Using Chatbot Technologies to Support Argumentation

Luis Henrique Herbets de Sousa[1], Guilherme Trajano[1], Analúcia Schiaffino Morales[1], Stefan Sarkadi[2] and Alison R. Panisson[1]

[1]*Department of Computing, Federal University of Santa Catarina, Santa Catarina, Brazil*

[2]*Department of Informatics, King's College London, London, U.K.*

Keywords: Argumentation, Chatbots, Human-Agent Interaction.

Abstract: Chatbots are extensively used in modern times and are exhibiting increasingly intelligent behaviors. However, being relatively new technologies, there are significant demands for further advancement. Numerous possibilities for research exist to refine these technologies, including integration with other technologies, especially in the field of artificial intelligence (AI), which has received much attention and development. This study aims to explore the ability of chatbot technologies to classify arguments according to the reasoning patterns used to create them. As argumentation is a significant aspect of human intelligence, categorizing arguments according to various argumentation schemes (reasoning patterns) is a crucial step towards developing sophisticated human-computer interaction interfaces. This will enable agents (chatbots) to engage in more sophisticated interactions, such as argumentation processes.

1 INTRODUCTION

Argumentation is one of the most significant components of human intelligence (Dung, 1995). Arguing allows different individuals to exchange relevant information during dialogues, enabling rich communication and a high level of understanding. When it comes to developing Artificial Intelligence (AI), human beings and the phenomena that emerge from their intelligence serve as inspiration. In this context, there are initiatives to develop argumentation-based techniques in intelligent agents, that is, artificial intelligence capable of communicating (and reasoning) using arguments (Maudet et al., 2006; Rahwan and Simari, 2009; Panisson and Bordini, 2017; Panisson et al., 2021b).

Recently, argumentation techniques have also been used as a method for intelligent agents to provide explanations (in the context of explainable AI) to humans about their decision-making (or decision-making suggestions) (Panisson et al., 2021a). These directions aim to develop technologies that enable applications in the context of hybrid intelligence (Akata et al., 2020), where humans and intelligent agents work together, sharing their intelligence capabilities and constantly communicating to make joint decisions. A challenge to achieve the ambitious objectives of those researches is the natural language communication interfaces. Among these challenges is the ability of an agent to understand an argument communicated by a human. One of the steps towards this understanding is to develop the ability of an agent to classify a received argument from a human according to commonly used patterns of reasoning used in that application domain, called argumentation schemes (Walton et al., 2008).

Argumentation schemes are patterns of reasoning used in specific or general domains (Walton et al., 2008), and they are gaining increasing attention from those interested in exploring the vast and rich interdisciplinary area between argumentation and AI (Girle et al., 2003). One of the reasons for this interest is the possibility of identifying common types of arguments in everyday discourse and conversations (Walton et al., 2008), using, for example, argument mining, which also allows for the automatic identification and extraction of components and structures of an argument (Lawrence and Reed, 2020). Moreover, argumentation schemes contain great potential for solving AI problems. For instance, the works of (Carbogim et al., 2000; Reed, 1997; Walton, 2000) demonstrate that argumentation provides a powerful means of dealing with non-monotonic reasoning problems, moving away from purely deductive and monotonic approaches to reasoning and towards presumptive and defeasible techniques.

Herbets de Sousa, L., Trajano, G., Morales, A., Sarkadi, S. and Panisson, A.
Using Chatbot Technologies to Support Argumentation.
DOI: 10.5220/0012578300003636

In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 2*, pages 635-645
ISBN: 978-989-758-680-4; ISSN: 2184-433X

In this paper, we investigate human-computer interaction technologies with the aim of classifying arguments provided by humans in natural language according to argumentation schemes (patterns of arguments) used in that domain of application. Thus, from this study and the developed technologies, arguments can be extracted, understood, and translated from human speech/text to a computational agent, enabling advances in the development of applications in the context of hybrid intelligence (Akata et al., 2020). To achieve this constant and rich communication between humans and machines (which contextualises hybrid intelligence), it is essential for a machine (AI) to be able to understand arguments communicated by humans. A technology with great potential to meet this need is chatbot technology, which is widely used for its ability to communicate very similarly to human communication. Therefore, considering this proximity to human communication, the use of chatbot technologies as an agent interface and human has great compatibility. A chatbot is software capable of holding a conversation with a human user in natural language, through messaging applications, websites, and other digital platforms and communication interfaces.

One of the essential elements in a chatbot is NLU (Natural Language Understanding), where a chatbot is trained to understand user inputs in natural language. This understanding is the result of a process of classifying user inputs according to a set of user intentions that can be identified in the corresponding application domain, also making the extraction (and identification) of entities that are relevant to understanding that user input. Thus, in this paper, we will investigate whether chatbot technologies also provide support for the classification of arguments communicated by humans in natural language into argumentation schemes (patterns of arguments), allowing an agent to understand the meaning of those arguments. In particular, we use 8 argumentation schemes, chosen according to their differences and complexities, and they are used to train the chatbot. Then, we evaluate whether the Rasa framework could classify different forms of arguments according to argumentation schemes (which are used to instance those arguments).

2 BACKGROUND

2.1 Argumentation Schemes

In the article presented by (Panisson et al., 2021a), the authors propose an approach to develop explainable agents using argumentation-based techniques. However, to achieve the development of applications in the context of explainability, as described in Section 1, there is a need to investigate human-computer interaction interfaces, such as making it possible for computational agents to understand arguments communicated in natural language by humans. Following the work proposed by the authors in (Panisson et al., 2021a), a very interesting direction of investigation is the use of argumentation schemes.

