Theoretical Foundations

version 1.2.0 · 17 December 2024

This document and its contents were created by Cesar Gonzalez-Perez, Martín Pereira-Fariña and Beatriz Calderón-Cerrato.

Partial funding was provided by Incipit · CSIC, USC, and project ACME (grant number PID2020-114758RB-I00 funded by MCIN/AEI/10.13039/501100011033, Retos de la Sociedad, Plan Estatal 2017-2020, Spain).

21 February 2025 0:51 · revision 1430

Table of Contents

Introduction

IAT/ML is a method for the joint analysis of discourse that combines ontological, argumentation and agency perspectives. It is based on Inference Anchoring Theory (IAT) and the ConML conceptual modelling language.

This document describes the theoretical foundations of IAT/ML, and explains why each major component of the methodology is designed the way it is.

For an introductory and brief overview of the methodology, please see the *IAT/ML White Paper*. If you are interested in a process-oriented description of IAT/ML, please see the *IAT/ML Analysis Process Guidelines*. If you are interested in a technical specification of IAT/ML, please see the *IAT/ML Technical Specification* document.

For more information on IAT/ML and additional documentation, please visit www.iatml.org.

Methodology Focus

IAT/ML is a methodology for discourse analysis. As such, it deals with discourse, understood as the practical and socially situated use of human language for communicative, pragmatic and symbolic purposes. Although discourse analysis in IAT/ML is approached mostly from the discipline of linguistics, other fields such as philosophy of language, ontology engineering or psychology are also relevant.

EMDA mini-theory

As a human activity, discourse implies the existence of **agents**, which include at least one speaker, who produces the discourse, and a receiver. The existence of agents, in turn, implies at least two additional phenomena: **mental states** and **actions** (see *Agency Analysis*, p. 18). Mental states include, for example, beliefs, emotions, desires, and intentions. Discourses produced by agents are based on their mental states, that is, we often say what we think, feel, want or plan to do. However, this doesn't mean that discourses faithfully follow our mental states all the time; in fact, people often speak words that do not match what is in their minds.

Similarly, the actions carried out by agents are often compatible with their mental states and discourses but, again, not always, as we sometimes do things that are not aligned with our thoughts or verbal commitments.

Reverse connections are also relevant. For example, behaving in certain ways usually compels us to produce certain discourses, and even perhaps to think in certain ways.

Mental states, discourses and actions occur within a given **environment**, composed of social and cultural elements that mediate what we think, say and do. Actions, in turn, modify this environment.

Mental states, discourses and actions, as well as the discrepancies between them, constitute human universals [5].

In this manner, environment-situated mental states, discourses, actions (EMDA for short) constitute the focus of IAT/ML, as shown in Figure 1.

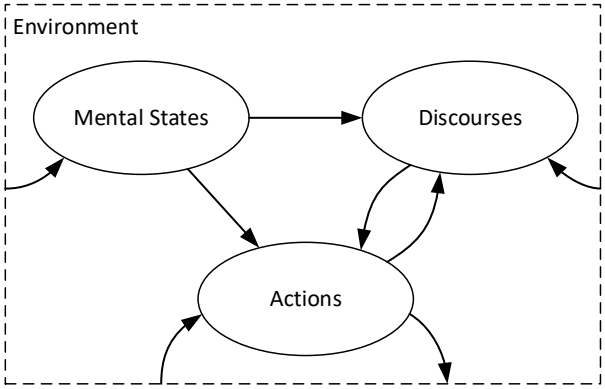

Figure 1. Environment, mental states, discourses and actions, plus the relations between them, constitute the focus of IAT/ML. Arrows represent influences.

Consider the following example. Imagine a person who believes that everyone should have the same rights and be treated equally regardless of their ethnicity, sex or religion. This person is likely to state these beliefs when asked. However, they may prefer locals as opposed to immigrants when looking for a carer for their children, perhaps due to mistrust and prejudice fuelled by mass media. This preference may be manifested as a systematic trend to hire only

locals despite of the availability of immigrant carers. When confronted with this by peers, this person may rationalise their behaviour by using a discourse that justifies their choices.

In this scenario, mental states, discourses and actions occur in inter-related manners, sometimes aligned, sometimes not. IAT/ML is designed to look at how discourses are aligned or misaligned with mental states, and to what extent actions are aligned or misaligned with mental states and discourses. By "aligned" here we mean that a manifestation is complete and truthful. For example, a fully aligned discourse is one that exposes everything that the agent thinks about something, and nothing that the agent doesn't think. Similarly, a fully aligned action is one that exercises everything the agent thinks and has said about something, and nothing that the agent doesn't think or hasn't said. Of course, alignment is gradual and nuanced, rather than binary.

In this manner, IAT/ML focuses primarily on analysing the discourses, and from them reaches into the mental states and actions of agents. In particular, ontology analysis focuses on representing the mental states (and, especially, the beliefs) that are revealed by discourses. Argumentation analysis, in turn, tries to understand how agents justify their claims and what strategies they use to support or attack other discourses. And, finally, agency analysis aims to shed light into agents' mental states and actions. This inclusive conception of discourse analysis, involving mental states and actions as well, is compatible with Gee's description of discourse in terms of "saying, doing and being" [11], where "saying" refers to the discourses themselves, "doing" to the actions, and "being" to the mental states of the agents. It also aligns with van Dijk's account of the relations between social cognition and discourse [10].

