Analysing debates on climate change with textual entailment and ontologies

Roxana Szabo and Adrian Groza
Intelligent System Group,
Department of Computer Science,
Technical University of Cluj-Napoca, Romania
Roxana.Szabo@isg.cs.utcluj.ro, Adrian.Groza@cs.utcluj.ro,

***Abstract*—The difficult task of recognising textual entailment aims to check if a natural language text T entails a smaller statement H. Current methods rely on machine learning and various lexical resources. Our aim is to include domain knowledge when searching for entailment or non-entailment. As most available knowledge comes in form of ontologies, we focused on translating description logic axioms into lexical rules suitable for existing textual entailment algorithms. We apply the developed system in the climate change domain, where many pro and counter arguments do exist. The performed experiments indicate an increasing of performance when including domain knowledge into the existing textual entailment algorithms.**

***Index Terms*—textual entailment, ontologies, climate change, argumentation**

I. Introduction

Climate experts agree on human-caused global warming [11]. The general opinion is that climate change is caused by factors such as biotic processes, solar radiation, plate tectonics, and volcanic eruptions. Certain human activities have been identified as primary causes of ongoing climate change, often referred to as global warming [14].

Different from experts, the agreement on the existence of climate change has not reached all public arena. In this line, climate change is still a very popular subject of online debates, with many people arguing on climate change related issues. To support people to clarify this doubt, we propose a system which can analysis the arguments from a debate. We can use this tool to decide if an opinion, or a short description is related to a specified domain or subject or not. In our paper, we use a knowledge base generated from ontologies in domain of the climate change.

To make sense of the arguments in natural language, we rely here recognizing textual entailment (RTE). RTE aims to recognise the relationship between sentence pairs, specifically if they entail or contradict each other. RTE tasks may involve 2-way or 3-way inference judgements. In case of a 2-way judgement, the class to predict is either $Entailment$ or $Nonentailment$. On the 3-way judgement scheme, the Nonentailment class is further differentiated into $Contradiction$ and $Unknown$ [10]. Various algorithms proposed to recognize textual entailment such as Alignment-based Entailment Recognition.

978-1-5386-3368-7/17 $31.00 © 2017 IEEE

Our corpus of arguments contains pairs of $text - hypothesis$. The corpus was automatically extracted from online debate sites. Such debate platforms provide valuable data about public opinions since all have a common structure: someone adds a question or statement, (which is the $hypothesis$ in our case), and other people convey pro and cons arguments (which are the $text$ in our RTE task).

RTE uses different edit distance algorithms to map the hypothesis H into the text T [9]. Mappings are performed as sequences of editing operations (i.e. insertion, deletion, substitution of text portions) needed to transform T into H, where each edit operation are made based on various linguistic knowledge resources like lexical (i.e., WordNet, VerbOcean, Wikipedia) and other linguistic levels of information.

In this paper, we present an approach for extending the the available lexical resources with ontologies. Our solution exploits domain knowledge by automatically converted OWL ontologies into lexical resources. We validate the solution against various pro and con arguments from debates on climate change. Finally, we present the results of our experiments to measure the utility of domain knowledge in the RTE task.

The rest of the paper is organized as follows: Section II presents the system architecture, the NLP tools, algorithms and resources enacted. Section III details a running scenario on various debates in climate change domain. Section IV discusses related work, while section V concludes the paper.

II. System architecture

Our system relies on two technical instrumentations: 1) *textual entailment* for making sense of climate change debates in natural language, respectively 2) *climate change ontology* to support the analyse through domain knowledge. We use two inputs: a corpus of labelled arguments for training the textual entailment and the climate change ontology. These two inputs are detailed in the following two subsections. Finally, this section presents how we apply the textual entailment instrumentation to our task.

A. Training corpus for experiments

The first task is to obtain a corpus of arguments on climate change. We use the debate corpus related to climate change from [7]. The corpus contains 142 debate topics with 877 pro and against arguments extracted from three debate sites: *Debatepedia* (http://www.debatepedia.org), *debate.org* (http://www.debate.org), and *For and Against* (http://www.forandagainst.com). The advantage of using debate sites is that arguments are already annotated by their creators with “pro or ”against“ labels. Hence, we have a large corpus of annotated arguments that can be used for training the heavy textual entailment machinery.