Argumentation schemes are patterns of reasoning used to instantiate (create) arguments (Walton et al., 2008). Argumentation schemes also represent forms of arguments that are revocable, which allows the implementation of highly sophisticated reasoning mechanisms on uncertain, incomplete, and conflicting information (Maudet et al., 2006). These characteristics are quite valuable when it comes to classifying human arguments into argumentation schemes (where arguments will have various forms, omission of information, etc.).

In artificial intelligence, especially in multi-agent systems, argumentation schemes have been used to provide means for intelligent agents to perform reasoning on conflicting or uncertain information, obtaining conclusions with solid grounding, and consequently, well-supported decisions (Panisson et al., 2021b). This argument-based reasoning process consists of a sophisticated process of monological argumentation reasoning. On the other hand, in the implementation of argumentation-based communication between intelligent agents, arguments are used for agents to justify/support their positions in different types of dialogues (Panisson et al., 2021b). In argumentation-based dialogues, it is customary to highlight the idea that agents can explain their positions in a deliberate dialogue, or even use arguments to persuade other agents in a negotiation, among others. For example, below we have the argumentation scheme *role to know*:
An agent 'ag' has the role 'R', and role 'R' knows things related to domain 'S' containing proposition 'A' ***(major premise)****. 'ag' asserts that 'A' is true* ***(minor premise)****. Then 'A' can be considered as true* ***(conclusion)****.*

An example in natural language, which characterises an instance of the argumentation scheme presented above, is exemplified below:

Fernando *is a* ***doctor*** *and knows things related to the domain about* ***cancer (major premise)****.* ***Fernando*** *asserts that* ***smoking causes cancer (minor premise)****. Then we conclude that* ***smoking causes cancer (conclusion)****.*

Note that the variables *ag*, *R*, *S*, and *A* are instantiated by elements of the application domain – $\{ag \mapsto Fernando,\ R \mapsto \text{doctor},\ S \mapsto \text{cancer},\ A \mapsto$

smoking causes cancer} – instantiating that pattern of reasoning. Also note that arguments have a slight variation of the considered standard and catalogued pattern of reasoning, which becomes a challenge for the classification process of the arguments, as well as the recognition of specific entities that were used to instantiate the argumentation scheme. These challenges are explored in this work.

2.2 Chatbot Technologies

Chatbots are artificial intelligence software that are capable of communicating with humans through human-computer interaction models. They utilise natural language processing for communication, which can be either in written or spoken form. Currently, chatbots are being utilised in numerous applications in various ways, primarily serving as a user-friendly communication interfaces. These technologies are considered friendly as they simulate natural conversation during dialogues with humans. A chatbot can respond based on pre-programmed guidelines or make use of AI to learn and adapt its responses.

Currently, chatbots have gained popularity as tools for replacing human agents in tasks such as customer service. They have also gained more attention as a human-computer interface with more general aspects, with applications of integration with distributed artificial intelligence systems, such as (Engelmann et al., 2021c; Engelmann et al., 2021b). Given this integration and the previous context, it is essential to study whether human-computer interaction technologies such as chatbots are capable of recognising, and at what level of detail, these more complex structures of human communication (recently applied to intelligent agents) such as argumentation. Therefore, in this paper, chatbot technologies will be used as a form of communication between agents and humans, investigating whether chatbot technologies provide support for identifying and detailing arguments communicated by humans to an intelligent agent, so that the agent can later understand and process these arguments uttered by humans in natural language.

There are various chatbot technologies available that could be used, such as IBM Watson[1], DialogFlow[2], and Rasa[3]. In this work, the Rasa framework will be used, justified by its various qualities such as having a solid documentation, allowing the integration of the chatbot on websites and applications, being highly customisable, and being open source. Also, it was decided to use Rasa, given the development of integration of this chatbot technology with multi-agent system development platforms (Custódio, 2022). The next section will be dedicated to describing the framework used and the necessary steps for implementing a chatbot.

[1]https://www.ibm.com/br-pt/products/watson-assistant
[2]https://dialogflow.cloud.google.com/
[3]https://rasa.com/

3 RASA

The Rasa framework is an open-source machine learning software for developing chatbots capable of automated text and voice conversations. Rasa can be used to automate human-computer interactions in numerous ways, from websites to social media platforms.

The Rasa framework has three main functionalities, namely Natural Language Processing (NLP), dialogue management, and integrations with other systems. For this work, we mainly used the NLP functionality, investigating whether it provides means for argument classification[4].

There are two components that interest us regarding the NLP functionality for the development of this work, which correspond to training the NLP unit and to the chatbot pipeline that allows us to improve its performance by introducing, for example, language models. Below we describe each of these topics.

3.1 NLU Training

The Natural Language Understanding (NLU) module of the Rasa framework provides natural language processing capabilities that transform user messages into intents and entities, allowing chatbots to understand those interactions.

For example, in order for a chatbot developed with Rasa to understand a greeting spoken in natural language, it is necessary to train the chatbot with a set of examples of natural language sentences that have that meaning. An example of such an implementation is shown below:

```
- intent: greeting
    example: |
    - Hello
    - Hi
    - Good afternoon!
    - Good morning
    - Good evening
    - Hey
```

In this way, any greeting, whether it be one of those presented above or variations of them, will

[4]The other functionalities may be explored in the future, by extending this work.

be recognized by the chatbot as the user's intention `greeting`.

