Classification of Modal Meaning in Negotiation Dialogues

Valeria Lapina[1] and Volha Petukhova[2]
Computational Linguistics[1], Spoken Language Systems[2], Saarland University, Germany
vlapina@coli.uni-saarland.de, v.petukhova@lsv.uni-saarland.de

Abstract

This paper addresses modality classification for multi-issue bargaining dialogues in order to model human-like negotiation behaviour and to efficiently compute negotiation strategies. We propose a modality annotation and classification scheme comprising semantically distinguishable categories applied reliably by humans and machines. Our classification of modality varieties is based on both the traditional dichotomy of epistemic and root modalities, and on the analysis of the available corpus data. Our modality scheme has been used for annotating human-human dialogues and training SVM-based classifiers. We built predictive models that show accuracies in the range between 73.3 and 82.6%.

1 Introduction

In any communicative situation, interlocutors communicate their beliefs, desires, expectations, interests and obligations by means of certain communicative actions, i.e. dialogue acts. These actions are used by the speaker to signal his or her intentions concerning events, objects, relations, properties involved in the communicative situation. Speaker's intentions can be rather complex, vague and ambiguous. They may also be emotionally qualified expressing particular attitudes towards their communicative partners, third parties and message content. In negotiation interactions, partners do not just negotiate through a sequence of offers. It is observed that negotiators actually rarely make concrete offers as binding commitments (Raiffa et al., 2002). Rather, participants' actions are often focused on obtaining and providing preference information. In multi-issue bargaining, a special form of negotiation, parties have the possibility to simultaneously bargain about several goods and attributes. In interest-based (win-win) bargaining, interlocutors search for integrative potential (Fisher and Ury, 1981). They have partially competitive and partially cooperative goals, conflicting, identical or partly overlapping preferences. All this allows bargainers to have a wide array of strategies. Such strategies are often communicated in natural language by means of various modal expressions. The present study addresses modality classification for multi-issue bargaining dialogues with the purpose to adequately model human-like negotiation behaviour and to efficiently compute negotiation strategies. The main objective of this study is to establish a reliable modality classification model for negotiation domain.

The paper is structured as follows. Section 2 defines modality and its types. Section 3 provides an overview of previous annotation and classification efforts. Section 4 focuses on the description of a multi-issue bargaining scenario and the role of modality in it. Section 5 addresses machine-learning data-oriented approach to modality classification by discussing experiment design, presents the obtained results and their analysis. Section 6 brings it all together by specifying the semantics of negotiation actions, extending the ISO 24617-2 dialogue act specifications with the specification of semantic content in terms of negotiation moves and their arguments. The ISO 24617-2 dialogue act metamodel is extended accordingly. Finally, we summarize our findings and outline future research directions.

2   Defining modality and its types

Linguistic modality is an omnipresent phenomenon in communication that, broadly speaking, is concerned with the speaker's subjective beliefs. Modality corresponds to the speaker's evaluation of probability of events; it concerns with what the speaker believes to be possible, necessary or desirable.

Utterances that express the speaker's subjective beliefs are modalised. A modalised utterance has a propositional and a modal content. For example, in *it must be raining* the propositional content is *it is raining*, while the modal content suggests that it is a personal perception expressing a high degree of certainty: *it must be*. If we apply a construction-centred approach (Ghia et al., 2016) to modalised utterance analysis, we may say that [It must] is a *trigger*, [be raining] is its *target*, and between the *holder* (the speaker of an utterance in this case) and the target there is a *modal relation* as depicted in Figure 1. In modal logic, a trigger corresponds to an operator, that semantically qualifies the truth of an utterance in its scope, e.g. ◇ stands for ability, □ for preference, etc. Any modalised utterance can be described in these terms. Modality triggers can be expressed verbally, prosodically, and multimodally.
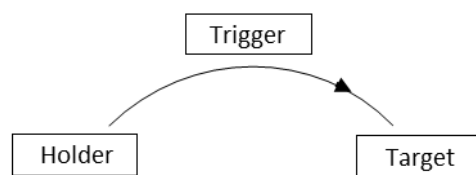

Figure 1: Modal relation between the *holder* and the *target*.

In linguistics, modality has been extensively studied by Lyons (1970), Palmer (1979), Calbert (1975), Kratzer (1981), Leech (1983), Bybee and Fleischman (1995), and many others. Previous studies distinguish between *epistemic* and *root* modalities.