In this line, the next step is to translate the available corpus into a suitable format for textual entailment. Each argument contains of pairs of texts and hypothesis. The pairs are annotated with two relations: *entailment* or *contradiction*. One entailment example follows:

```
<pair id="1" entailment="ENTAILMENT">
   <t> The arctic vegetation zones are affected
       by the climate changes. </t>
   <h> The polar deserts is affected by
       the global warming.</h>
</pair>
```

Similarly, the following example illustrates a non-entailment (or contradiction) relation:

```
<pair id="401" entailment="NONENTAILMENT">
   <t> CA rise in sea levels worldwide due
       to the global warming.</t>
   <h> Human activities cause global warming.
   </h>
</pair>
```

The corpus is used by the machine learning component of the entailment algorithms to learn entailment and nonentailment relations between sentences.

B. Converting OWL ontologies into lexical rules

The second task is to obtain the domain knowledge. This is in our case a climate change ontology. The ontology supports the process of identifying entailments in natural language arguments. Consider the following axioms in description logic[1]:

$$\begin{aligned}
&CarbonDioxid \sqsubseteq Gas \sqcap \\
&\quad \forall hasNaturalSource.(Volcanoes \sqcup HotSprings) \sqcap \\
&\quad \forall producedBy.AerobicOrganism \sqcap \\
&\quad \forall hasColour.\bot \sqcap \\
&\quad \forall hasSmell.\bot \sqcap \\
&\quad \exists hasFormula.\{CO2\} \sqcap \\
&\quad \forall hasSynonyms.(CarbonicAcidGas \sqcup \\
&\qquad\qquad CarbonicAnhydride \sqcup \\
&\qquad\qquad CarbonicOxide \sqcup \\
&\qquad\qquad DryIce)
\end{aligned}$$

The above axiom formalises $CarbonDioxid$ as a gas having natural sources as volcanoes or hot springs, that is produced by all aerobic organisms, colourless and odourless, and which can be found under various names such as ”Carbonic acid gas“ or ”Dry Ice“ in solid phase. This knowledge supports recognising textual entailment task in statements like the following pairs:

[1]We assume the reader is familiarised with basic terminologies of description logics. For a detailed explanation about families of description logics, the reader is referred to [1].

TABLE I
MAPPINGS OF RELATION TYPES:

No.	*Ontology relationships*	*EOP Relation Type*
1.	Subclass Axioms	STRONGER THAN
2.	Equivalent Classes	SIMILAR
3.	Disjoint Classes	OPPOSITE OF
4.	AnnotationProperty: exact synonym	SIMILAR
5.	AnnotationProperty: related synonym	SIMILAR

T: *CarbonDioxid is a main factor in climate change.*
h: *CO2 is the main cause in global warming.*

Here the relation between $CarbonDioxid$ and $CO2$ is provided from the domain ontology. By applying the lexical alignment $CarbonDioxid \rightsquigarrow CO2$, the distance between T and h decreases.

Such domain ontologies are available on the web in OWL format. Our task is to automatically convert OWL to something usable by the textual entailment machinery. Therefore, we need to generate new lexical resources from an OWL ontology. The newly created lexical resource contains triplets structured as follows: $\langle word_1 - relation - word_2 \rangle$. For this task, we extract noun-base triplets from the given ontology. to create triplets like: $\langle noun_1 relation noun_2 \rangle$.

Firstly, the algorithm identifies the concepts and then the relations between them based on the ontology’s structure. When the ontology classes and their relations are identified, we propose to map each relation to a specified relation type. This relation type will used in order create the triplets for the new lexical resource file. Table I lists the mappings for each relation into ontology relation:

The basic heuristic employed to extract concepts and instances is that nouns represent concepts and proper nouns are instances, as Table II illustrates.

