Introduction

We are presenting an approach involving dissecting and analyzing text through a hierarchical stack of modalities ranging from core meaning to complex sociocultural and psychological layers, integrating concepts from multiple scientific disciplines and methodologies.

This multidisciplinary approach doesn't fit neatly into a single existing category but spans several areas of study. To define the underlying science or methodology we call it **"Computational Semiotics within Multimodal Discourse Analysis."**

This name captures the integration of computational methods with the study of signs and symbols, including cultural references (semiotics) as they operate across various modes of communication (discourse analysis).

By leveraging advanced computational techniques, embedding vector mathematics, and knowledge graph technologies, we seek to develop a suite of tools capable of capturing the full spectrum of human expression—from the core meanings of phrases to their most nuanced cultural, emotional, and contextual layers.

This approach, which could be seen as a cutting-edge fusion of computational and humanistic sciences, aims to enhance our understanding of language and communication in their fullest complexity. By naming and framing this methodology, we acknowledge the complexity of human communication and the interdisciplinary effort required to capture its nuances computationally. ### Global trends and principles

Information overload, or the difficulty in understanding an issue and making decisions caused by the presence of too much information, is exacerbated in the digital age by the rapid proliferation of data, misinformation, and the complexity of global issues. This overload contributes to the 'infocalypse'—a state where fake news, biased information, and digital manipulation overwhelm factual, reliable content, making it increasingly difficult for individuals to discern truth and make informed decisions.

A human desire to long for The Same causes many of us to produce and consume the same type of content with slight variations. The same little dance is performed by thousands of vloggers, but even movies and series tend to follow carefully crafted sameness rules in characters, storylines, and audiovisual representations to be essentially near-duplicate productions. The proliferation of the mass amount of content only worsens the information overload problem.

Solving problems by doing more of the same is a trend we need to break. Many approaches to information overload focus on creating even more, but better targeted, information on new platforms, resulting in even more information overload. We have to look into principles to actually compress and reduce the amount of available information.

UX as a method of creating time, space, and friction is an approach that counters the movement of trying to smooth every interaction into a frictionless, highly personalised experience. A meaningful experience requires a challenge and obstacles to overcome, as

well as invite people to adapt and re-evaluate their way of thinking and acting to achieve the desired outcome.

Humanity Centered Design is another way of designing for playing **The Longest Game**, where we keep the long-term goals of sustainable humanity in the long term in mind, and not conservatism or problem solving on short-term problems. This also means sometimes taking the non-obvious route and divergent research without direct applications in mind.

Envisioned products

Large Language Models (LLMs), with their deep understanding of context and ability to generate coherent text, excel at detecting nuances and dissecting modalities within text. They can infer underlying perspectives, detect subtle shifts in worldviews, and even mimic personality traits in their responses. This capability stems from their extensive training on diverse datasets, enabling them to recognize and replicate the myriad ways in which meaning is constructed and communicated across different cultures and languages.

1. **Universal Translator Concept** - Building on these capabilities, the concept of a universal translator that goes beyond mere linguistic translation to encompass perspectives, worldviews, and personalities is both groundbreaking and complex. Such a translator would not only convert text from one language to another but also adapt the translation to reflect the cultural nuances, societal norms, and individual personalities involved in the communication.
2. **Opinion Difference Analyzer** - Opinions are layered with personal experiences, emotions, cultural background, and socio-political context. These layers add depth and complexity but can also obscure the underlying agreement or common ground between differing opinions. By comparing the neutral cores and the modality layers of the opinions separately, it becomes possible to identify exactly where the disagreements lie and what aspects are actually in agreement. In public policy and diplomacy, understanding the core concerns and separating them from rhetorical flourishes can aid in crafting compromises that address the fundamental needs of all parties.

Underlying Concept

Embedding vectors are high-dimensional spaces where words, phrases, or even entire documents are mapped to vectors (lists) of real numbers. This mapping captures not just the semantic meaning of a word but also its association with other words, contextual usage, and connotations, effectively making these vectors nuggets for a nexus of meaning. The strength of these embeddings lies in their ability to encapsulate nuances in meaning that arise from context, a crucial aspect of understanding modalities like perspectives and worldviews.

The power of the embedding vectors is that you can do mathematical operations on them which have meaningful consequences for the meaning of the text.

Design

The Modality Stack

Creating an exhaustive and hierarchical stack of modalities, starting from the core meaning and expanding to incorporate increasingly nuanced layers, involves categorizing the various dimensions of language and communication in a logical order. This hierarchy begins with the most fundamental, neutral aspects of meaning and progressively includes more context dependent and nuanced modalities.

The exact layers and stack is in the design phase not of the utmost importance and can be fine-tuned later on. Since the system is backed by vectors, we can validate mathematically of its layer and constituents, as is discussed in the next section of the plan. We note this assumes the validity of the embedding vector.