As a consequence, analysis in IAT/ML is tackled from four different perspectives, each one focusing on one element in the EMDA mini-theory: context (E), ontology (M), argumentation (D) and agency (A). Each of these is supported by its own theoretical grounding, described over the next sections.

Marker-phenomenon mini-theory

One of the goals of IAT/ML is to be rigorous and traceable. To accomplish this, it must produce results that are strongly anchored on the text being analysed. In this regard, it is assumed that linguistic phenomena manifest through markers in the text. For example, insistent use of the pronouns "us" and "them" may suggest that the speaker is trying to divide or differentiate their group from another one.

Consequently, analysis with IAT/ML often proceeds by identifying markers in the text and recognising the phenomena that they point to. Informally speaking, phenomena are "what is going on" in the discourse, whereas markers are the particular way in which phenomena manifest in the text.

Markers include linguistic elements such as:

- Lexical choices, e.g. using "sex worker" instead of "prostitute", or specific phrases such as "is a type of" or "is composed of"
- Modals, e.g. deontic constructions such as "we should…" or "it is imperative that…"
- Grammatical structures, e.g. using the passive voice over the active
- References, e.g. to temporal continuity (such as in "our forebears") or to victimhood (such as in "we are the victims here")
- Ethotic or pathetic targets such as "he's not reliable" or "you're going to love this"
- Metaphors, e.g. the boat or house metaphor such as "we're all on the same boat" or "this country is our house"
- Figures of speech, e.g. procatalepsy or hyperbole
- Absence of any of the former in a situation where it would be expected, e.g. omitting an ethotic statement when one is asked about the reliability of someone

Particular markers or combinations of markers suggest the presence of particular linguistic phenomena. **Phenomena** include:

- Ontological structures such as subtyping or instantiation
- Argumentation relations such as inferences or rephrases
- Rhetorical strategies such as *Discrediting* or *Portrait in black & white*

The relationship between markers and phenomena is many-to-many, that is, a phenomenon may manifest via multiple markers, and the same marker may indicate multiple phenomena. For example:

- The words “because”, “therefore” and “consequently” suggest that the speaker is probably making an inference.
- A *Discrediting* rhetorical strategy may manifest as a negative attribution in the present (e.g. “they are criminals”) but also as procatalepsy (e.g. “although conservatives will argue that raising taxes will stifle economic growth, it is crucial to recognise that it will benefit the country in the long run”).

The presence of a marker does not determine the presence of the associated phenomena. Rather, it works as a suggestion. The analyst must evaluate the marker itself, additional surrounding markers, and the context where they appear, in order to decide whether a particular phenomenon is present or not.

The marker-phenomenon mini-theory provides the basic concepts to anchor analysis decisions on to the text and thus produce results that are rigorous and traceable.

Context Analysis

As a socially situated phenomenon, discourse occurs in a context. For this reason, the first step towards understanding the context of a discourse is to describe the social milieu where it takes place, that is, the social issues or problems being discussed, the themes of the discourse, the positions being supported or opposed, and the agents doing this.

In addition, the support or opposition of a position by an agent is often mediated by identity and polarisation phenomena. Understanding identity and polarisation is thus crucial to understanding discourses and their contexts.

Groups

In this regard, a **group** is *a set of people united by common traits*. There are two types of groups. The first correspond to classical categories, in the sense that the traits that unite their members are properties shared by all of them. For example, the group of "inhabitants of Norwich" is given by all the people who share the fact of residing in Norwich. The second type of groups corresponds to Lakoff's radial categories [21], in which different pairs of members have something in common, but not necessarily the same thing. For example, the group "mothers" is given by women who have given birth, gestated, adopted, temporarily cared for, or lived with the father or mother of their children. Although their definitions are different, both types of groups function the same for the purposes of context analysis. Both groups and individual people are **agents**.

A group can be determined in several ways:

- Non-discursively:
    - Through immutable traits of people, such as their age, sex, or ethnicity. Some examples are "the elderly," "women," or "African-Americans".
    - Through mutable traits of people, such as their political affinity or their hobbies. Some examples are "labour voters" or "birdwatching enthusiasts".
- Discursively and, specifically, positionally, that is, through the exercise of a certain position. Some examples are "prostitution abolitionists" or "defenders of an open-borders policy".

Non-discursive groups, especially those based on immutable traits, can be usually determined by any observer. For example, it is relatively easy to determine whether a person belongs to the groups "the elderly," "women," or "African Americans", without the need for complex research. Determining group membership based on mutable traits is often more difficult, as these traits are often linked to properties of people's behaviour, which are difficult to infer from their looks or appearance. Finally, determining a person's membership of a positional group can only be achieved by analysing that person's discourses.

Often, a person belongs to many groups. For example, person P can belong simultaneously to the group "mothers", the group "inhabitants of Norwich", and the group "labour voters".

Groups can be combined by union and by intersection.

- Regarding the **union**, and following the previous example, we can say that P belongs to the group resulting from joining all the mothers, all the inhabitants of Norwich, and all the labour voters. The union of the groups to which a person belongs is manifested, for example, in his or her network of personal affinities; if we look at the contact list on P's mobile phone, for example, we will surely see people who are mothers (although they are not inhabitants of Norwich or labour voters), others who are inhabitants of Norwich

(although they are neither mothers nor labour voters), and others who are labour voters (although they are neither mothers nor neighbours of Norwich).