Another important aspect of chatbot technologies related to training the natural language processing unit is the extraction of entities from user interactions. For example, in a broad context of chatbots, it is highly advisable that the chatbot have interactions that establish a connection with the user, for instance, by addressing the user by name. In order to achieve this, the chatbot must be able to extract this entity - the user's name - from its interactions with the user. For instance, after the chatbot asks for the user's name, the following intent could be used to process the user's response and extract the entity "`name`":

```
- intent: provide_name
    example: |
    - My name is [John](name)
    - I am [Mary](name)
    - [Juca](name) is my name
    - They call me [Mateus](name)
    - I am [Thiago](name)
```

The structure of the training examples for entity marking follows the example template above, where the portion of the sentence that represents an entity is marked with square brackets, and the name of the entity is annotated between parentheses.

3.2 Pipeline

Within Rasa, received messages are processed by a sequence of components, which are executed one after another in a process called pipeline. The possibility of choosing a pipeline for the NLU allows for customizing the chatbot model and better adapting it to the data that will be used in the application domain.

When no pipeline is defined, the Rasa framework automatically defines a pipeline based on the language defined in the framework's configuration file. However, it is possible to add several elements to the pipeline, depending on the project's needs. One of these elements is the *WhitespaceTokenizer*, which processes words separated by white space. Another example of an element is the *LexicalSyntacticFeaturizer*, which creates features for entity extraction from a message. The *LexicalSyntacticFeaturizer* element can be configured to better extract entities from messages expected by a created model. There are several other components that can be added, such as pretrained language models from spaCy[5], among other components.

[5]https://spacy.io/

4 USING CHATBOT TECHNOLOGIES TO RECOGNISE ARGUMENT IN NATURAL LANGUAGE

Recognizing arguments in natural language is a crucial step in developing artificial intelligence applications that will interact with humans, in a context where agents (intelligent software) and humans will work together to solve problems, which has been contextualized as hybrid intelligence (Akata et al., 2020).

There are already approaches that develop this type of interaction between intelligent agents and humans, where agents use arguments (properly translated into natural language) to explain their conclusions (resulting from their reasoning processes) and decision-making to human users (Panisson et al., 2021a; Ferreira et al., 2022), in applications such as hospital bed allocation (Engelmann et al., 2021c), task allocation in groups of collaborators (Schmidt et al., 2016), data access control (Panisson et al., 2018), among others. These works are developed using the multi-agent systems development platform Jason (Bordini et al., 2007), supported by the framework developed on this platform that supports the use of argumentation-based reasoning and communication techniques (Panisson et al., 2021b; Panisson and Bordini, 2020).

However, these works contextualize only one side of communication, where an intelligent agent can generate and communicate arguments in natural language to a human user. It is necessary to investigate how an intelligent agent would actually be able to engage in an argumentation process with human users, allowing users to counter-argue a decision explained by it, argue about their own conclusions, etc. In this context, there are already works that implement interfaces with well-known natural language processing technologies, such as chatbot technologies. Among these interfaces are Dial4Jaca (Engelmann et al., 2021b; Engelmann et al., 2021a) and Rasa4Jaca (Custódio, 2022), which implement interfaces between chatbot technologies such as Dialogflow and Rasa, and the Jason agent development platform (Bordini et al., 2007).

Thus, as an initial point of investigation for the presented context, in the development of natural language communication interfaces between intelligent agents and human users, we investigated how chatbot technologies could support the understanding of arguments presented by humans in natural language, classifying them into reasoning patterns used for argument instantiation, i.e., argumentation schemes.

This section presents the study that aimed to evaluate whether chatbot technologies provide the necessary means for an intelligent agent (or chatbot referred to here) to recognise arguments in natural language, providing an understanding of these complex structures to an intelligent agent. For the study purpose, 8 argumentation schemes extracted from the book (Walton et al., 2008) were chosen and modelled, and they will be presented below. For each scheme, 16 examples of arguments were defined for training the NLU, resulting in a dataset of 128 examples of natural language arguments. The 8 modelled argumentation schemes are presented below:

1 - Argumentation Scheme Role to Know: This argumentation scheme is adapted from the argumentation scheme *position to know* by (Walton et al., 2008), as presented in the work (Panisson et al., 2021a) and has the structure presented in Table 1.

Table 1: Argumentation scheme role to know.

Major Premise	Agent Ag is in a position to know about things in a particular subject domain S containing the proposition Ar.
Minor Premise	Ag asserts that Ar is true(false).
Conclusion	Ar is true(false).

An example argument that corresponds to an instance of this argumentation scheme is presented below:

[Joseph is a engineer](premise1) and [says bricks are better than blocks](premise2), so it can be concluded that [bricks are better than blocks](conclusion).

where `premise1` would be the major premise properly annotated in the example, `premise2` would be the minor premise, and `conclusion` the conclusion.

2 - Argumentation Scheme Classification: This argumentation scheme was directly extracted from the book (Walton et al., 2008), and has the structure presented in Table 2, and an example is presented below:

Table 2: Argumentation scheme classification.

Major Premise	All *F's* can be classified as *G's*.
Minor Premise	A is an F.
Conclusion	Therefore, A is a G.

[All developers from Company X are senior developers](premise1). [Todd work at Company X](premise2), so [Todd is a Senior Developer](conclusion).