Epistemics deal with possibilities that follow from the speaker's knowledge, whereas roots deal with possibilities that follow from the circumstances surrounding the main event and its participants; epistemics are taken to be speaker-oriented, roots subject-oriented, see Bybee et al. (1994). An epistemic modality (Lyons, 1970; Palmer, 1979; Nuyts, 2001) is generally understood as a weak commitment of the speaker to the truth of the proposition based on evidence or personal beliefs.

There are several types of root modalities distinguished. For instance, a *deontic* (Palmer, 1979; Peters et al., 2009; Nuyts, 2001) modality is concerned with rules, norms, and principles of either ethical or legal nature and it often expresses permissions or obligations. Scholars commonly distinguish between deontic and *dynamic* modalities (Palmer, 1979). Dynamic modality corresponds to expression of possibility that does not depend on rules, but either on the laws of nature (for example, physics) or personal abilities of the modality holder. Dynamic and deontic modal meanings may share common triggers, such as *can* and *possible*.

Modality that is concerned with expressing liking or disliking an event is called *bouletic* (Rubinstein et al., 2013) or *boulomaic* (Kratzer, 1981; Nuyts, 2001). The modal meaning that corresponds to expressing one's goals is referred to as *teleological* modality. In practice, it can be challenging to distinguish between boulomaic and teleological modalities, that is why some scholars (Portner, 2009; Rubinstein et al., 2013) propose to treat modal meanings that generally express speaker's priorities (both goals and preferences) as *prioritizing* modality. In addition, *volitive* modality indicates the desires and intentions of the speaker.

To sum up, most modality classification approaches account for *epistemic, deontic, dynamic, prioritizing (boulomaic, teleological)*, and *volitive* modal meanings.

3   Modality classification

To be practically useful, modality taxonomies need to facilitate reliable human and machine annotations. Existing automatic modality classification approaches rely on classification schemes of different granularity and complexity levels, with specific domain properties and constraints. For example, Nirenburg and Raskin (2004) proposed modality taxonomy to classify attitudinal propositional meaning. Their classification includes seven classes: epistemic, deontic, volitive, potential, epiteuctic (describing success of an event), evaluative, and saliency (highlighting important information).

	Nirenburg & Raskin (2004)	Medlock & Biscoe (2007)	Kilicoglu & Bergler (2007)	Baker et al. (2010)	Rubinstein et al. (2013)	Kobayakawa et al. (2009)	Lavid et al. (2016)
Epistemic	v	v	v	v	v	v	v
Deontic	v			v	v	v	v
Dynamic				v	v	v	v
Volitive	v			v		v	v
Bouletic				v	v	v	
Teleological					v		
Epiteuctic	v			v			
Evaluative	v						
Saliency	v						
Circumstantial					v		

Table 1: Overview of modality categories defined in various annotation schemes.

Medlock and Briscoe (2007) developed an annotation scheme for automatic classification of hedging (i.e. speculative language) in scientific texts. This annotation comprises two categories: speculative and non-speculative. The authors admit that even for human annotators it is a challenge to distinguish speculative from non-speculative sentences. The annotation scheme was applied to automatic modality classification using weakly supervised learning methods, achieving 75% accuracy.

Kilicoglu and Bergler (2008) designed a modality classification scheme to analyse biomedical texts. Their classification included the epistemic modality and its subcategories following Palmer: speculatives (uncertainty), deductives (inference), assumptives (inference from what is generally known). The highest accuracy of 93% was achieved based on syntactic features.

Kobayakawa et al. (2009) used a fine-grained annotation scheme and trained an SVM-based classifier to predict modality classes. Their classification includes 18 categories, some of which correspond to sub-categories of epistemic, deontic and dynamic modalities such as request, recommendation, will, wish, judgment, unnecessary, permission, possible. Other categories are more language and domain specific, e.g. unexpected, meaningless, hearsay, emphasis, admiration, duty, properness, qualifiedness, tentative, and natural occurrence. The classifier showed an accuracy of 78% on this multi-class classification task.

The scheme for modality annotation proposed by Baker et al. (2010) includes eight main categories: permissive, success, effort, intention, ability, want and belief. A string-matching tagger enriched with syntactic patterns has been implemented and showed a reasonable performance result in terms of precision (86%).

Rubinstein et al. (2013) proposed a hierarchical taxonomy inspired by Kratzer (Kratzer, 1981). The scheme defines three main coarse-grained categories (*epistemic, ability, priority*) and seven fine-grained classes: epistemic, circumstantial, ability, deontic, bouletic, teleological, and bouletic/ teleological. It has been observed that annotation was challenging for human annotators due to the ambiguity of modal verbs. The measured inter-annotated agreement was rather moderate in terms of Krippendoff's α .49.