- Regarding the **intersection**, we can say that P belongs to the group resulting from the intersection of all the mothers, inhabitants of Norwich and labour voters, resulting in a much smaller group of people who meet all these criteria at the same time. The intersection of groups can serve to define the individuality of the person, since each person has a combination of groups that is unlikely to be found, in the same way, in most other people. This is the approach followed by intersectionality [8].

Positions

A **position** is *a proposition that is relatively well known in a group and for which there are supporters and opponents*. Every position is about a particular **topic**, which contextualises it. For example, the position "children of immigrant origin have more school problems due to racial discrimination" can be framed within the topic "immigration", or perhaps "racism", or even "education".

Positions are characterised by two variables: compatibility and prevalence.

- Regarding **compatibility**, there are positions that are compatible with each other, and others that are incompatible. For example, "children of immigrant origin have inferior school grades due to racial discrimination" is compatible with "racialised people are systematically discriminated against," but is incompatible with "children of immigrant origin do not obtain inferior academic results than those of local origin". To determine whether two positions are compatible or incompatible, their propositional contents must be evaluated, and it must be decided whether they contradict each other or not.
- Regarding **prevalence**, there are positions defended by many agents, while others are defended by very few. Prevalence indicates the degree of support that a position has within a group. To determine the prevalence of a position, it is necessary to carry out surveys or use other sociological techniques.

An agent's support of a position is accomplished by exercising it. An agent's **exercise** of a position is *the agent's commitment to that position at the level of beliefs, desires, intentions, and/or actions* (see *Agency Analysis*, p. 18).

Positions can be formulated in two modes:

- In **indicative** mode, such as "racialised people are systematically discriminated against". Positions like these can be exercised at the level of beliefs, but not desires, intentions or actions.
- In **deontic** mode, such as "children of immigrant origin should receive extra school care". Positions like these can be exercised at the levels of beliefs, desires, intentions and actions.

Exercising a deontic position at the level of desires, intentions or actions implies modifying the world (or wishing or planning to modify it) so that its state becomes closer to the propositional content of the position. The extent to which the world matches a deontic position's propositional content is called *correspondence*. For example, an agent that exercises the position "children of immigrant origin should receive extra school care" at the level of intentions is likely to attempt to improve its correspondence, that is, to change the world so that children of immigrant origin actually receive extra care at school.

In this manner, a position can be exercised by an agent to different levels of *engagement*, from least to greatest intensity:

- **Level 0**, when the agent does not believe in the position, nor, in the case of deontic positions, has any desires or intentions regarding it, nor do they carry out any actions to achieve it.
- **Level B**, when the agent *believes* in the position, but, in the case of deontic positions, has no desires or intentions regarding it, nor do they carry out any actions to improve its correspondence.
- **Level D**, when the agent has the *desire* to improve the correspondence of the deontic position, but has no intentions regarding it, nor does he/she carry out any actions to achieve it. This engagement level is independent of whether the agent believes in the position or not.
- **Level I**, when the agent has specific *intentions* to act to improve the correspondence of the position, but does not carry out such actions. This level is independent of the agent's beliefs and desires regarding the position.
- **Level A**, when the agent *acts* to improve the correspondence of the position. This level is independent of the agent's beliefs, desires, or intentions regarding the position.

Thus, a person **belongs** to a positionally determined group *if they exercise the corresponding position at least to engagement level B*. The higher the level to which they exercise the position, the stronger their belonging to the group.

There are sets of positions that tend to be supported or opposed as a whole. In other words, these sets of positions usually come as clusters or packages so that anyone supporting one position in a package is likely to support the other positions in the package. For example, if person P supports feminism, it is likely that they will also support green energy (as opposed to fossil fuels) and animal rights.

Thus, a **position package** is *a set of positions that are usually supported or opposed by the same agents as a block*. Positions in a package are often related to very different themes. Position packages constitute a common grounding for strong group identity. The concept of position package is related to that of ideology [10].

Identity and polarisation

We call **identity** to *the self-declared membership of a person in relation to a group or set of groups*. For example, if a person P says that they belong to the groups "mothers" and "inhabitants of Norwich", then we say that this person identifies with the mothers who live in Norwich. In this way, identity is defined in terms of group membership.

Often, different groups have different weights when it comes to shaping a person's identity. For example, a mother from Norwich who votes labour may feel strongly identified with other Norwich neighbours, but not as much with other labour voters.

Polarisation is any situation characterised by the following:

- Two or more groups support incompatible positions.
- Each group sees the positions of the others as being extreme, and do not contemplate a middle ground or a reconciliation of views.
- Emphasis is made on the positions that mark differences between groups rather than any common position that mark commonalities.
- The exercise of the incompatible positions involves appeals to group identity and judgments of value about the opposing groups.

Often, but not always, the positions that are object of polarisation are part of strongly identitarian packages.

Domination, resistance and acquiescence

A position is **dominant** *if its exercise necessarily causes harm to another agent and, in addition, the agent exercising it obtains a benefit directly caused by this harm*.

For example, the position "women should stay at home and do housework" is dominant, because its exercise harms the women it affects and, in addition, the male partners of these women benefit directly from this harm, as they are freed from the domestic tasks that their female partners carry out.

A position is **resistant** regarding another position *if it is incompatible with that other position, and that position is one of domination*.

For example, the position "women have the right to work outside the home" is resistant against "women should stay at home and do housework", because it is incompatible with it and the latter is dominant.