C. Alignment-based entailment recognition

Having both the training corpus and the domain ontologies converted for textual entailment, we can enact Excitement Open Platform (EOP) tool [12]. The textual entailment is used to analyse the domain hypotheses on the available text-arguments. For experiments, we created a dataset in the climate change domain containing 800 pairs of text/hypothesis divided into 50”%” of entailment pairs and 50”%” of non-entailment pairs. All the data sets are distributed in RTE-3 style format that is fully compatible with the EOP.

Data set consists of a training data set for training the system and a test data set for evaluating it. Based on the model generated after training, EOP computes the confidence of hypotheses entailment within the text. This confidence constitutes one of the main factor in the results analysis. Also, an important factor is the accuracy of the system.

The textual entailment process is enriched with various lexical knowledge bases such as VerbOcean [4] or Wordnet, as Fig. 1 bears out. The system has two main functionalities.

Firstly, the system aims to help the user to decide if a given text is for or against experts theory about global warming.

TABLE II
GENERATING LEXICAL RULES FROM CLIMATE CHANGE DOMAIN ONTOLOGY.

1.OWL axiom: Inclusion axiom ($A \sqsubseteq B$)	Lexical relation: STRONGER THAN ($\twoheadrightarrow$)
Well $\sqsupseteq$ OilWell	well $\twoheadrightarrow$ oil well
Well $\sqsupseteq$ WaterWell	well $\twoheadrightarrow$ water well
SalineLake $\sqsupseteq$ AlkalineSaltLake	saline lake $\twoheadrightarrow$ alkaline salt lake
VegetatedArea $\sqsupseteq$ Tundra	VegetatedArea $\twoheadrightarrow$ Tundra
OrganicAromaticCompound $\sqsupseteq$ Arene	organic aromatic compound $\twoheadrightarrow$ arene
OrganicAromaticCompound $\sqsupseteq$ BenzenoidAromaticCompound	organic aromatic compound $\twoheadrightarrow$ benzenoid aromatic compound
OrganicAromaticCompound $\sqsupseteq$ Heteroarene	organic aromatic compound $\twoheadrightarrow$ heteroarene
FuelOil $\sqsupseteq$ Paraffin	fuel oil $\twoheadrightarrow$ paraffin
MarineCurrent $\twoheadrightarrow$ DeepOceanCurrent	marine current $\twoheadrightarrow$ deep ocean current
MarineCurrent $\twoheadrightarrow$ OceanCurrent	marine current $\twoheadrightarrow$ ocean current
2.OWL axiom: Equivalent classes ($A \equiv$)B	Lexical relation: SIMILAR ($\sim$)
River $\equiv$ DryRiver	river $\sim$ dry river
3.OWL axiom: Disjoint classes ($A \sqsubseteq \neg B$)	Lexical rule: OPPOSITE OF ($\nsim$)
Clay $\sqsubseteq \neg$ B	clay $\nsim$ silt
4.Exact synonym: $hasSynonym$	Lexical relation: SIMILAR ($\sim$)
SalineLake $\sqsubseteq \forall$ hasSynonym.SaltLake	saline lake $\sim$ salt lake
ObsoleteUnpasteurizedMilkProduct $\sqsubseteq \forall$ hasSynonym. RawMilk	obsolete unpasteurized milk product $\sim$ raw milk
CarbonDioxid $\sqsubseteq \forall$ hasSynonym.CarbonicAcidGas	carbon dioxid $\sim$ carbonic acid gas
5.Related synonym: $hasRelatedSynonym$	Lexical relation: SIMILAR ($\sim$)
SalineLake $\sqsubseteq \forall$ hasRelatedSynonym.Salina	saline lake $\sim$ salina
SalineLake $\sqsubseteq \forall$ hasRelatedSynonym.SodaLake	saline lake $\sim$ soda lake
Garden $\sqsubseteq \forall$ hasRelatedSynonym.Park	garden $\sim$ park
VolcanicCave $\sqsubseteq \forall$ hasRelatedSynonym.LavaCave	volcanic cave $\sim$ lava cave
CarbonDioxid $\sqsubseteq \forall$ hasFormula.{CO2}	carbon dioxid $\sim$ CO2