Because of the nature of embeddings, we could start with an utterance and peel away layer by layer (in any particular order as a matter of fact) to create a new utterance with some of its features stripped away, until only the core remains. This is the decomposition or deconstruction part of the analysis. We can also create new utterances by transforming the modality at one particular layer, e.g. transforming from positive to negative sentiment, or changing from a rational to traditional world view, while keeping the core meaning intact.

A single utterance can therefore be transformed to a multitude of decomposed and transformed new utterances, each with its own embedding vector. We will exploit this feature.

Core

1. **Nugget Core Meaning:** The fundamental, context-independent meaning of the text. This is the most abstract representation of an idea or fact, stripped of any emotional, cultural, or stylistic influences.

2. **Lexical Semantics:** The dictionary meanings of words and phrases and their relationships (e.g., synonyms, antonyms). This layer adds the first level of specificity to the core meaning.

3. **Syntactic Structures:** Grammar and sentence structure that organize words and phrases into coherent statements. This modality includes understanding parts of speech and their arrangements.

These first three layers are so elementary linked to the medium text that together we can represent this as text and preserve the most neutral specific core meaning. We could also represent it in a knowledge graph. For this project, we consider layer 3 as the lowest level of our stack.

Neutral Linguistic Intent and Quality

4. **Pragmatic Context:** The intended use of language in situational contexts, including speech acts (e.g., requests, offers, commands) and implicatures, which require an understanding of the speaker's intentions and the conversational context.

5. **Referential Context:** The specific entities, locations, times, and real-world references mentioned in the text. This layer anchors abstract meanings to concrete instances.

6. **Discourse Coherence:** The logical flow and connectivity of ideas across sentences and paragraphs, ensuring that the text forms a coherent whole rather than disconnected fragments.

The following three layers add a neutral context to the core message placing it in a particular application context and whether the message has a coherent application at all.

Social Code

7. **Sentiment and Emotional Tone:** The emotional layer of the text, which includes sentiments (positive, negative, neutral) and more specific emotional states (joy, anger, sadness).

8. **Cultural References and Symbolism:** Implicit and explicit references that require cultural knowledge to decode, including idioms, proverbs, cultural symbols, and allusions.

9. **Sociolinguistic Variations:** Variations in language use influenced by social factors, including dialects, sociolects, and registers. This modality reflects the social identity and status of the speaker or writer.

The next three levels add meaning that resonate in a specific social and cultural context.

Personal identity

10. **Stylistic Features:** The choice of words, tone, and rhetorical devices that reflect the author's personal style, genre conventions, or the text's intended effect on the reader.

11. **Intertextuality:** References to and influences from other texts, which require knowledge of those texts to understand fully.

12. **Philosophical and Ideological Stances:** The underlying beliefs, worldviews, and ideologies that shape the perspective from which the text is written.

13. **Temporal and Spatial Contexts:** The historical time, geographical place, and cultural setting in which the text was produced and is interpreted.

14. **Psychological and Personality Traits:** Indications of the psychological state or personality traits of the speaker or writer, as inferred from language use patterns.

The last layers represent the nuances from a personal experience, values, take on the world and context in which the author operates.

This hierarchical stack represents a comprehensive approach to dissecting the multifaceted nature of text and communication. Starting from the abstract core meaning, it progressively layers on the nuances that give text its rich, complex, and context-dependent meanings.

Mathematical Modeling

Using mathematical operations on embedding vectors, it is possible to manipulate and transform these vectors to achieve specific goals, including the extraction of a more neutral, core meaning from text by reducing the influence of modality-specific nuances.

To mathematically model the stack of modalities as a linear combination of vectors within a vector space, we can start by considering each modality as a vector in a high-dimensional space, with the entire set of modalities forming a basis for this space. This approach allows us to represent texts or linguistic entities as combinations of these basis vectors, encapsulating the nuanced layers of meaning, context, and interpretation.

Vector Space Model

Assume we have a set of modalities $M_1, M_2, \ldots, M_n$, where each modality M_i is represented as a vector in a (d)-dimensional space, (d) being sufficiently large to capture the nuances of language. The core meaning can be represented as the base vector M_1, with subsequent vectors representing increasingly nuanced layers, as previously described.

A text (T) at a certain level of nuance can then be represented as a linear combination of these modality vectors:

$$T_n = a_1 M_1 + a_2 M_2 + \cdots + a_n M_n$$

where $a_1, a_2, \ldots, a_n$ are coefficients indicating the extent to which each modality contributes to the text's overall meaning.