3 - Argumentation Scheme Sign: Like the previous argumentation scheme, this one, and all the following ones, are schemes from the book (Walton et al., 2008). Its structure is presented in Table 3, and below we present an argument instantiated from it:

Table 3: Argumentation scheme sign.

Minor Premise	The data represented as statement A is true in this situation.
Major Premise	Statement B is generally indicated as true when its sign A is true.
Conclusion	Therefore, B is true.

[Here are some tracks](premise2) that [look like they were made by a bear](premise1). Therefore, [a bear possibly passed this way](conclusion).

4 - Argumentation Scheme Effect to Cause: its structure is presented in Table 4, and an example is presented below:

Table 4: Argumentation scheme effect to cause.

Premise 1	Generally, if A occurs, then B will occur.
Premise 2	In this case, B did in fact occur.
Conclusion	Therefore, A presumably occurred.

[Fred has high temperature](premise2). So, [Fred has a fever](conclusion).

Note that this scheme only includes premise 2 and the conclusion, which is also an adaptation of the original scheme. Premise 1 still serves as a way to understand and explain the argument's classification and how it works, and premise 2 is the one actually present in the argument.

5 - Argumentation Scheme Threat: This scheme has its structure defined with 3 premises, shown in Table 5, along with its example:

Table 5: Argumentation scheme threat.

Premise 1	If you bring A, bad consequences B will happen.
Premise 2	I am in a position to cause B.
Premise 3	I hereby assert that B will occur if you provoke A.
Conclusion	Therefore, it is better for you not to provoke A.

[If Jason bring a dog to this park](premise1) [he should pay a fee](premise2). [Jason can not afford this fee](premise3). [Therefore, is better for Jason that he does not bring a dog to the park](conclusion).

6 - Argumentation Scheme Guilt by Association: Its structure is presented in Table 6, and an example is presented below:

Table 6: Argumentation scheme guilt by association.

Premise	*Ag* is a member of or associated with the group *G*, which is morally condemned.
Conclusion	Therefore, *Ag* is a morally bad person.

[Jake is a member of Mur Family, and all members of Mur family are killers](premise1). [Therefore, Jake is a killer](conclusion).

This scheme originally has 2 conclusions, but in this work an adaptation was used, in which only conclusion 1 was used, as the other conclusion only serves as an extension of the used conclusion.

7 - Argumentation Scheme Positive/Negative Scheme for Practical Argument from Analogy: This argumentation scheme is a combination of two schemes, the positive and negative, but since the difference is only in the polarity of the sentence, both positive and negative examples were used. Its structure is presented in Tables 7 and 8, and below we present examples of both polarities:

Table 7: Argumentation scheme positive scheme for practical argument from analog.

Major Premise	The right thing to do in *S1* was to do *A*.
Minor Premise	*S2* is similar to *S1*.
Conclusion	Therefore, the right thing to do in *S2* is to do *A*.

[The righteous thing to do on the miners case was to help](premise1). [The wreckage building case is similar to miners case](premise2). Therefore, [the right thing to do in wreckage building case is to help](conclusion).

Table 8: Argumentation scheme negative scheme for practical argument from analogy.

Major Premise	The wrong thing to do in *S1* was to do *A*.
Minor Premise	*S2* is similar to *S1*.
Conclusion	Therefore, the wrong thing to do in *S2* is to do *A*.

[The wrong thing to do on first game was not communicating with the team](premise1). [The second game will be similar to the first game](premise2). Therefore, [the wrong thing to do in the second game is to not communicate](conclusion).

8 - Argumentation Scheme Necessary Condition: This scheme is divided into the premise of the objective and the necessary premise to achieve the objective, with the structure presented in Table 9, followed by an instantiation example:

Table 9: Argumentation scheme necessary condition.

Objective Premise	Making *Sn* is my objective.
Necessary Premise	To make *Sn*, it is necessary to do *Si*.
Conclusion	Therefore, I need to do *Si*.

[Jake wants to produce Wine](premise1). [In order to produce wine, planting grapes is necessary](premise2). Therefore, [Jake needs to plant grapes](conclusion)

To evaluate whether the chatbot technology used would be capable of correctly classifying arguments in the argumentation schemes presented above, two chatbot projects were developed, both using the 8 argumentation schemes, with 16 sentence examples for each of them.

In the first project, there was no marking of premises and conclusion, only the sentences were expressed in natural language, with the aim of verifying the classification in relation to the intention that corresponds to the argumentation scheme. Below we show an example of the training of an intention for argument classification:

```
- intent: classification
  example: |
  - All people who lives in Switzerland
    are rich. Nomu lives in Switzerland.
    So Nomu is rich
  - All animals that produce milk can be
    classified as mammals.
    A buffalo produces milk.
    Therefore a buffalo is a mammal
  - ...
```

In the second project, simple structures of the arguments were marked, in which it is interesting not only to recognise the intention that classifies the argument according to the modelled argumentation schemes, but also to identify which parts of the sentences are premises and which part of the sentence is the conclusion of the argument. Below is an example of training an intention for argument classification with simple structure:

```
- intent: classification
  example: |
  - [All people who lives in Switzerland
```

```
    are rich](premise1).
    [Nomu lives in Switzerland](premise2).
    So [Nomu is rich] (conclusion)
  - [All animals that produce milk can be
    classified as mammals] (premise1).
    A [buffalo produces milk](premise2).
    Therefore a [buffalo is a mammal]
    (conclusion)
  - ...
```