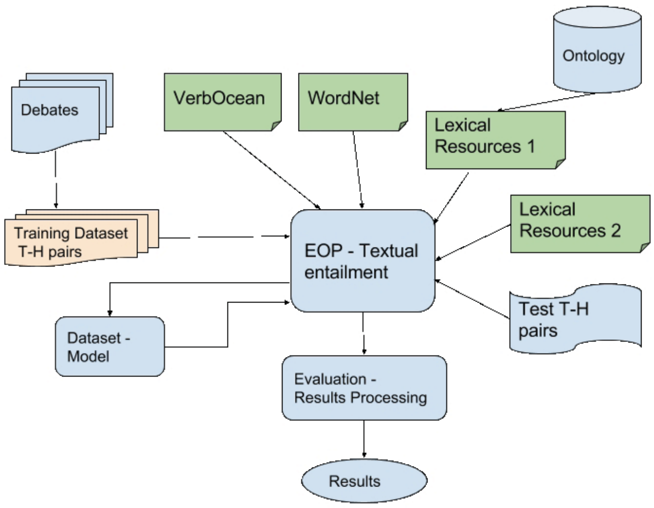

Fig. 1. System architecture.

Input : $\mathcal{O}$ - domain ontology;
$\mathcal{LR} \leftarrow$;set of lexical rules
foreach *verb* $v \in s$ **do**
find the longest sequence of words rv such that:
(1) rv starts at v,
(2) rv satisfies the syntactic constraint, and
(3) rv satisfies the lexical constraint
end
if $\exists$ *pair of matches adjacent or overlap in* s **then**
merge them into a single match;
end
foreach *relation phrase* r **do**
find the nearest noun phrase x to the left of $r \in s$ such that $x \notin RelativePronoun \cup \{There\}$,
Find the nearest noun phrase y to the right of $r \in s$.
end
return $\langle x, r, y \rangle$.

Algorithm 1: OWL2EOP: Converting OWL ontology in lexical resources.

The hypotheses are selected from a predefined list of experts' theories, which is considerated to be the topic of the debate. The text is represented by the opinion written by a human agent. Secondly, the system evaluates metrics based on the given text-hypothesis pair. In this case, the text, and also,the hypothesis are introduced by the user. The user checks the entailment relation type, the confidence and other metrics between the given text and hypothesis.

EOP provides functionality for the automatic identification of entailment relations among texts. The EOP's use is open both to users interested in using textual inference in applications and to developers willing to extend the current functionalities. This is the main reason, why we proposed to extend the current functionalities of the P1EDA sub-module by adding new lexical resources to improve the final decision if the T entails H. The platform combines the linguistic pipelines, entailment algorithms and linguistic resources such as knowledge resources.

The EOP architecture [13] consists of two main parts. The Linguistic Analysis Pipeline (LAP) is a series of linguistic annotation components such as tokenization to part of speech tagging or chunking. The Entailment Core (EC) performs the actual entailment recognition based on the Entailment Decision Algorithms (EDAs) and more subordinate components. EDA computes an entailment decision for a given Text/Hypothesis pair. It uses standardized algorithms: transformation based, edit-distance based, and classification based, and knowledge resources. Scoring components accept a Text/Hypothesis pair as an input and return a vector of scores.

P1EDA is a new alignment based EDA added recently on the EOP platform. One main goal of the EDA was making future contribution and adding of linguistic knowledge as easy as possible. EDA is designed to utilize diverse sources of lexical, syntactical and other linguistic levels of information. Annotation Components are used to add different annotations to the Text/Hypothesis pairs.