In order to exercise resistance, an agent does not have to be harmed by the corresponding domination. For example, men who defend "women have the right to work outside the home" are not necessarily harmed by the opposite position ("women should stay at home and do housework"), but they exert resistance to it just the same.

Finally, a position can be resistant and also dominant at the same time. For example, "the citizens of Gaza must be martyrs in the fight against Israel", as exercised by Hamas, dominates these citizens. But, at the same time, it resists against "the territory of Gaza must be Israeli".

A position is **acquiescent** to another position *if it is compatible with that other position, and that other position is dominant*.

For example, the position "women have the right to receive a salary in exchange for their work" is acquiescent to "women should stay at home and do housework," which is dominant, since they are compatible with each other.

If the two positions belong to different themes, the acquiescent one is said to be *trivially acquiescent*, as it has a different domain of discourse as compared to the dominant position. For example, "women have the right to work outside the home" is trivially acquiescent in relation to "the territory of Gaza must be Israeli".

If the two positions belong to the same theme, however, the acquiescent one is said to be *significantly acquiescent*, because the two share a common domain of discourse.

Context of discourse

In this manner, the context of a discourse is composed of the issues being addressed, the themes that are being discussed, the positions being supported or opposed, and the agents that do it.

Ontology Analysis

Discourse has content, that is, it is *about* something. In other words, discourses refer to things in the world via representation or, more linguistically, though semantics [26]. Understanding what things a discourse refers to, what properties of these things are mentioned, and how these things are portrayed to relate to one another is crucial to proper discourse analysis.

In this regard, the concept of ontology becomes useful. The word "ontology" has a long tradition in philosophy; however, it is also used in computer science and information systems with a different meaning. According to the influential [16], an ontology is an explicit specification of a shared conceptualisation, that is, a formal (or semi-formal) model of the set of concepts that a group agrees to. Ontologies in this sense are closely related to conceptual models [22], and it has been argued that they are essentially equivalent although they emphasise different things [12], [17].

Ontology analysis in IAT/ML builds upon the vast literature in conceptual modelling and ontology engineering, as described in the next sections.

Types and tokens

The basic and foremost kind of things that a piece of text may be about is **entities**. An entity is *a thing in the world that we can distinguish from others* [13]. Traditionally, two kinds of entities are distinguished: types and tokens [30]. Types, also called universals, classes or categories, are entities that work as "templates" that can be instantiated into other entities. For example, the concept of *Cat* in my mind is a type, as it doesn't refer to any particular cat but to the overall concept of cat. When I say things like "cats are fluffy and beautiful", I am not referring to any specific cat, but to the overall *Cat* type. In IAT/ML, types are called **categories**.

On the other hand, tokens, also called individuals, particulars, instances, objects or atoms, are things that cannot be instantiated into others. For example, when I talk about my cat Valentina, I am referring to an individual cat rather than the *Cat* category. When I say things like "Valentina is a bit grumpy to strangers", I am referring to a token. In IAT/ML, tokens are called **atoms**.

The kinds of relationships that connect types and tokens have been much debated in ontology engineering and conceptual modelling. Traditionally, the orthodoxy has been that each token is an instance of a type (instantiation), and each type can be a sub-type of one or more types (subtyping or subsumption). This means that tokens can be instances of types, but types cannot be instances of other types, but just subtypes. More recently, the trend known as *multi-level modelling* has become popular [1], and advocated for a more powerful and expressive scheme by which types can be instances of other types as well. This is the approach adopted by IAT/ML.

In summary, and regarding IAT/ML:

- The world is composed of entities.
- There are two kinds of entities: categories and atoms.
- An entity (either an atom or a category) can be an instance of a category. For example, the *Valentina* atom is an instance of the *Cat* category, and the *OakTree* category is an instance of the *TreeSpecies* category.
- A category, in addition, can be a subtype of another category. For example, the *Cat* category (together with *Dog*, *Cow*, etc.) is a subtype of the *Animal* category which, in turn (and together with *Plant*, for example), is a subtype of the *LivingBeing* category. Categories, in this manner, can be organised into subtyping hierarchies.

Features and facets

Entities can represent things in the world, but not their properties. To achieve this, IAT/ML adds facets. A **facet** is a concept that represents a predication on an entity. For example, when I say that "Valentina is quite grumpy", I am predicating something (being grumpy) on an entity (the atom *Valentina*). Similarly, when I say that "Valentina is my pet", I am also predicating something (belonging to me) on the same entity. In this manner, there are two kinds of facets. **Values**, on the one hand, correspond to a quantity or quality, such as being grumpy, red or beautiful. **References**, on the other hand, correspond to a directed relationship to another entity, such as belonging to me (*me* is the other entity) or being located in France (*France* being the other entity).

What can be predicated on an entity depends on its type. For example, it doesn't make sense to predicate a colour on an idea, or a mood on a rock. In this manner, the facets that an entity may have are regulated by the features of its type category. A **feature** is a concept that represents a type of predication on entities of a given category. For example, the *Cat* category may have an *Age* feature, because it makes sense to predicate ages on cats. Similarly, it may have a *BelongsTo* feature, because cats usually belong to someone. Features like *Age*, which are instantiated in the form of values, are called **properties**. And features like *BelongsTo*, which are instantiated in the form of references, are called **connections**. Connections, in addition, specify what category is the target of its potential references. In this manner, we could say that the *Cat* category has two features: the *Age* property and the *BelongsTo* reference towards Person.

