BERT analyses text in both directions, offering a more nuanced interpretation of modality.

**Word Representation**: LLMs represent words as high-dimensional vectors, capturing semantic meaning and syntactic relationships. These representations are dynamic, allowing for the modulation of meaning based on surrounding text, which is crucial for identifying mood.

Main Characteristics of LLMs:

* **Attention Mechanism**: Allows the model to focus on relevant parts of the text when predicting the next word or classifying text, which is essential for understanding mood and its triggers.
* **Positional Encoding**: Gives the model a sense of word order, which can change the meaning and mood conveyed by a sentence.
* **Training on Large Datasets**: LLMs are trained on vast corpora of text, enabling them to recognize a wide range of moods and modalities in different contexts.
* **Fine-Tuning**: After pre-training, LLMs can be fine-tuned on specific tasks, including mood detection, by adjusting the model to the nuances of the new dataset.

## Ethical and Societal implications

In our previous presentations, we have navigated through the intricate workings of Large Language Models (LLMs) and their profound capabilities in representing mood within textual data. As we continue to unravel the layers of this advanced technology, it becomes imperative to address the ethical and societal implications that emerge specifically from the application of LLMs in mood analysis.

The implications we are about to explore are a subset of a broader spectrum, each deeply entwined with the nuanced ways in which AI systems understand and interact with human emotions. These selected areas of focus are particularly pertinent, given our discussions on mood representation in LLMs, and they highlight the ethical and societal ripples that such technology casts across the fabric of our daily lives.

As we delve into these topics, let us remember that our exploration of mood representation in LLMs is not just a technical journey but also a path that runs parallel to ethical reflection. The decisions we make today in shaping these AI systems will echo in the societal narratives of tomorrow.

* **Bias and Fairness**: Bias in AI refers to systematic errors that create unfair outcomes, such as favouring one demographic over another. In mood detection, this could manifest as misinterpreting language use or emotional expression from underrepresented groups.
  + **Example**: An LLM consistently misinterprets expressions of emotion in African American Vernacular English (AAVE), leading to incorrect assessments of mood.
  + **Consequences**: Such biases can perpetuate stereotypes and discrimination, affecting everything from job hiring to legal sentencing.
  + **Solutions**: Diversifying training data, involving multidisciplinary teams in development, and conducting regular audits for bias can help create fairer AI systems.
* **Emotional Manipulation**: Emotional manipulation through AI occurs when systems use mood detection to influence a user's emotions or actions, often without their awareness.
  + **Example**: A music streaming service changes playlists based on detected mood to keep users listening longer, potentially affecting their emotional state.
  + **Consequences**: This can lead to manipulation of consumer behaviour, impacting mental well-being and personal autonomy.
  + **Solutions**: Establishing ethical guidelines, transparency in AI operations, and giving users control over if and how their mood data is used can prevent manipulation.
* **Mental Health**: The intersection of AI and mental health in mood detection can have profound implications, especially regarding the accuracy and use of such sensitive data.
  + **Example**: A mobile app analyses text messages for signs of depression and sends alerts to users or their contacts, sometimes based on out-of-context information.
  + **Consequences**: Misinterpretation can lead to unwarranted concern, privacy invasion, and stigma around mental health issues.
  + **Solutions**: Collaborating with mental health professionals, ensuring accuracy, and providing clear opt-out options can enhance the responsible use of mood detection for mental health.
* **Accountability and Transparency**: Accountability in AI necessitates that developers and companies are responsible for the outcomes of their systems, while transparency involves clear communication about how AI systems operate and make decisions.
  + **Example**: An LLM used for hiring misjudges a candidate's mood as negative, leading to an unfair rejection.
  + **Consequences**: Lack of accountability and transparency can result in mistrust and harm to individuals unfairly judged by these systems.
  + **Solutions**: Creating explainable AI, documenting decision-making processes, and establishing clear accountability structures can improve trust and fairness.
* **Surveillance and Monitoring**: Surveillance and monitoring refer to the systematic observation of individuals, often without their consent, which can be facilitated by mood detection technologies.
  + **Example**: Employers monitor employee communications for mood and sentiment, potentially leading to punitive measures or pressure to display certain emotions.
  + **Consequences**: This can create a culture of fear and self-censorship, impacting freedom of expression and workplace morale.
  + **Solutions**: Enforcing strict limitations on surveillance, ensuring transparency, and protecting employee rights can mitigate these issues.
* **Long-Term Societal Effects**: The widespread use of mood detection technology can have lasting effects on societal norms and individual behaviour.
  + **Example**: If people know their mood is constantly being analysed, they might alter their behaviour online to conform to perceived norms.
  + **Consequences**: This could lead to a homogenization of emotional expression and a reduction in personal authenticity.
  + **Solutions**: Promoting ethical use, public discourse on the implications of mood detection, and establishing societal norms for technology use can help navigate these long-term effects.

# Conclusion

* LLMs have changed everything. Compared to previous methodologies, they have been capable of representing modality like never before.
* However, benchmarks confirm that there still a lot to do.
* The future of AI is a mystery, and nobody really knows what will happen in the long and short terms.

# References

<https://direct.mit.edu/coli/article/38/2/223/2142/Modality-and-Negation-An-Introduction-to-the> (Modality and Negation: Introduction to the Special Issue)

<https://aclanthology.org/W04-0905.pdf> (Nirenburg et. Al 2004)

<https://arxiv.org/pdf/1410.4868.pdf> (Baker et. Al 2010)

<https://web.mit.edu/fintel/fintel-2006-modality.pdf> (Von Fintel 2006)

<https://mirror.aclweb.org/ijcnlp11/downloads/tutorial/tu3_present.pdf> (Tutorial on Modality and Negation in NLP)