Knowledge is needed to recognize cases where T and H use different textual expressions (words, phrases) while preserving entailment:

climate change → *global warming*
Hawaii → *America*
is affected by → *suffer from*
cause → *implies*

EOP contains a wide range of knowledge resources, including lexical and syntactic resources. *Lexical knowledge components* describe semantic relationships between words. The knowledge is represented as directed rules made up of two word - POS pairs, where the left-hand side (LHS) entails the right-hand side (RSH):

synthetic compounds → *chemicals compounds*
shooting star → *meteorite*

Syntactic knowledge components capture entailment relationships between syntactic and lexical-syntactic expressions. We represent such relationships by entailment rules that link dependency tree fragments that can contain variables as nodes. For example, the rules:

X is affected by Y → *Y affects X* or
X affects Y → *Y suffers from X* or
X causes Y → *Y is due to X*

express general paraphrasing patterns at the predicate-argument level that cannot be captured by purely lexical rules. Formally, each syntactic rule consists of two dependency tree fragments plus a mapping from the variables of the LHS tree to the variables of the RHS tree [8].

Various linguistic knowledge are annotated as alignment. Alignment links connect components of Text and Hypothesis. Each link represents a relation and any relation can be represented as an alignment. From LAP level, CAS holds generic language processing results like POS, lemma, syntactic dependencies, NER, and so on. All individual components add alignments in a CAS that holds a T-H pair. Then, EDA uses all formalised knowledge in CAS to determine entailments.

Figure 2 exemplifies an alignment-based entailment at the algorithmic level. Links are identified between words or phrases across the two texts: words/phrases of T can explain words/phrases of H. The identified links are are highlighted by the same colors. Broad-coverage knowledge needed to align words/phrases consists of i) align identical words; ii) align equivalent/related phrases - using paraphrase resources; and iii) align lexically related words: using lexical resources (i.e, WordNet, VerbOcean).

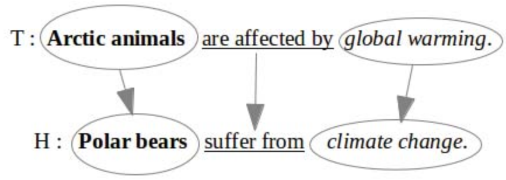

Fig. 2. Alignment-based Entailment.

Figure 3 presents the workflow for the P1EDA module. Labelled $T - H$ are processed by various aligners. Based on the identified links, the scores return a vector of features that is passed to a classifier and it generates the final entailment decision and a confidence score for it. The underlying classifier can be changed by any other Weka machine learning algorithm. The basic setup for P1EDA module uses all four aligners A_i: (A_1) IdenticalLemmaPhrase, (A_2) MeteorPhrase, (A_3) WordNet, respectively (A_4) VerbOcean.

The *IdenticalLemmaPhrase* aligner is a surface level aligner that aligns identical lemma sequences found in text and hypothesis. The module annotates only the longest lemma sequence. As an illustration, consider example 1.

Example 1 (Identical lemma sequences)**.**
T: *The melting of ice can produce a rise in sea levels.*
H: *A rise in sea level is due to the climate changes.*
Here, only one link is added that connects five tokens from T to five tokens from H:

$$['a\ rise\ in\ sea\ levels'] \xrightarrow{A_1} ['A\ rise\ in\ sea\ level]$$

The *MeteorPhrase linker* is a lexical level paraphrase aligner, lookup-aligner, that adds all links and it does not select the best one. It simply adds everything the underlying resource knows. The resource is based on the English Meteor Paraphrase table [5].

The *WordNet linker* is a lexical aligner based on WordNet. WordNet groups English words into sets of synonyms called synsets, provides short definitions and usage examples, and records a number of relations among these synonym sets or their members.

The *VerbOcean linker* is a lexical aligner that links tokens based on VerbOcean [4], which is a semantic network of verbs. VerbOcean extractor reads simple text files of VerbOcean file