Lavid et al. (2016) propose a linguistically-motivated annotation model of modality in English and Spanish. Their annotation scheme is hierarchical and comprises a core tagset (epistemic, deontic, dynamic, and volitional modal meanings) and a two-tiered extended tagset that specifies each core modality. The epistemic modality is subdivided into evidential (perception, cognition, and communication) and non-evidential (possibility, probability, certainty, doubt, apprehension). The deontic modality in the extended tagset is divided into obligation, recommendation, permission, prohibition, and absence of obligation. The dynamic modality is represented as necessity, tendency, and possibility which may mean either ability or situational possibility. The volitional modality in the extended tagset can either mean willingness or acceptance. The authors report having achieved a high inter-annotator agreement (the Cohen's kappa coefficient is 0.854).

The proposed schemes differ with respect to the number, level and nature of the defined concepts (see Table 1). However, most manual and automatic classification efforts showed that annotation success does not depend so much on the scheme complexity and granularity, but rather on the clarity and semantic distinctiveness of the defined concepts. Well-defined categories facilitate effective human and machine classification process, the annotation should fulfill certain criteria.

The design of an annotation scheme should be based on the principle of *semantic adequacy*, which

Linguistic modality	Annotation	Definition	Example
Prioritizing	Preference	agent A expresses that he is in favour of action α[1]	I *like* anti-smoking television advertisements
Prioritizing	Dislike	agent A expresses that he is not in favour of action α	This is even *worse* for me
Prioritizing	Necessity	agent A expresses that action α is necessary for him	I *have to* have at least all outdoor smoking allowed
Dynamic/Deontic[2]	Ability	agent A expresses that action α is possible for him	We *can* go for no change in tobacco taxes
Dynamic/Deontic	Inability	agent A expresses that action α is not possible for him	It's *impossible* for me to accept no smoking in public transportation
Volitional	Acquiesce	agent A expresses that action α is possible, but not favourable for him	*$In-breath* okay, it is still possible

Table 2: Defined modality categories with definitions and examples. The modality triggers are marked *italic*.

requires that semantic annotations should have a well-defined semantics (Bunt and Romary, 2002). We based our modality annotation scheme on the following criteria (Bunt, 2014; Petukhova, 2014): (1) compatibility: it incorporates categories of existing schemes, see (Sections 2 and 3; (2) theoretical and empirical validity: each category is semantically defined and observed in the corpus, Section 4; (3) completeness: the scheme provides good coverage of the phenomena in question, Section 4.2; (4) distinctiveness: each category is clearly distinct; and (5) effective usability: both humans and machines can understand and distinguish the categories, Section 5. It has been demonstrated in the past that the fulfillment of these criteria supports well-founded decisions when designing the conceptual content and structure of an annotation scheme (Petukhova, 2011).

4   Modality expressions in negotiations

Modality is an important tool in negotiations that structures the interaction and enables participants to interpret each others' intentions and to evaluate their dynamically changing goals and strategies efficiently. In negotiation, participants introduce their options. When establishing jointly possible values, a bargaining range, the participants' actions are focused on obtaining and providing information about preferences and abilities. Parties also tend to mention the least desirable events. Apart from preferences and dislikes, a negotiator has certain goals to achieve, those are signalled by teleological modal expressions. Thus, the use of prioritizing modality is frequent.

Cooperative negotiators adjust their offers taking the partner's priorities into consideration, non-cooperative ones prefer to stick to their initial offers. Cooperative behaviour may be characterized by acknowledging other parties' preferences and making concessions where possible. Verbally, it can be communicated by utterances expressing acquiescence (see Table 2). Non-cooperative (adversarial, competitive) behaviour, by contrast, may be articulated by expressing inability and dislike.

4.1   Data collection

The data used in our modality classification experiments is referred to as the Metalogue Multi-Issue Bargaining (MIB) corpus (Petukhova et al., 2016). The data was collected in simulated negotiations in which two participants (City Councilor and Business Representative) negotiate the city's 'smoking-ban' policy based on a list of negotiation preferences that have been randomly assigned to them. Each participant in the experiment received a background story and instructions, as well as a preference profile. Their task was to negotiate an agreement which assigns exactly one value to each issue, exchanging and eliciting offers concerning $\langle ISSUE; VALUE \rangle$ options. The participants were randomly assigned their roles. They were advised to start with the highest possible values according to their preference

---

[1]In negotiation domain, action α mostly corresponds to *offers* expressed in the semantic content of an utterance.