Each Modality vector is also linear combination of its own base vectors:

$$M_i = b_1 M_{1,1} + b_2 M_{1,2} + \cdots + b_n M_{1,n}$$

For the core modality M_1 the base vectors are simply the unit vectors of the embedding vector and we consider this to be rather black box. For higher levels, they can represent clear dimensions of the modality such as for sentiment ($i = 7$) where the base vectors $M_{7,1}, M_{7,2}, M_{7,3}$ represent negative, neutral, and positive sentiment.

Defining, choosing or finding the exact modality vector corresponding to our modality stack is a complex challenge and near impossible to do perfectly. We feel that this is not required for the model and the application to be powerful. However, there are metrics to validate some aspects of the model. For instance, ideally the vectors $M_1, M_2, \ldots, M_n$ form an orthonormal base, meaning that each modality addresses an aspect of the utterance that is independent of the others and the coefficients a indicate the weighting of the modality on the whole. Similarly, for each modality vector, its base vectors should also form an orthonormal base. We can use the dot product to measure the orthogonality of a set of vectors.

Vector operations

By subtracting modality or adding modality vectors we can add or subtract the presence of the modality in the text. With a scalar multiplication, we can modulate a modality's influence on the text. We can isolate a specific modality vector using vector projection. We can compare two texts at any modality level by doing a distance measure, which for large embedding sizes we will use the cosine similarity measure. Once we have a set a of modality vectors we can do a complete decomposition to obtain all the weightings at once by solving a straightforward linear algebra problem.

Some of these vector operations depend on whether we can isolate clear modality vectors at all. The easiest way to get started is by cleverly prompting an LLM fine-tuned on instructions as will be discussed in the approach section.

By building up a dataset we can use dimension reduction techniques and semantic clustering and averaging (which work on the basis of the above mentioned vector operations) to try and isolate generalized modality vectors. With that we could model and transform the text purely using the vector embeddings and use a embedding-to-text generator as the universal translator. However, this requires fine-tuning our own encoder-decoder transformer models, instead of using off the shelf decoder-only models such as GPT.

Conceptual approach

We start with a naive text deconstruction and transformation approach:

1. **Deconstruction:** starting with a text T_n, we use a clever prompt to remove a modality and generate a new text T_{n-1}. Continue this until we reach T_3 which is our core level.
2. **Isolate case particular modality vectors** M_i: using the embedding vectors that can be created for all T_n.
3. **Generate transformation stack:** use clever prompting to generate transformed texts side-steps per modality (e.g. from positive to negative sentiment), meaning, resulting in: $T_n^j = b_j M_{n,j} \sum_{i=1}^{n-1} a_i M_i$
4. **Isolate case particular sub-modality vectors** $M_{i,j}$**.**
5. **Visualize and validate:** We now have all information available to visualize the modalities for each text, and also show the myriad of forms the same core meaning could be uttered. The measure of orthogonality of the modality vectors indicates the validity of the transformation.

This can already be applied to a case of comparing two different opinions:

1. **Naive text deconstruction and transformation approach** for both texts containing opinion A and B, T^A, T^B.
2. **Distance measurement** based on embedding vectors only, to detect for each layer and sub-modality the alignment between the two opinions,

3. **Visualize and validate:** Show how the distance between the opinion evolves in the stacked modalities and what the path is to bring both closer together.

If the results of the naive approach look promising, we can take a more rigorous and abstract approach.

1. **Data generation:** By performing the naive deconstruction approach for a large dataset, we create multiple meaning trees.
2. **Generalized modality (sub)-vectors:** Based on these trees we can construct and isolate more generalized modality (sub)-vectors using semantic clustering.
3. **Free-form meaning vectors:** we can now play with the modality weightings to fine-tune any transformation mathematically.
4. **Fine-tune an encoder-decoder model:** The trees also allow training a decoder to produce the different transformed texts based on a embedding vector alone (something decoder-only LLMs cannot do out-of-the-box).

This would make generating all the transformed translations obsolete. And open the road to filtering and removing a lot of similar and

It is interesting to convert the texts at different levels into Knowledge Graphs, using named entity recognition to offer a different way to explore the semiotic space.

Another interesting avenue would be to use the constructed language *Lojban* as an additional representation. This language was designed to be unambiguous and highly expressive in terms of the here described modalities. Utterances could perhaps be more efficiently and accurately represented in Lojban than in a natural language such as English, and could serve as such as a (moderately) human-readable intermediate between embedding vectors and natural language.

Phasing

Phase I: Foundation and Framework Development

- Define the hierarchical stack of modalities and establish the mathematical models for embedding vector manipulation.
- Construct and prototype with relevant prompting techniques.
- Functionally define the steps of the naive deconstruction and transformation approach.

Phase II: Tool Development

- **Universal Translator:** Build the initial version focusing on key languages and cultural contexts, integrating modality-specific translation capabilities and a visual interface.
- **Opinion Difference Analyzer:** Create a prototype that can dissect text into modal layers, providing insights into the structure and basis of differing opinions.