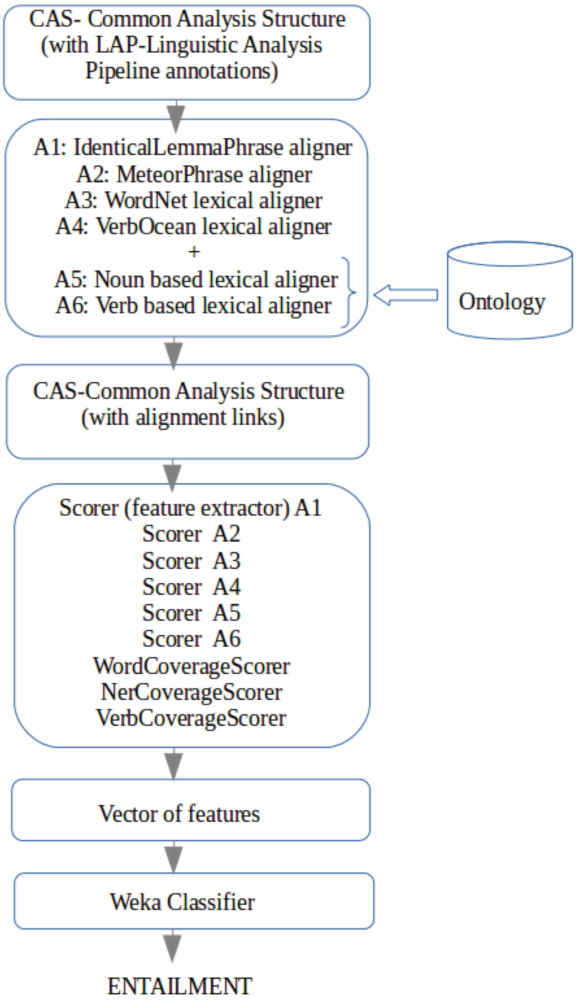

Fig. 3. Extending P1EDA workflow with ontology-based aligners.

format and converts the file content to a Topic Map. VerbOcean file format is a text file format where each line contains a quartet with two verbs and a relation, and a strength value.

These two lexical aligners are lexical aligners based on lexical resources. They compute alignment links based on lexical rules: if T contains a phrase t, H contains a phrase h and a lexical resource contains one of the rules $t \rightarrow h$ or $h \rightarrow t$, then an alignment link between t and h is created.

As an extension for the P1EDA, we created two new lexical aligner that uses generated lexical knowledge resources from ontologies. Our lexical resources consist of two types of lexical rules using the format: $\langle word_1, relation, word_2 \rangle$.

The lexical aligner based on new verb lexical resource uses the lexical resource for VERBS (A_5) which was generated manually using verbs from climate change domain. The lexical file contains triplet with two verbs and a relation between them.

TABLE III
ACCURACY FOR THE ENTAILMENT DECISION ALGORITHM.

EDA	*Accuracy (%)*
Transformation-based English RTE-3	67.13
Edit-Distance English RTE-3	64.38
Classification-based English RTE-3	65.25

Example 2 (Lexical resource based on verbs)**.**

cause	*[stronger-than]*	*imply*
cause	*[stronger-than]*	*produce*
affect	*[stronger-than]*	*suffer*

The second, lexical aligner based on new noun lexical resource uses the lexical resource for NOUNS (A_6) generated automatically from a climate change domain ontology. The resource file contains triplets with two nouns or nouns phrases.

The lexical file contains triplets with two verbs and a relation between them.

Example 3 (Lexical resource based on nouns)**.**

clay	*[opposite-of]*	*silt*
CO2	*[stronger-than]*	*gas*
carbon dioxid	*[similar]*	*CO2*
global warming	*[stronger-than]*	*climate change*

The features extractor generates a vector of features based on the identified alignment links, such as: *wordCoverageScorer*. *simpleProperNounCoverageScorer* or *verbCoverageScorer*.

The *word coverage scorer* is an alignment evaluator which returns four values: count covered tokens, count all tokens, count covered content-tokens, count all content-tokens. The *NER coverage scorer* is a POS-based coverage feature extractor that extracts how much of hypothesis named entities are covered. It returns two values: number of NER words covered in H and number of all NER words in H. The *verb coverage counter* returns also two values: number of covered verbs in H and number of all verbs in H). These attributes are used by the classifier to learn entailment and nonentailment relations.