[2]In our scenario, it is not always possible to distinguish between expressions of agent's personal physical abilities (dynamic) and possibilities imposed by norm and conventions including those related to the agent's membership in a certain professional group or political party (deontic).

SCOPE	**TAXATION**
○ All outdoor smoking allowed	○ No change in tobacco taxes
○ No smoking in public transportation	○ 5% increase in tobacco taxes
○ No smoking in public transportation and parks	○ 10% increase in tobacco taxes
○ No smoking in public transportation, parks and open air events	○ 15% increase in tobacco taxes
	○ 25% increase in tobacco taxes
CAMPAIGN	**ENFORCEMENT**
○ Flyer and billboard campaign in shopping district	○ Police fines for minors in possession of tobacco products
○ Anti-smoking posters at all tobacco sales points	○ Ban on tobacco vending machines
○ Anti-smoking television advertisements	○ Police fines for selling tobacco products to minors
○ Anti-smoking advertisements across all traditional mass media	○ Identification required for all tobacco purchases
	○ Government issued tobacco card for tobacco purchases

Figure 2: Example of preferences in four issues presented to participants as a colour.

information. The participants were not allowed to show their preference cards to each other. No further rules on the negotiation process, order of discussing issues, or time constraints were imposed. They were allowed to withdraw or re-negotiate previously made agreements within one session, or terminate a negotiation.

The anti-smoking regulations were concerned with four main *issues*: (1) smoke-free public areas (smoking ban scope); (2) tobacco taxes (taxation); (3) effective anti-smoking campaign (campaign); and (4) enforcement policy and police involvement (enforcement), see Figure 2. Each of these issues involves four to five most important negotiation *values* with preferences assigned representing parties negotiation positions, i.e. preference profiles. Nine cases with different preference profile were designed. The preferences strength was communicated to the negotiators through colours. Brighter orange colours indicated increasingly negative options; brighter blue colours increasingly positive options.

The collected corpus consists of 24 dialogues with 8 participants involved, of a total duration of about 2.5 hours, comprising approximately 2.000 speaking turns (about 10.000 tokens). To study modality we extracted 1145 task-related utterances.

4.2 Data annotations

The recorded data was transcribed, segmented and annotated with dialogue act information following the ISO standard. The ISO 24617-2 taxonomy ISO (2012) distinguishes 9 dimensions, addressing information about a certain *Task*; the processing of utterances by the speaker (*Auto-feedback*) or by the addressee (*Allo-feedback*); the management of difficulties in the speaker's contributions (*Own-Communication Management*) or that of the addressee (*Partner Communication Management*); the speaker's need for time to continue the dialogue (*Time Management*); the allocation of the speaker role (*Turn Management*); the structuring of the dialogue (*Dialogue Structuring*); and the management of social obligations (*Social Obligations Management*). Additionally, to capture the negotiation task structure, *Task Management* acts are introduced. These dialogue acts explicitly address the negotiation process and procedure. This includes utterances for coordinating the negotiators' activities (e.g., "Let's go issue by issue") or asking about the status of the process (e.g., "Are we done with the agenda?"). Task Management acts are specific for a particular task and are often similar in form but different in meaning from Discourse Structuring acts, which address the management and monitoring of the interaction. Examples of the later are utterances like "To sum up", and "Let's move to a next round".

With respect to modality, three main types of utterances are observed in the corpus: non-modalised utterances (41%), utterances containing triggers of prioritizing modality (preference - 30%, necessity - 2%, dislike - 3.1%, and acquiescence - 3%), and dynamic (ability - 19% and inability- 1.2%), see Table 2 for examples. The developed modality scheme has these seven modality categories. They were

Feature type	Training set			
	Original data		+Simulated Data	
	Baseline (MNB accuracy, in %)	SVM accuracy, in %	Baseline (MNB, accuracy, in %)	SVM accuracy, in %
unigrams	67.6	74.3	65.8	79.5
TFIDF unigrams	69.5	79.0	66.9	77.0
bigrams	69.5	80.0	67.9	82.6
TFIDF bigrams	69.5	75.2	71.4	79.7
1-skip-bigrams	73.3	73.3	68.0	79.8

Table 3: Classification results in terms of accuracy (in %) obtained on collected 'original' human-human and when adding artificially 'simulated' utterances.

chosen after studying the existing modality annotation schemes (see Section 3), the domain of multi-issue bargaining (see Section 4), and the corpus.