In addition, in order to better understand the classification process of the technology in both projects and for all argument schemes, the order of the elements in the argument structure was interchanged, as commonly found in different writing or speech styles. These non-standardized structures allowed for a better evaluation of the technology for the desired task. For example, in the *Role to Know* argument scheme, in the two examples presented below, we can observe the different forms of an argument, focusing on the order in which the premises and conclusions are presented in each example provided for the chatbot's training:

```
- intent: role_to_know
  example: |
  - [Jaime is a engineer](premise1)
    and [says bricks are better than
    blocks](premise2), so [it is
    concluded that bricks are better
    than blocks](conclusion)
  - [Today will rain](conclusion),
    [because Todd told me it will rain
    today](premise2), [Todd is a
    weatherman](premise1), wheathermans
    know about weather forecast.
  - ...
```

Note that in the first example above, `premise1`, `premise2`, and `conclusion` were marked, respectively, according to the argument structure. In the second example, the argument structure has the order of `conclusion`, `premise2`, and `premise1`.

After defining the intentions corresponding to the argumentation schemes, providing examples for the training of the NLU in both projects, some empirical experiments were carried out to evaluate how accurate the technology would be for argument classification with and without simple structure. The results obtained are presented in the next section.

5 RESULTS AND DISCUSSIONS

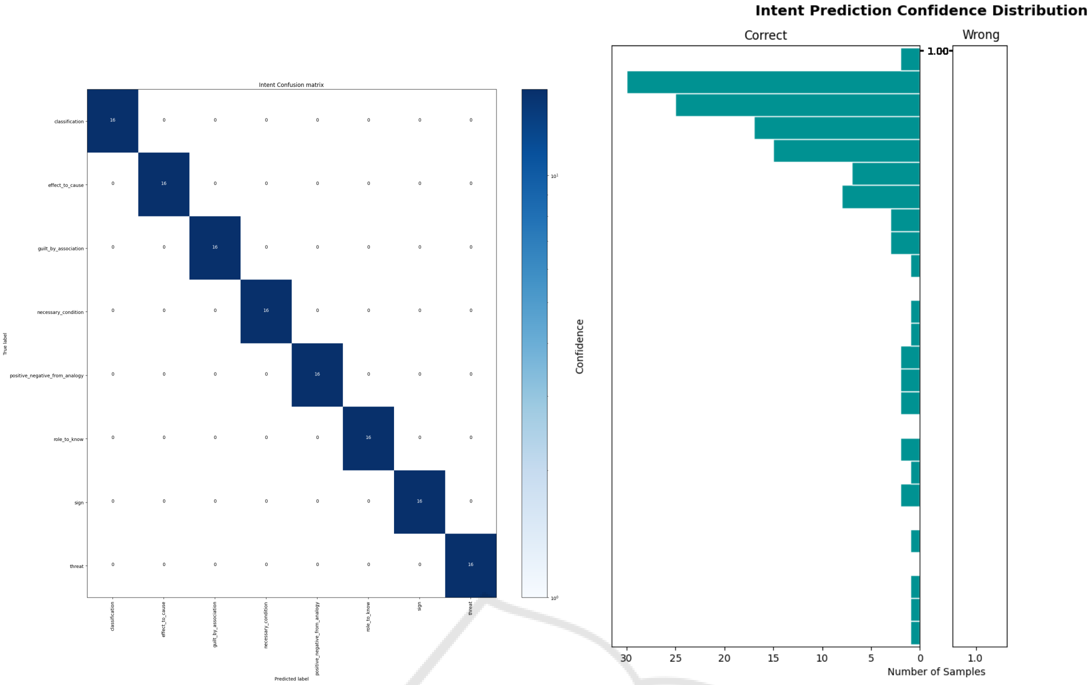
The evaluation was performed using the tools provided with the Rasa framework. The process followed the standard procedure for training a chatbot, where the NLU training was carried out primarily using data that corresponds to 128 examples classified into 8 argumentation schemes presented in the previous section, with 16 examples for each argumentation scheme.

After NLU training, tests were performed to certify the quality of the developed NLU. In this work, these tests provide indications of the ability of these technologies to classify arguments into argumentation schemes. To perform this analysis, the standard testing configuration established by Rasa was used, which automatically separates 80% of the sentence dataset examples for NLU training and 20% for validation and NLU verification. This separation is important because the tests are carried out with data that was not provided to the machine learning technique during training, providing results that resemble those that would be observed when the system is put into production, considering that the inputs will also often be data that were not present in the training dataset.

From the execution of the tests, the Rasa testing tool generates a confusion matrix and a histogram of the confidence distribution of the intention prediction, which correspond to the classification of arguments into argumentation schemes in this work. For the first project, where there is no marking of argument structures, we have the following results, presented in Figure 1 (a) and Figure 1 (b), respectively. As can be seen in Figure 1 (a), which presents the result for tests of argumentation scheme classification without marking premises and conclusion, all tests were correctly classified. It is important to understand that in a confusion matrix, the expected classes (in this case, argumentation schemes) for the example provided to the model during tests (rows of the matrix) are correlated with the resulting classification of the model (columns of the matrix). Thus, the correct classifications will be positioned on the main diagonal of the matrix, as observed in the matrix of Figure 1 (a). In case of wrong classifications, the confusion matrix allows to observe which classes received wrong classification and what was the wrong classification provided by the model.