To sum-up, our methodology for improving performance of entailment algorithm by means of domain ontology is as follows: Firstly, we load the domain knowledge, that is the ontology in climate change domain domain shown in Figure II. Secondly, we generate the list of lexical rules which are extracted from the ontology. The resulted lexical resources file contain nouns, and we use a second lexical resources for verbs, created manually. Then, we train our system using the prepared dataset, and using the additional lexical resources. Finally, we test the obtained model against a smaller set of data and then run different experiments using all six aligners, or a combinations of aligners.

III. Running Experiments

Consider the samples of T-H pairs in the climate change domain in Table IV. The tests are performed using 15 different

TABLE V
APPLYING ALIGNERS ON PAIR P_1.

Notation	*Aligner*	*Decision for P_1*	*Confidence*	*Accuracy (on train set)*	*Precision (on train set)*
A_1	IdenticalLemmaPhrase aligner	Entailment	0.632	0.475	0.473
A_2	MeteorPhrase aligner	NonEntailment	0.600	0.5	0.5
A_3	Lexical Aligner based on WordNet	Entailment	0.622	0.675	0.675
A_4	Lexical Aligner based on VerbOcean	NonEntailment	0.527	0.5	0.5
A_5	Lexical Aligner based on new verb lexical resource	NonEntailment	0.540	0.5	0.5
A_6	Lexical Aligner based on new noun lexical resource	NonEntailment	0.540	0.5	0.5
A_{12}	$A_1 + A_2$	NonEntailment	0.506	0.7	0.766
A_{13}	$A_1 + A_3$	Entailment	0.632	0.475	0.474
A_{123}	$A_1 + A_2 + A_3$	NonEntailment	0.509	0.7	0.766
A_{1356}	$A_1 + A_3 + A_5 + A_6$	NonEntailment	0.551	0.475	0.474
A_{1256}	$A_1 + A_2 + A_5 + A_6$	Entailment	0.645	0.7	0.766
A_{1234}	$A_1 + A_2 + A_3 + A_4$	NonEntailment	0.509	0.7	0.766
A_{12345}	$A_1 + A_2 + A_3 + A_4 + A_5$	Entailment	0.572	0.7	0.766
A_{12346}	$A_1 + A_2 + A_3 + A_4 + A_6$	Entailment	0.572	0.7	0.766
A_{123456}	$A_1 + A_2 + A_3 + A_4 + A_5 + A_6$	Entailment	0.651	0.7	0.766

TABLE IV
SAMPLE OF PAIRS USED FOR TESTING.

Pair	$\langle T, H\rangle$
P_1	T: *The CO2 causes global warming.* H: *The gas implies global warming.*
P_2	T: *Warmer global temperatures increase intensity of storms.* H: *Global warming causes an increase in storms.*
P_3	T: *Global warming is not caused by human activity.* H: *Climate change is manmade.*
P_4	T: *The melting of ice can produce a rise in sea levels.* H: *The humans activities cause the melting of ice.*

aligners in Table V. Here, aligners A_1, A_2, A_3, and A_4 are provided by the EOP framework. A_5 and A_6 represent our contribution and they are obtained by converting the climate change ontology. The remaining aligners represent various combinations of the base aligners A_1 to A_6.

If we use only the A_1 aligner based on identical lemma phrase or the A_3 aligner based on WordNet, the decision will be $Entailment$ with 0.62-0.63 confidence. Note that the accuracy is larger when using WordNet compared to the IdenticalLemmaPhrase aligner.

With the aligner A_2 (MeteorPhrase) or A_4 based on VerbOcean, the decision will be $NonEntailment$ with 0.5 accuracy similar precision.

By combining the aligners A_1, A_2 and A_3 we obtained in genera the $Nonentailment$ decision with a smaller confidence score, only 0.5 excepted the case when we use A_1 and A_3, when the result is $Entailment$ but the accuracy and the precision is small enough.

An interesting test case is when we combine the aligners A_1, A_2, A_5, and A_6. In this case, decision is $Entailment$ with 0,64 confidence, that is a good score. Also the accuracy and precision with a fair value of 0.7. Differently, when we aggregate A_1, A_3, A_5, and A_6, the result is $Nonentailment$ with 0.55 confidence and the accuracy and precision is just 0,47, that is quite small.