The modality types were annotated by two independent annotators using audio recordings and manually produced transcriptions. The annotators were instructed to look only at utterances expressing task-related acts and assign to them one of the seven modality categories. A list of verbal and paralinguistic triggers was provided for each category. A near perfect inter-annotator agreement on average was reached in terms of Cohen's kappa of 0.91 (Cohen, 1960).

5 Assessing automatic modality classification and learnability

To classify modal meanings in multi-issue bargaining dialogues, SVM-based learning experiments were conducted in stratified ten-fold cross-validation setting. Support Vector Machines (SVM) are known to generalize well when applied to small training samples sets and show a rather robust multi-class performance when using (Gaussian) radial basis function kernel (RBF kernel), see Chang et al. (2010).The obtained performance has been compared to the previously undertaken efforts reported in the literature (Section 3) and to the baseline system built on the training data using different features. As the baseline, Multinomial Naive Bayes (MNB) classifier was trained to predict modality classes. The MNB algorithm assumes conditional independence between features which makes it suitable to be used as a strong baseline. Moreover, MNB is robust, fast, and easy to implement.

Features were computed from speech transcriptions, such as token unigrams and bigrams, 1-skip bigrams of lemmas, tf-idf weights for unigrams and bigrams, and various combinations of those. We measured the trained classifier's accuracy assessing its performance on different types of features and a set of tuned default exhaustive grid search parameters. Additionally, error analysis was performed by studying the confusion matrices.

Our feature selection experiments showed that the classifier's accuracy ranges between 72 and 82.6%, and does not differ significantly when features are varied (see Table 3). It should be noticed, however, that the use of bigrams resulted in the highest accuracy scores.

The collected human-human data set was rather small to train robust predictive models in data-oriented way. To extend a training set with more data, *user simulation* is often used and is efficient when domain and task structure are well-defined, and the user model truly reflects what real users are likely to do (Paek, 2006). Applied to our domain and tasks, we based the generation of simulated modalised/non-modalised utterances on trigger patterns extracted from transcribed human-human dialogues and changing possible targets. For example, the originally recorded utterance was I prefer *a smoking ban* resulted in the generated one I prefer *a discount*, substituting the value *a smoking ban* by *a discount*. The larger training set comprises 6145 utterances. All trained models were tested on original human-human data.

When training on the data extended with simulated utterances, the results in terms of accuracy are consistent with those obtained on the original data: the use of unigrams and bigrams of tokens results in the highest accuracy scores, and SVM clearly outperforms the baseline Multinomial Naive Bayes classifier.
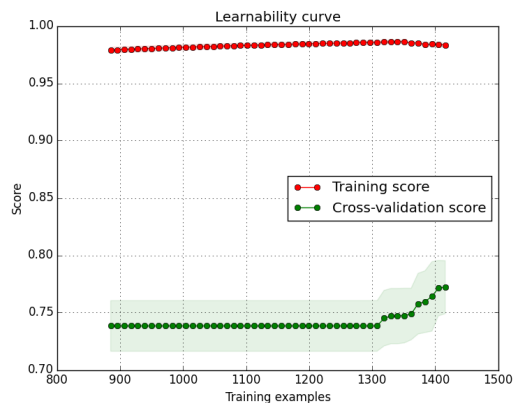
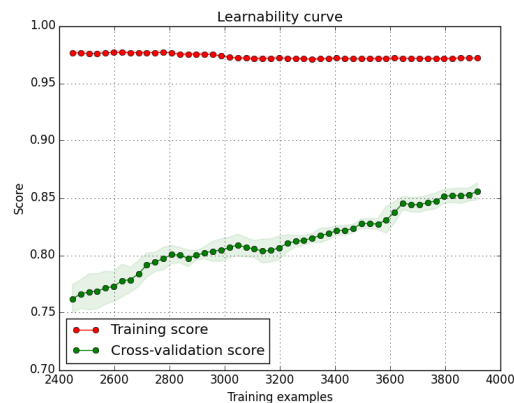Figure 3: Learnability curved using original human-human data.

Figure 4: Learnability curved using all available data.

Error analysis was performed by a detailed study of the confusion matrices, as well as assessing classifier precision for each class. The analysis of the confusion matrices showed that generally there are not many classification errors, and some errors were expected. For instance, discrimination between non-modalised utterances and those expressing preference presented the biggest problem to the classifier. This may be attributed to the absence of a verbal modality trigger in many 'preference' utterances. Elliptical forms are quite common as speakers rely heavily on the context and paralinguistics, e.g. intonation, and other prosodic features. In other cases an utterance has a verbal modality trigger, however, its meaning is ambiguous and can be disambiguated when certain contextual parameters are known. For example, *I think public transport and parks*, where *I think* indicates preference expressing boulomaic modality rather than being a trigger of epistemic modality.