In Figure 1 (b), the histogram of the confidence distribution of the intention prediction is presented, which, for this study, corresponds to the confidence in the classification of an argument in relation to the argumentation scheme. Higher bar represent higher confidence, starting from 1 at top to 0 at bottom. The horizontal bar indicate the number of samples classified with that particular confidence. Green bars indicate correct classifications, and red bars indicate wrong classification[6]. As can be observed, there are

[6]The main goal here is to avoid incorrect classifications, but when they do occur, they are typically associated with



(a) Confusion matrix.

(b) Histogram.

Figure 1: Intention (arguments) classification for the model with argumentation schemes without marked premises and conclusion.

no wrong classifications, and most of the classifications have a high prediction confidence.

For the second project, where premises and conclusions were marked in the examples for entity extraction, in addition to the confusion matrix and histogram to evaluate the classification of arguments into argumentation schemes, the testing tool also provides the confusion matrix of the entities and the histogram of the confidence distribution of the entity prediction, indicating after the correct classification of the argument into its respective argumentation scheme whether the entities (marked premises and conclusions) were also correctly extracted by the NLU.

The results obtained for classification are presented in the confusion matrices and histograms of Figure 2 (a) and Figure 2 (b), respectively. The results obtained for entity recognition are presented in the confusion matrix and histogram of Figures 3 (a) and 3 (b), respectively.

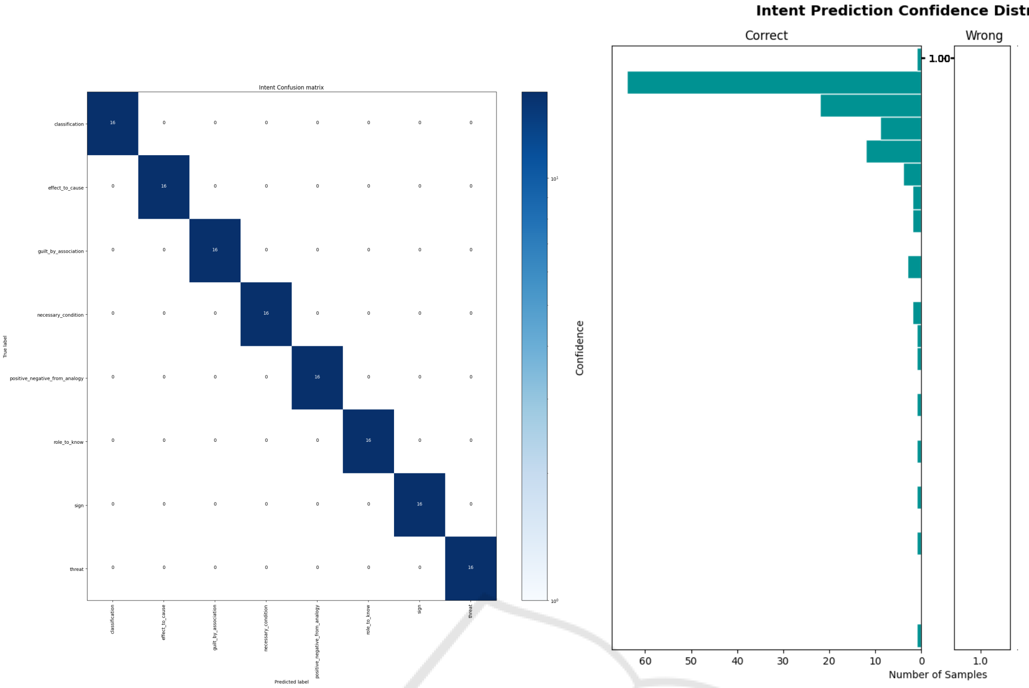
As can be seen from the confusion matrix in Figure 2, obtained through premise-marked classification tests, the results are similar to those obtained in tests without premise markings, i.e., the markings do not influence the classification of the argument in its respective argumentation scheme. Regarding the histogram of confidence distribution of intention prediction presented in Figure 2, it can be noted that premise-marked arguments increased the confidence of the model on classifying those arguments, where most arguments are correctly classified with higher confidence than the previous test (more samples close to confidence 1 (at top)).

Regarding the correct recognition of entities, in our case marked premises and conclusion, the results are presented in Figures 3 and 3. It can be observed that only 5 entities (premises and conclusions) were recognised incorrectly, which represents a very small proportion compared to the number of correctly recognised entities. The histogram in Figure 3 shows that the incorrectly recognised entities have a prediction confidence of less than 54%.

6 RELATED WORK

The author in (Wells, 2014) suggests that the literature in argumentation schemes and dialogue games can be classified as follows: (i) Games unable to utilise (i.e., represent and manipulate) argumentation schemes; (ii) Games able to utilise a single argumentation scheme; and (iii) Games able to utilise multiple/arbitrary argumentation schemes. Also, the author describes that there is no game at the third level, con-

low confidence.



(a) Confusion matrix. (b) Histogram.

Figure 2: Intention (arguments) classification for the model with argumentation schemes with premises and conclusion labelling.