Once features are in place, facets become easier to understand. For example, *Valentina.Breed = Siamese* is a value that predicates the Siamese breed on to the *Valentina* atom. Since *Valentina* is of type *Cat*, and *Cat* has a *Breed* property, it makes sense to assign a breed value to *Valentina*. We could not possibly assign a breed value to an atom of type *Building* or *City*, for example, as neither of these categories are likely to have a *Breed* property.

References work in pairs. For example, if Valentina the cat holds the reference *BelongsTo* towards *Alice*, then we can say that *Alice* holds the reference Owns towards *Valentina*. In other words, the existence of a reference implies the existence of an inverse in the opposite direction. A pair of references like this is called a **link**. At the category level, connections work in an analogous way: if *Cat* has a *BelongsTo* connection towards *Person*, then *Person* has a Owns connection towards *Cat*. A pair of connections that work together like this is called an **association**.

Entities, features and facets are collectively called *concepts*. Concepts constitute the basic building blocks from which we can describe almost anything. An ontology in IAT/ML is thus composed of concepts. Most conceptual modelling languages such as ConML [13], [18], many programming-oriented languages such as UML [23], and many semantic web technologies such as OWL [32] adopt a similar structure for the description of the things in the world.

Temporality and subjectivity

Any model of the world becomes obsolete when the things that it represents change. For example, an ontology that states that Valentina the cat belongs to Alice will become obsolete when Valentina moves to live with Bob. In other words, an ontology represents the world at a given time, but may not be valid at other times, unless we take extra precautions. Similarly, an ontology represents the world according to someone, but may be invalid according to someone else. For example, an ontology stating that the Colosseum in Rome is very well maintained may be valid for those who agree with that statement, but invalid for others.

To cater for the passage of time and the different points of view of people, ontologies in IAT/ML make use of *existence* and *predication qualifiers* [14]. A qualifier is a piece of text that is added to an entity or a facet to qualify it in terms of temporality of subjectivity:

- A temporality qualifier on an entity indicates when the entity exists. For example, if the atom *vp* represents the Viceroyalty of Peru, then we could attach the temporal qualifier "1542–1824" to indicate that this political entity existed only during that period.
- A subjectivity qualifier on an entity indicates according to whom the entity exists. For example, the atom *Shiva*, representing the Hindu god of the same name, may have attached the subjective qualifier "believers in Hinduism" to indicate that Shiva exists only according to this group.
- A temporality qualifier on a facet indicates when the predication applies. For example, we could attach "June 2017 to September 2020" to the *Alice.LivesIn = London* reference to indicate that this is the period of time during which Alice lived in London.
- A subjectivity qualifier on a facet indicates according to whom the predication is valid. For example, we could attach "Kremlin" to the *Crimea.BelongsTo = Russia* reference to indicate that Crimea belongs to Russia according to the Kremlin but probably not others.

When writing a temporal qualifier after an entity or facet, we use the @ separator (read "at"). For example, *Alice.LivesIn @ June 2017 to September 2020 = London* indicates that Alice lived in London only during that period. Similarly, we use the $ separator (read "according to") when writing a subjective qualifier after an entity or facet. For example, *Shiva $ believers in Hinduism* indicates that Shiva exists only according to that group.

Temporality qualifiers are useful to express the passage of time in an ontology. For example, you can have entities with values for the same property and different moments or periods to represent how the entity changed over time. Similarly, subjectivity qualifiers are useful to represent the points of view of different agents. You can have, for example, entities with values for the same property and different agents to indicate disagreements or conflicts, all within the same ontology.

Denotations

Conventional approaches to semantics tell us that terms denote concepts. In terms of IAT/ML, this means that some of the words that make up the texts that we analyse point to concepts in the ontology. In fact, if an ontology contains a particular concept it is likely due to the fact that a term pointing at that concept exists in the text. For example, an ontology constructed from a text that includes terms such as "Israel" and "Palestine" is likely to contain atoms for *Israel* and *Palestine*, among others.

A **denotation** in an IAT/ML ontology is a connection between a term in the text and a concept in the ontology. Denotations capture semantics. For example, you can describe a denotation that connects the term "Palestine" in the text to the *Palestine* atom. You can also record a denotation that connects other terms, such as "the Palestinian territories" or "the Promised Land" to the same concepts, thus reflecting the fact that they denote the same thing.

Reference ontologies

When carrying our ontological analysis, you can start from scratch and create your ontology by finding relevant terms in the text and constructing concepts from them. If you create ontologies about the same topics repeatedly, you will find that this is a cumbersome task, often repetitive. In this regard, it is worth having a general ontology for your topic, and customising it as needed for each particular text by adding, modifying or removing concepts.

After a long time doing ontological analysis on a particular topic, it is likely that common patterns and concepts will emerge. For example, almost every text that discusses immigration is likely to mention migrants, their countries of origin, their countries of destination, and the reasons why immigration happens. Instead of re-creating these concepts and the relationships between them every time you develop an ontology about immigration, you can have them ready in the form of a reference ontology. A **reference ontology** is a very abstract representation of the major concepts in a particular domain of discourse, intended to be used via refinement. Starting from a reference ontology instead of from scratch is good, because it saves time and gives ontologies a common structure for better interoperability.