Phase III: Testing, Refinement, and Expansion

- Conduct comprehensive testing with diverse inputs to identify areas for improvement and refinement.
- Refine the modalities based on validation metrics and other tests.

Phase IV: Launch and Ongoing Development

- Open use of the tool to external users (within contained user group).
- Seek connection to use cases in other projects.

Phase V: Ongoing Development

- Design and populate the multilayered knowledge graph with initial datasets, ensuring robust representation of languages, cultures, and modalities.
- Integrate the knowledge graph with both the Universal Translator and Opinion Difference Analyzer, enhancing their analytical depth and accuracy.
- Add Lojban as representation.
- Isolate generalized modality vectors based on the larger test dataset.
- Train a encoder-decoder to produce texts based on embeddings alone.

Challenges and Disclaimers

We are well aware that describing subtle highly complex language using a relatively simple and finite abstract model is fraught with danger. It is important that our project is self-aware and open about these limitations and take a descriptive rather than prescriptive attitude. Here, we list several of the challenges and pitfalls:

- While embedding vectors provide a rich semantic understanding, they are not without limitations. One key challenge is the inherent bias that can be encoded within these vectors, reflecting the biases present in the training data.
- Additionally, the abstraction of meaning into high-dimensional spaces can sometimes lead to a loss of nuance, especially in interpreting complex human emotions and cultural specifics, which are critical when considering perspectives and worldviews.
- The deconstructive approach may not always perfectly isolate or remove all nuances of a given modality, especially in complex texts where modalities are deeply intertwined with the factual content.
- The success of this method heavily relies on the skill in crafting prompts and the model's understanding of nuanced instructions.
- In its naive presentation the system is not very scalable and adaptable to accommodate future advancements in technology and shifts in global discourse. Practically, the tree of possible text transformations scales exponentially with the number of (sub)-modalities.
- It may be hard to maintain sensitivity to cultural differences and ethical considerations in both translation and opinion analysis.
- The process of identifying a truly neutral core is fraught with difficulties, as biases and assumptions can influence what is considered "neutral."

- Technical Limitations: The current state of NLP and machine learning might not always accurately dissect and recombine the nuances of human opinions, especially given the subtlety and complexity of human beliefs and communication styles.

Concluding thoughts

Deconstructivism posits that meanings are not fixed but are constantly deferred through an endless chain of signifiers. Algorithms should be flexible and nuanced in handling the dynamic nature of language and meaning, especially in the context of translating nuanced modalities. In addition, meaning is heavily context dependent. Deconstructivism also emphasizes the role of difference and opposition within language.

Deconstructivist philosophies encourage the development of a translator that accounts for the fluidity of meaning across different contexts, rather than relying on static, one-to-one correspondences between words across languages. The translator could be more sensitive to cultural nuances, recognizing that translations must go beyond literal meanings to capture the full spectrum of cultural and contextual significances. The tool can be designed to offer flexible interpretations of opinions, highlighting the potential for multiple readings and the importance of considering various contextual factors.

Postmodernism, with its skepticism towards grand narratives and emphasis on the plurality of truths and interpretations, underscores the complexity of achieving accurate and meaningful communication across diverse linguistic and cultural landscapes. This perspective highlights several key points relevant to the project: multiplicity of meanings, cultural relativism, deconstruction of Text.

We adopt the ethos of moving beyond the cynicism of postmodernism towards a synthesis of opposing perspectives which can inspire the development of tools that seek not just to translate or dissect, but to bridge understanding between disparate viewpoints.

- **Constructive Dialogue:** Emphasizing both sincerity and irony, we encourage a constructive dialogue that acknowledges complexity while striving for authentic communication, mirroring the project's goal to enhance dialogue and understanding.
- **Embracing Contradictions:** By embracing contradictions, the project can develop innovative approaches to handling the ambiguities and complexities inherent in translating and analyzing opinions, reflecting the balance between various poles of thought.

So why are we relevant?

1. **Reflecting Contemporary Realities:** The project is timely and relevant, directly addressing the fragmented, diverse, and interconnected world that postmodern and **metamodern** perspectives describe. It acknowledges the complexity of modern communication and seeks innovative solutions.
2. **Bridging Cultural Divides:** By leveraging these philosophical insights, the project positions itself as a bridge-builder, capable of navigating and reconciling the cultural and linguistic diversity that characterizes our era.

3. **Fostering Inclusive Dialogue:** The tools developed under this project embody the principles of inclusivity and multiplicity of perspectives, essential for fostering a global dialogue that respects and incorporates diverse viewpoints.
4. **Innovating Communication:** The project represents an innovative leap forward in communication technology, inspired by postmodern skepticism and metamodern optimism, aiming to create more nuanced, understanding, and constructive interactions across cultural and linguistic barriers.