Collecting and annotating data is costly. Knowing the minimal amount of data needed for training a robust predictive model will save efforts. Therefore, we conducted a series of learnability experiments and plotted learnability curve, see Figures 3 and 4. The experiments show that when training on all available data (original and simulated) the curve does not level off. This suggests that the classifier performance will further benefit from adding more training data.

6 Modal negotiation semantics

Negotiations are commonly analysed in terms of certain actions, such as offers, counter-offers, and concessions, see Watkins (2003), Hindriks et al. (2007). We considered two possible ways of using such actions, also referred to as 'negotiation moves', to compute the semantic meaning of partners' contributions in negotiation dialogues. One is to treat negotiation moves as task-specific dialogue acts. Due to its domain-independent character, the ISO 24617-2 standard does not define any communicative functions that are specific for a particular kind of task or domain, but the standard invites the addition of such functions, and includes guidelines for how to do so. For example, a negotiation-specific kind of $Offer_N$ function could be introduced for the expression of commitments concerning a negotiation value.[3] Another possibility is to use negotiation moves as the semantic content of general-purpose dialogue acts. For example, a negotiator's statements concerning his preference to a certain option can be represented as $Inform(A, B, \Box offer(X; Y))$.

We specified 8 basic negotiation moves: offer, counterOffer, exchange, concession, bargainIn, bargainDown, deal and withdraw, see Petukhova et al. (2017).

Negotiators often communicate their cooperativity by using modal utterances that express preference and ability. Non-cooperative behaviour, by contrast, may be articulated by expressing inability and dislike. Modality expressions are mainly observed in *Inform* and *Answer* acts.

According to ISO 24617-2, the representation of a dialogue act annotation with the ISO Dialogue

---

[3]Negotiation 'Offers' may have a more domain-specific name, e.g. *Bid* for selling-buying bargaining.

Act Markup Language (DiAML) makes use of the XML element `<dialogueAct>`. This element has the following attributes:

- `@target`, whose value is a functional segment identified at the second level;
- `@sender,@addressee,@otherParticipant`;
- `@communicativeFunction, @dimension`;
- `@certainty, @conditionality`, and `@sentiment` qualifiers;
- `@functionalDependence` and `@feedbackDependence`, which have `<dialogueAct>` elements and `<functionalSegments>` as values.

Additionally, rhetorical relations among dialogue acts are represented by means of `<rhetoLink>` elements.

`<NegotiationSemantics>` element has been added to DiAML to represent the semantic content of a dialogue act. A shallow negotiation semantics is defined in terms of `<NegotiationMove>` with attributes defined for different types of such moves. For example:

```
<dialogueAct xml:id="dap1TSK38" sender="#p1" addressee="#p2"
   dimension="task" communicativeFunction="inform"
   target="#fsp1TSKCV38">
     <NegotiationSemantics>
       <NegotiationMove type="counterOffer"/>
     </NegotiationSemantics>
   <rhetoricalLink rhetoAntecedent="#dap2TSK37"
   rhetoRel="substitution"/>
 </dialogueAct>
```

Additionally, dependent on annotation goals, approach, granularity and type of semantic processing, `<NegotiationSemantics>` elements are extended with `<Arg>` elements for negotiated issues and values, and `<Operators>` for logical operators between arguments. Modal relations can be represented by `<modalLink>` linking the holder (e.g. speaker) and target (semantic content) with values describing the speaker's attitudes to the necessity or probability of the events, and the speaker's abilities. Consider Figure 5 for the ISO 24617-2 meta-model extended with a modality relation between one or more participants (i.e. holder or actor) and one or more targets that mainly consist of an event and its arguments forming a semantic content (a negotiation move).

The full proposed DiAML representation of utterance *P1: I prefer all outdoor smoking allowed* produced by the sender P1 addressed to P2 is a task-related Inform act with the semantic content $\Box offer(1A)$ is as follows:

```
<dialogueAct xml:id="da1" sender="#p1" addressee="#p2"
dimension="task" communicativeFunction="inform"
target="#fsp1TSKCV38 qualifier="certain">
    <NegotiationSemantics>
    <NegotiationMove xml:id="nm1" type="offer"/>
       <Arg>issue-1; option-A</Arg>
       <modalLink holder="#p1" target="#nm1" modalRel="preference"/>
    </NegotiationSemantics>
</dialogueAct>
```