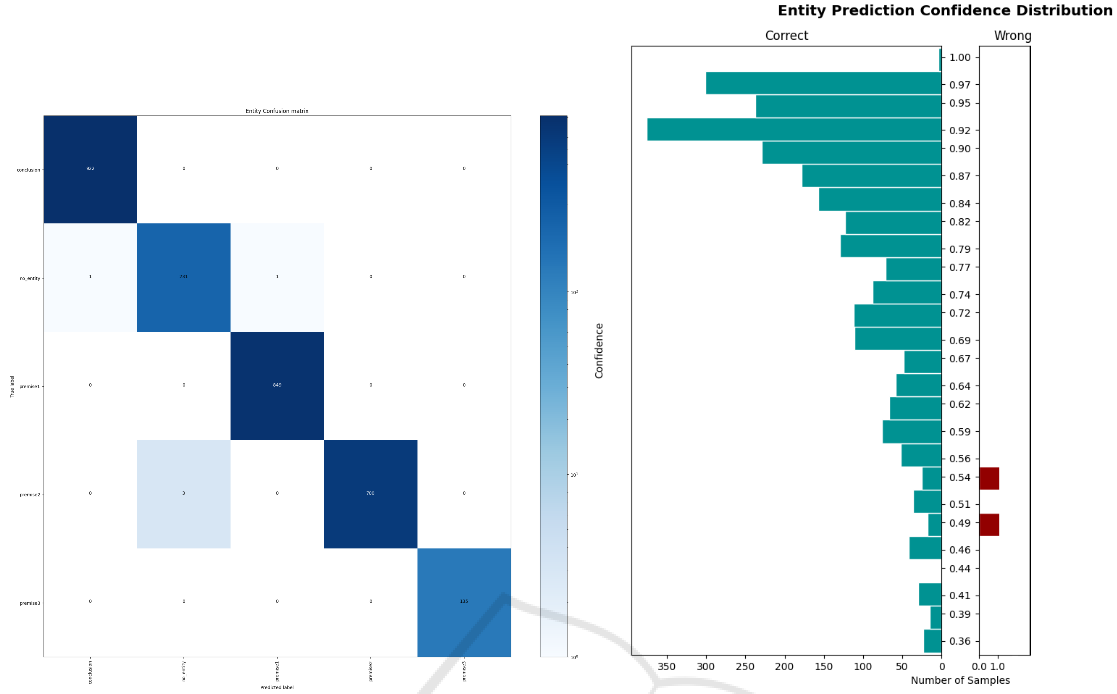
sidering multiple argumentation schemes. Our work moves towards filling this gap, providing argument (and argumentation scheme) understanding to agents.

In (Walton, 2012), the authors highlight the use of argumentation schemes for argument extraction, advocating a bottom-up approach wherein arguments are grouped based on their similarities. In (Katzav and Reed, 2004), the authors provide the conception of argument and arguments types, suggesting them as the basis for developing a natural classification of arguments. The work is based on the idea that stereotypical forms (or schemes) of arguments are found in natural discourse, but there is no formal and well-defined basis for it such that it could be employed in IA systems.

In (Macagno and Walton, 2015), the authors argue that classifying the structure of natural arguments has an important role in dialectical and rhetorical theories, which also reflects to AI systems based on those theories. Also, the authors describe that argumentation schemes are the main representation for arguments based on those theories. They also argue that argumentation scheme could be selected according to the intended purpose of the argument during dialogues.

Also, argumentation schemes have been used to argument mining (Lawrence and Reed, 2016), which aims to extract argumentative structures form text. The authors describe that argumentation schemes provide a rich information to extract argumentative structures from natural language texts. They also explain that by training various classifiers, it becomes feasible not only to classify the argumentation scheme being used but also the components of the arguments, including their respective roles.

While our approach is inspired by all the mentioned work in this section, it differs from all of them. Our approach focuses on classifying and extracting argument structure based on well-known argumentation schemes, using chatbot technologies to support this natural language understanding task. After this classification and component extraction, arguments could be grouped according to different views, such as (Walton, 2012; Katzav and Reed, 2004; Macagno and Walton, 2015). Additionally, our approach could be used for argument mining (Lawrence and Reed, 2016), considering that the developed chatbot works as a classifier and component extractor, capable of, for example, filtering only those arguments classified with higher accuracy from discourses/texts (interactively providing sentences from them to the chatbot). However, it is important to note that our work focuses on providing an interface for argumentation-based dialogues between intelligent software agents and humans, wherein our approach supports the understanding of arguments by intelligent agents.



(a) Confusion matrix. (b) Histogram.

Figure 3: Prediction of the extracted entities from argumentation schemes with premise and conclusion markings.

7 CONCLUSION

In this work, we investigated whether chatbot technologies are capable of classifying arguments presented in natural language into their respective reasoning patterns (argumentation schemes) used to instantiate arguments. Two lines of evaluation were explored for the proposed investigation. The first was where arguments only need to be classified according to the argumentation scheme used to instantiate them. The second was where elements of argument composition, i.e., premises and conclusion, were also identified, in addition to their classification in relation to the argumentation scheme.

The results obtained from our study were very promising, and we concluded that yes, chatbot technologies have great potential for implementing this type of problem. However, the limitation of the number of argumentation schemes used is a limitation of the results obtained, and future work will aim to extend the number of modeled patterns. Through the investigation carried out in this work, possibilities for great technological advances are opened up, such as using existing works, such as Dial4JaCa (Engelmann et al., 2021b) and the argumentation framework centred on the use of argumentation schemes (Panisson et al., 2021b), providing “scheme awareness” to agents (Wells, 2014), to develop agents capable of understanding an argument presented by a human user. This would make an intelligent agent capable of counter-arguing or even better understanding the user, considering that the arguments presented by them support their position and provide justifications for it, among other emerging possibilities of these sophisticated communication phenomena.