Different authors have proposed reference ontologies for a number of domains. For example, CHARM [13], [19] and CIDOC CRM [7] are well known reference ontologies in the field of cultural heritage. When carrying out ontology analysis in this field, you can start from one of these instead of from scratch.

Furthermore, if you work with a complex corpus that makes extensive use of topics and sub-topics, then you can gradually refine your reference ontology, so that you have a particular ontology for each hierarchical level in the corpus: the reference ontology for the corpus as a whole, a refined version for each main topic, a further refined version for each sub-topic, and so on until you reach very specific ontologies for each document. This approach, called gradual refinement of models, has been proven to work well in different scenarios [15].

Argumentation Analysis

Discourse often makes arguments, that is, presents reasons why certain claims are said to be true. Argumentation, in this regard, is about how speakers justify their claims, or attack others [27]. Understanding what argumentation devices are used, what claims are supported or attacked, and how different statements relate to each other in terms of argumentation relations is crucial to proper discourse analysis.

A well-known approach to argumentation analysis is the Inference Anchoring Theory (IAT) [20], [24], which emphasises the need to anchor argumentation structures to elements in the text. Argumentation analysis in IAT/ML builds upon IAT, also relying on ontology analysis (see *Ontology Analysis*, p. 11), as described in the next sections.

Locutions and transitions

Locutions and transitions describe segments of the text that operate as argumentative units. For example, the text "I am happy because it's sunny" would be divided into two locutions, "I am happy" and "it's sunny", plus a transition indicating the argument being made and anchored to the "because" marker. In this regard, a **locution** in IAT/ML is a segment in the text representing an utterance made by a speaker. Locutions usually correspond to sentences or clauses within sentences. Dividing a text into locutions is a complex task; please see the see the *IAT/ML Analysis Process Guidelines* and *IAT/ML Argumentation Patterns Guidelines* for a comprehensive explanation.

A **transition**, in turn, corresponds to a discursive relationship between locutions. In the previous example, a transition is employed to describe the fact that a speaker first said "I am happy", and then added "it's sunny". IAT/ML contemplates different types of transitions, such as "adding", "reporting" or "turn-taking", depending on the nature of the discursive connection between locutions. Please see the *IAT/ML Analysis Process Guidelines* or the *IAT/ML Technical Specification* for a comprehensive list of transition types.

A simple transition connects an initial to a final locution, but more complex structures may exist. For example, a transition may involve multiple final locutions. Also, a locution may have multiple outgoing (but not incoming) transitions. An example of the latter would be "I'm happy because it's sunny and I'm on holiday"; here, "I'm happy" is connected to "it's sunny" via a transition, and to "I'm on holiday" by another transition, because "it's sunny" and "I'm on holiday" constitute two reasons that the speaker is providing to support the premise "I'm happy" and, from an argumentative point of view, they could have been provided in any order.

In any case, transitions must be compatible with the time ordering of the discourse. In other words, the final locutions in a transition cannot be prior to the initial locution.

Propositions

In IAT/ML, a **proposition** is a statement about a state of affairs in the world. Propositions are worded as canonical and self-contained sentences so that they can be maximally understood in isolation. Propositions are *reconstructed* from locutions by the analyst, that is, the analyst must reword a locution so that it makes up a canonical and self-contained sentence that expresses the same, resolving deictic particles or anaphora. For example, the locution "I am happy" as uttered by Alice could be reconstructed as a proposition "Alice is happy". Reconstructing propositions is a complex task that is comprehensively described in the *IAT/ML Argumentation Patterns Guidelines*.

Typically, each locution is reconstructed into one proposition, although there may be locutions that yield no propositions (e.g. interjections or other utterances lacking argumentative power) or multiple propositions (such as loaded questions, which provide both an implicit statement plus a question).

Propositions constitute the core of argumentation analysis in IAT/ML, and can be characterised through many variables such as their factual and ontological aspects, their modality or their truth value.

Argumentation relations

Propositions work as nodes in a mesh of **argumentation relations**, which connect propositions to describe argumentation devices of three kinds: inferences, conflicts and rephrases.

In IAT/ML, an **inference** is an argumentation relation involving one or more premises (propositions) that are provided by a speaker to support a conclusion (another proposition). For example, the text "I am happy because it's sunny" would be reconstructed as two propositions connected by an inference, so that "Alice is happy" would be the conclusion and "It's sunny" would be the premise. IAT/ML supports many types of inferences, as described by [28], [29].

A **conflict**, on the other hand, is an argumentation relation involving a source proposition that is provided by a speaker to contradict, disagree or negate a target argumentation unit, whether this is another proposition or an argumentation relation. In this regard, a conflict may target a proposition, such as in "Alice: It's warm today; Bob: Well, I am cold", but also an inference, rephrase, or even another conflict. For example, in the text "Alice: My friend's daughter got autism because she was vaccinated; Bob: But vaccines don't cause autism", there is conflict from Bob's proposition towards the inference that Alice makes.

Finally, a **rephrase** is an argumentation relation involving a source proposition that is provided by a speaker as a reformulation of another proposition. For example, in "Alice: The president gave a glorious speech; Bob: Yeah, she was amazing", Bob is rephrasing what Alice said. Rephrasings can be of many types, depending on the nature of the reformulation. Common types include concretion, answer, agreement or paraphrasis. Please see the *IAT/ML Analysis Process Guidelines* or the *IAT/ML Technical Specification* for a comprehensive list of rephrasing types.