7 Conclusion and future work

There have been numerous attempts to automatically classify modal meanings in written, mostly scientific discourse. The objective of our study required designing a modality classification approach that can be applied to the modelling of spoken (multi-modal) negotiation dialogues. The present negotiation
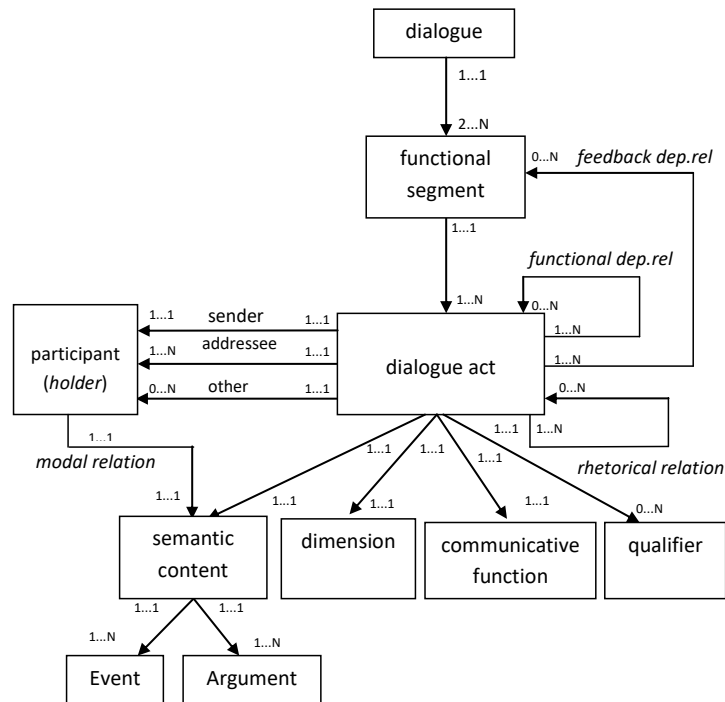
Figure 5: ISO 24617-2 metamodel for dialogue act annotation extended with modality elements.

modality annotation scheme includes categories related to expressions of speaker's *necessity, preferences, acquiescence* and *abilities*. The categories have been proven to be observable in data, semantically distinguishable and self-explanatory, which facilitated efficient human annotation and machine classification process.

The trained SVM classifier has generally showed good results in terms of accuracy. When it was trained only on the original data, accuracy ranged between 72% and 78%, outperforming a baseline multinomial Naive Bayes classifier in most cases. After extending the training set with artificially simulated utterances, we achieved accuracy range between 77% and 82.6%.

In future, we will test our predictive models within other negotiation domains and settings in order to assess their generalisability. We also plan to include multi-modal triggers as features to the classifier input space.

Acknowledgments

References

Baker, K., M. Bloodgood, B. J. Dorr, N. W. Filardo, L. Levin, and C. Piatko (2010). A modality lexicon and its use in automatic tagging. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC10)*, pp. 1402–1407. European Language Resources Association.

Bunt, H. (2014). Annotations that effectively contribute to semantic interpretation. In H. Bunt, J. Bos, and S. Pulman (Eds.), *Computing Meaning*, Volume 4, pp. 49–69. Springer.

Bunt, H. and L. Romary (2002). Towards multimodal content representation. In K. S. Choi (Ed.), *Proceedings of LREC 2002, Workshop on International Standards of Terminology and Language Resources Management*, pp. 54–60. ELRA.

Bybee, J. and S. Fleischman (1995). *Modality in Grammar and Discourse*. Typological studies in language. J. Benjamins.

Bybee, J. L., R. D. Perkins, and W. Pagliuca (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*, Volume 196. University of Chicago Press Chicago.

Calbert, J. (1975). Toward the semantics of modality. In J. Calbert and H. Vater (Eds.), *Aspekte der Modalitt*, pp. 2–70. Gunter Narr.

Chang, Y.-W., C.-J. Hsieh, K.-W. Chang, M. Ringgaard, and C.-J. Lin (2010). Training and testing low-degree polynomial data mappings via linear SVM. *Journal of Machine Learning Research 11*(Apr), 1471–1490.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Education and Psychological Measurement 20*, 37–46.

Fisher, R., W. L. Ury, and B. Patton (2011). *Getting to yes: Negotiating agreement without giving in.* Penguin.

Ghia, E., L. Kloppenburg, M. Nissim, P. Pietrandrea, and V. Cervoni (2016). A construction-centered approach to the annotation of modality. In *Proceedings of the 12th ISO Workshop on Interoperable Semantic Annotation*, pp. 67–74.