For future work, it is possible to further increase the classification capability of arguments with Rasa by training more argumentation schemes beyond the eight used. For example, training with all existing argumentation schemes would create a general-purpose argumentative AI. This line of research also includes the need for numerous examples, such as various interpretations of arguments. It is also possible to explore other Rasa pipeline configurations to improve the classification ability, according to the various types of text interpretation available for the pipeline. There are many possibilities for advancement from this work, as even using the basic pipeline shows positive results. In addition, with the use of the Rasa framework, it is possible to integrate the trained chatbot into applications and use these applications with human users so that they can converse with the chatbot and validate how Rasa will behave in a real scenarios, in the classification of arguments created by humans, in argumentation schemes. It is also possible to integrate the developed chatbot with intelligent agent development technologies, such as the Jason framework, through integration interfaces, as studied by (Custódio, 2022).

REFERENCES

Akata, Z., Balliet, D., De Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., et al. (2020). A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(08):18–28.

Bordini, R. H., Hübner, J. F., and Wooldridge, M. (2007). *Programming multi-agent systems in AgentSpeak using Jason*. John Wiley & Sons.

Carbogim, D. V., Robertson, D., and Lee, J. (2000). Argument-based applications to knowledge engineering. *The Knowledge Engineering Review*, 15(2):119–149.

Custódio, M. (2022). Rasa4jaca: Uma interface entre sistemas multiagentes e tecnologias chatbots open sourc.

Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357.

Engelmann, D., Damasio, J., Krausburg, T., Borges, O., Cezar, L. D., Panisson, A. R., and Bordini, R. H. (2021a). Dial4jaca–a demonstration. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 346–350. Springer.

Engelmann, D., Damasio, J., Krausburg, T., Borges, O., Colissi, M., Panisson, A. R., and Bordini, R. H. (2021b). Dial4jaca–a communication interface between multi-agent systems and chatbots. In *International conference on practical applications of agents and multi-agent systems*, pages 77–88. Springer.

Engelmann, D. C., Cezar, L. D., Panisson, A. R., and Bordini, R. H. (2021c). A conversational agent to support hospital bed allocation. In *Brazilian Conference on Intelligent Systems*, pages 3–17. Springer.

Ferreira, C. E. A., Panisson, A. R., Engelmann, D. C., Vieira, R., Mascardi, V., and Bordini, R. H. (2022). Explaining semantic reasoning using argumentation. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 153–165. Springer.

Girle, R., Hitchcock, D., McBurney, P., and Verheij, B. (2003). Decision support for practical reasoning. c. reed and t. norman (editors): Argumentation machines: New frontiers in argument and computation. argumentation library.

Katzav, J. and Reed, C. A. (2004). On argumentation schemes and the natural classification of arguments. *Argumentation*, 18:239–259.

Lawrence, J. and Reed, C. (2016). Argument mining using argumentation scheme structures. In *COMMA*, pages 379–390.

Lawrence, J. and Reed, C. (2020). Argument mining: A survey. *Computational Linguistics*, 45(4):765–818.

Macagno, F. and Walton, D. (2015). Classifying the patterns of natural arguments. *Philosophy & Rhetoric*, 48(1):26–53.

Maudet, N., Parsons, S., and Rahwan, I. (2006). Argumentation in multi-agent systems: Context and recent developments. In Maudet, N., Parsons, S., and Rahwan, I., editors, *ArgMAS*, volume 4766 of *Lecture Notes in Computer Science*, pages 1–16. Springer.

Panisson, A. R., Ali, A., McBurney, P., and Bordini, R. H. (2018). Argumentation schemes for data access control. In *COMMA*, pages 361–368.

Panisson, A. R. and Bordini, R. H. (2017). Argumentation schemes in multi-agent systems: A social perspective. In *International Workshop on Engineering Multi-Agent Systems*, pages 92–108. Springer.

Panisson, A. R. and Bordini, R. H. (2020). Towards a computational model of argumentation schemes in agent-oriented programming languages. In *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pages 9–16. IEEE.

Panisson, A. R., Engelmann, D. C., and Bordini, R. H. (2021a). Engineering explainable agents: an argumentation-based approach. In *International Workshop on Engineering Multi-Agent Systems*, pages 273–291. Springer.

Panisson, A. R., McBurney, P., and Bordini, R. H. (2021b). A computational model of argumentation schemes for multi-agent systems. *Argument & Computation*, (Preprint):1–39.

Rahwan, I. and Simari, G. R. (2009). *Argumentation in artificial intelligence*, volume 47. Springer.

Reed, C. (1997). Representing and applying knowledge for argumentation in a social context. *AI & SOCIETY*, 11(1):138–154.

Schmidt, D., Panisson, A. R., Freitas, A., Bordini, R. H., Meneguzzi, F., and Vieira, R. (2016). An ontology-based mobile application for task managing in collaborative groups. In *The Twenty-Ninth International Flairs Conference*.

Walton, D. (2000). The place of dialogue theory in logic, computer science and communication studies. *Synthese*, 123(3):327–346.

Walton, D. (2012). Using argumentation schemes for argument extraction: A bottom-up method. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 6(3):33–61.

Walton, D., Reed, C., and Macagno, F. (2008). *Argumentation schemes*. Cambridge University Press.

Wells, S. (2014). Supporting argumentation schemes in argumentative dialogue games. *Studies in Logic, Grammar and Rhetoric*, 36(1):171–191.