Illocutionary forces

Speech act theory [2], [25] establishes that utterances may involve a locutionary act (the actual utterance and its apparent meaning), an illocutionary act (the actual meaning of the utterance, as derived from the speaker's intention), and a perlocutionary act (the consequences of the utterance). Given the central role that intentions play in IAT/ML (see EMDA mini-theory, p. 4), the illocutionary acts of utterances are recorded in the form of illocutionary forces. In this regard, an **illocutionary force** in IAT/ML is a connection between a discourse element (a locution or transition) and an argumentation element (a proposition or argumentation relation) in terms of speaker intent. Illocutionary forces are the device that allows IAT/ML to anchor argumentation on to the actual text.

Two kinds of illocutionary forces can be distinguished, depending on whether they are anchored to locutions or transitions. **Locution-anchored illocutionary forces** express the intention of the speaker when issuing an utterance, and connect the locution to the corresponding proposition. There are four types: asserting, questioning, challenging and popular conceding. Similarly, **transition-anchored illocutionary forces** express the intention of the speaker when making a transition in the discourse, and connect the transition to the corresponding argumentation

relation (inference, conflict or rephrase). There are four types as well: arguing, agreeing, disagreeing and restating.

Denotations

Denotations in argumentation models work in a very similar way as denotations in ontologies (see *Denotations*, p. 13). A **denotation** in an IAT/ML argumentation model is a connection between a term in a locution and a concept in the associated ontology. For example, you can describe a denotation that connects the term “Rome” in the locution “Trump visited Rome in 2017” to the *Rome* atom in the ontology. You can also record a denotation that connects other terms, such as “the Eternal City” or “the capital” to the same concept, thus reflecting the fact that they denote the same thing.

In this manner, denotations work as connections between an argumentation model and an ontology, and allow their joint treatment.

Agency Analysis

As discussed as part of the EMDA mini-theory (see *EMDA mini-theory*, p. 4), discourse is produced by agents. Agents, furthermore, employ discourse in order to convey their mental states and related concerns. In this manner, discourse can unveil, at least potentially, the mental states of the participant agents.

Critical Discourse Analysis (CDA) [9] or, more recently, Critical Discourse Studies (CDS) [6], are well known approaches that have contributed to gaining insights into the speaker's mental states. However, they have also been criticised as being too reliant on the subjectivity of the analyst and therefore not too reliable [4], [31]. To alleviate these issues with CDA and CDS, agency analysis in IAT/ML adopts some ideas from them, but mostly constitutes an original development, as described in the next sections.

Agents and agency

In IAT/ML, *agency* is the combination of:

- The capability for mental states, at least beliefs, desires and intentions.
- The capability of action according to these mental states.

Mental states may also include perceptions, emotions or memory, but these are not required for agency.

In turn, an agent is anything that has agency. In other words, an **agent** is *something that has beliefs, desires and intentions, as well the capacity to act according to them*.

The conceptualisation of agency in terms of beliefs, desires and intentions (BDI) was popularised by [3] and later incorporated into the development of multi-agent systems. Although not part of the original BDI triad, actions are also part of the definition of agents in IAT/ML, so the four elements are combined as BDI+A.

Beliefs

A **belief** is *a subjective attitude of an agent in relation to the veracity of a proposition*. In other words, an agent believes P if they possess an attitude towards P's truth value. For example, an agent believes that "Paris is the capital of France" if they have a clear position on whether this is true or false.

Beliefs are closely related to exercising positions to engagement level B (see *Positions*, p. 8). In this regard, possessing a belief is equivalent to exercising the corresponding position to level B.

Desires

A **desire** is *a counterfactual state of the world that is judged to be good by an agent*. For example, an agent desires that "Quebec is an independent country" if Quebec is not actually an independent country (as in fact it isn't) and the agent judges Quebec being an independent country as something good.

Desires are closely related to exercising deontic positions to engagement level D (see *Positions*, p. 8). In this regard, possessing a desire is equivalent to exercising the corresponding position to level D. For example, an agent that exercises "Quebec should be an independent country" to level D actually possesses the desire that "Quebec is an independent country".

Intentions

An **intention** is *a plan held by an agent to modify the world so that the correspondence of a particular deontic position is improved*. For example, an agent has the intention to alleviate hunger in the world if they plan to make a donation to an NGO working on feeding the homeless, so that the correspondence of position "There shouldn't be hungry people in the world" is improved.

Intentions are closely related to exercising deontic positions to engagement level I (see *Positions*, p. 8). In this regard, possessing an intention is equivalent to exercising the corresponding position to level I. For example, an agent that exercises "There shouldn't be hungry people in the world" to level I actually possesses the intention to bring the world closer to this ideal.

Actions

An **action** is *a behaviour by an agent that changes the external world*. For example, an agent that donates money to an NGO to fight against hunger is carrying out an action.

Actions are closely related to exercising deontic positions to engagement level A (see *Positions*, p. 8). In this regard, and since an action may improve the correspondence of a particular position, carrying out an action may be equivalent to exercising the position to level A. For example, donating to the above-mentioned NGO probably improves the correspondence of position "There shouldn't be hungry people in the world", because it changes the world in a manner that makes it closer to the propositional content of the position.

References

[1] J. P. A. Almeida, U. Frank, and T. Kühne, "Multi-Level Modelling (Dagstuhl Seminar 17492)," Wadern, Germany, 2018. doi: 10.4230/DagRep.7.12.18.