Hindriks, K., C. M. Jonker, and D. Tykhonov (2007). Analysis of negotiation dynamics. In M. Klusch, K. V. Hindriks, M. P. Papazoglou, and L. Sterling (Eds.), *Cooperative Information Agents XI: 11th International Workshop, CIA 2007, Delft, The Netherlands, September 19-21, 2007. Proceedings*, pp. 27–35. Berlin, Heidelberg: Springer Berlin Heidelberg.

ISO (2012). *Language resource management – Semantic annotation framework – Part 2: Dialogue acts. ISO 24617-2*. Geneva: ISO Central Secretariat.

Kilicoglu, H. and S. Bergler (2008). Recognizing speculative language in biomedical research articles: a linguistically motivated perspective. *BMC bioinformatics 9*(11), 1.

Kobayakawa, T. S., T. Kumano, H. Tanaka, N. Okazaki, J.-D. Kim, and J. Tsujii (2009). Opinion classification with tree kernel SVM using linguistic modality analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pp. 1791–1794. ACM.

Kratzer, A. (1981). The notional category of modality. In H. Eikmeyer and H. R. (eds.) (Eds.), *Words, worlds, and contexts: New approaches in word semantics*, pp. 38–74. Berlin: de Gruyter.

Lavid, J., M. Carretero, and R. J. Zamorano-Mansilla (2016). A linguistically-motivated annotation model of modality in English and Spanish: insights from MULTINOT. *LiLT (Linguistic Issues in Language Technology), Special Issue on Modality in Natural Language Understanding 14*, 1–33.

Leech, G. (1983). *Principles of Pragmatics*. Longman linguistics library ; title no. 30. Longman.

Lyons, J. (1970). *New horizons in linguistics*. Number v. 1 in A Pelican original. Penguin.

Medlock, B. and T. Briscoe (2007, June). Weakly supervised learning for hedge classification in scientific literature. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, Prague, Czech Republic, pp. 992–999. Association for Computational Linguistics.

Nirenburg, S. and V. Raskin (2004). *Ontological semantics*. MIT Press.

Nuyts, J. (2001). *Epistemic Modality, Language, and Conceptualization: A cognitive-pragmatic perspective*. Human Cognitive Processing. John Benjamins Publishing Company.

Paek, T. (2006). Reinforcement learning for spoken dialogue systems: Comparing strengths and weaknesses for practical deployment. In *Proceedings of the Interspeech-06 Workshop on Dialogue on Dialogues - Multidisciplinary Evaluation of Advanced Speech-based Interactive Systems*, Pittsburgh, Pennsylvania. International Speech Communication Association (ISCA).

Palmer, F. (1979). *Modality and the English modals*. Longman linguistics library. Longman.

Peters, P., P. Collins, and A. Smith (2009). *Comparative Studies in Australian and New Zealand English: Grammar and beyond*. Varieties of English Around the World. John Benjamins Publishing Company.

Petukhova, V. (2011). *Multidimensional dialogue modelling*. doctoral thesis, Tilburg University.

Petukhova, V. (2014). Understanding questions and finding answers: semantic relation annotation to compute the expected answer type. In *In Proceedings of the Tenth Joint ISO - ACL SIGSEM Workshop on Interoperable Semantic Annotation (ISA-10)*.

Petukhova, V., H. Bunt, and A. Malchanau (2017). Computing negotiation update semantics in multi-issue bargaining dialogues. In *Proceedings of the 21stWorkshop on the Semantics and Pragmatics of Dialogue (SemDial 2017 - SaarDial)*, Saarbrücken, Germany, pp. 114–124.

Petukhova, V., C. Stevens, H. de Weerd, N. Taatgen, F. Cnossen, and A. Malchanau (2016). Modelling multi-issue bargaining dialogues: Data collection, annotation design and corpus. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, and S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, pp. 31–33. European Language Resources Association (ELRA).

Portner, P. (2009). *Modality*. Oxford Surveys in Semantics & Pragmatics. OUP Oxford.

Raiffa, H., J. Richardson, and D. Metcalfe (2002). *Negotiation analysis: The science and art of collaborative decision making*. Harvard University Press.

Rubinstein, A., H. Harner, E. Krawczyk, D. Simonson, G. Katz, and P. Portner (2013). Toward fine-grained annotation of modality in text. In *Proceedings of IWCS 2013 Workshop on Annotation of Modal Meanings in Natural Language (WAMM), Potsdam, Germany*, pp. 38–46.

Watkins, M. (2003). Analysing complex negotiations. *Harvard Business Review 9-903-088*, 1–22.