[2] J. L. Austin, *How to Do Things with Words*, Reprint. Martino Fine Books, 2018.

[3] M. E. Bratman, *Intention, Plans, and Practical Reason*. CSLI Publications, 1999.

[4] R. Breeze, "Critical Discourse Analysis and its Critics," *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, pp. 493–525, Jul. 2022, doi: 10.1075/prag.21.4.01bre.

[5] D. Brown, *Human Universals*. McGraw-Hill, 1991.

[6] T. Catalano and L. R. Waugh, *Critical Discourse Analysis, Critical Discourse Studies and Beyond*, vol. 26. Cham: Springer International Publishing, 2020.

[7] CIDOC, "The CIDOC Conceptual Reference Model," 2011. http://www.cidoc-crm.org/.

[8] B. Cooper, "Intersectionality," in *The Oxford Handbook of Feminist Theory*, L. Disch and M. Hawkesworth, Eds. Oxford University Press, 2015.

[9] T. A. Van Dijk, "Critical Discourse Analysis," in *Handbook of Discourse Analysis*, D. Tannen, D. Schiffrin, and H. Hamilton, Eds. Blackwell, 2001, pp. 352–371.

[10] T. van Dijk, *Discourse and Context: A Sociocognitive Approach*. Cambridge University Press, 2008.

[11] J. P. Gee, *An Introduction to Discourse Analysis: Theory and Method*. Routledge, 2014.

[12] C. Gonzalez-Perez, "How Ontologies Can Help in Software Engineering," in *Grand Timely Topics in Software Engineering*, vol. 10223 LNCS, no. 10223, J. Cunha, J. P. Fernandes, R. Lämmel, J. Saraiva, and V. Zaytsev, Eds. Springer, 2017, pp. 26–44.

[13] C. Gonzalez-Perez, *Information Modelling for Archaeology and Anthropology*. Springer, 2018.

[14] C. Gonzalez-Perez, "Conceptual modelling of temporality and subjectivity as cross-cutting concerns," *Softw Syst Model*, Dec. 2024, doi: 10.1007/s10270-024-01247-0.

[15] C. Gonzalez-Perez and P. Martín-Rodilla, "Integration of Archaeological Datasets through the Gradual Refinement of Models," in *21st Century Archaeology: Concepts, Methods and Tools - Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology*, F. Giligny, F. Djindjian, L. Costa, P. Moscati, and S. Robert, Eds. Archaeopress, 2015, pp. 193–204.

[16] T. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.

[17] B. Henderson-Sellers, "Bridging Metamodels and Ontologies in Software Engineering," *Journal of Systems and Software*, vol. 84, no. 2, pp. 301–313, 2011, doi: doi:10.1016/j.jss.2010.10.025.

[18] Incipit CSIC, "ConML Technical Specification," Incipit CSIC, 2020. [Online]. Available: http://www.conml.org/Resources/TechSpec.aspx.

[19] Incipit CSIC, "CHARM Web Site," 2016. http://www.charminfo.org (accessed Oct. 09, 2020).

[20] M. Janier, M. Aakhus, K. Budzynska, and C. Reed, “Modeling argumentative activity with Inference Anchoring Theory,” in *Argumentation and Reasoned Action. Volume I Proceedings of the 1st European Conference on Argumentation*, vol. 1, no. 62, D. Mohhamed and M. Lewinski, Eds. College Publications, 2016.

[21] G. Lakoff, *Women, Fire, and Dangerous Things*. University of Chicago Press, 1990.

[22] A. Olivé, *Conceptual Modeling of Information Systems*. Springer, 2007.

[23] OMG, “Unified Modeling Language 2.5.1.” 2017, [Online]. Available: https://www.omg.org/spec/UML/.

[24] C. Reed *et al.*, “The Argument Web: An Online Ecosystem of Tools, Systems and Services for Argumentation,” *Philos Technol*, vol. 30, no. 2, pp. 137–160, Jun. 2017, doi: 10.1007/s13347-017-0260-8.

[25] J. R. Searle and D. Vanderveken, *Foundations of Illocutionary Logic*. Cambridge University Press, 1985.

[26] J. Speaks, “Theories of Meaning,” in *The Stanford Encyclopedia of Philosophy*, Winter 2024., E. N. Zalta and U. Nodelman, Eds. Metaphysics Research Lab, Stanford University, 2024.

[27] S. E. Toulmin, *The Uses of Argument*. Cambridge University Press, 2003.

[28] J. Visser, J. Lawrence, C. Reed, J. Wagemans, and D. Walton, “Annotating Argument Schemes,” *Argumentation*, vol. 35, no. 1, pp. 101–139, Mar. 2021, doi: 10.1007/s10503-020-09519-x.

[29] D. Walton, C. Reed, and F. Macagno, *Argumentation Schemes*. Cambridge University Press, 2008.

[30] L. Wetzel, “Types and Tokens,” *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, 2018, [Online]. Available: https://plato.stanford.edu/archives/fall2018/entries/types-tokens/.

[31] H. G. Widdowson, “Discourse analysis: a critical view,” *Language and Literature: International Journal of Stylistics*, vol. 4, no. 3, pp. 157–172, Aug. 1995, doi: 10.1177/096394709500400301.

[32] World Wide Web Consortium, “OWL 2 Web Ontology Language.” World Wide Web Consortium, 2012, [Online]. Available: http://www.w3.org/TR/2012/REC-owl2-overview-20121211/.