The above results indicate that the MeteorPhrase aligner combined with our 2 new aligners A_5 and A_6 is more successful than the WordNet aligner combined with A_5 and A_6

We investigated the combinations of our aligners A_5 and A_6 with the other aligners. If we use the A_5 and A_6 aligners, $Entailment$ is deduced with highest precision and accuracy, excepting one case

If we use only the default aligners A_1, A_2, A_3 and A_4, the result for the P_1 pair is $NonEntailment$ with 0.5 confidence. Note that, when adding one of the aligners A_5 or A_6 the decision is the correct one - that is $Entailment$ - with 0.57 confidence, Moreover, with both A_5 and A_6, $Entailment$ is inferred with the highest confidence 0.65 and with the highest accuracy and precision.

The above experiments confirm that the two new added aligners A_5 and A_6 improve the entailment algorithm, while the score is increased with 0.15 points, which means that the confidence is increased by 30%.

The Nonentailment is exemplified in pair 3 and the results for the 15 different aligners are presented in Table VI. The default aligners A_1, A_2, A_3, and A_4 provide a Nonentailment result with 0.649 confidence. By adding the two new aligners A_5 and A_6, the result is the same because even if the alignments are identified between the words or pharses, there are some words or phrases which are in contradiction.

Similar results were obtained for the remaining pairs P_2, and P_4.

TABLE VI

APPLYING ALIGNERS ON PAIR P_2.

Notation	*Aligner*	*Decision for* P_1	*Confidence*	*Accuracy (on train set)*	*Precision (on train set)*
A_1	IdenticalLemmaPhrase aligner	NonEntailment	0.598	0.65	0.73
A_2	MeteorPhrase aligner	NonEntailment	0.584	0.55	0.63
A_3	Lexical Aligner based on WordNet	NonEntailment	0.580	0.475	0.25
A_4	Lexical Aligner based on VerbOcean	NonEntailment	0.556	0.5	0.25
A_5	Lexical Aligner based on new verb lexical resource	NonEntailment	0.562	0.5	0.25
A_6	Lexical Aligner based on new noun lexical resource	NonEntailment	0.562	0.5	0.25
A_{12}	$A_1 + A_2$	NonEntailment	0.627	0.7	0.812
A_{13}	$A_1 + A_3$	NonEntailment	0.615	0.6	0.696
A_{123}	$A_1 + A_2 + A_3$	NonEntailment	0.648	0.7	0.766
A_{1356}	$A_1 + A_3 + A_5 + A_6$	NonEntailment	0.615	0.6	0.696
A_{1256}	$A_1 + A_2 + A_5 + A_6$	NonEntailment	0.627	0.7	0.812
A_{1234}	$A_1 + A_2 + A_3 + A_4$	NonEntailment	0.649	0.7	0.766
A_{12345}	$A_1 + A_2 + A_3 + A_4 + A_5$	NonEntailment	0.649	0.7	0.766
A_{12346}	$A_1 + A_2 + A_3 + A_4 + A_6$	NonEntailment	0.649	0.7	0.766
A_{123456}	$A_1 + A_2 + A_3 + A_4 + A_5 + A_6$	NonEntailment	0.649	0.7	0.766

IV. DISCUSSION

There are several studies which aim to process the arguments from social debates, in order to identify the position of each person regarding the selected premise in the context of information verification in user-generated content . Also, the goal is to detect if the user's text is in contradiction or not.

For application domains such as information extraction, question answering or summarisation, for which evidence from multiple sentences needs to be combined the entailment recognition methods are very useful [10]

In the PHEME project [10] the authors used the data corpus collected from the Twitter social media platform. Firstly, the data set has been normalized and then manually annotated to identify the three type of relations: contradiction, entailment and unknown. Differently, we use a corpus in which the arguments were already annotated by their creators.

Previous RTE research has mainly focused on achieving good performance on the Entailment relation, whereas this method is motivated by the need for a resource that facilitates the development of processing approaches specifically targeting the Contradiction relation.

Analysis texts from social media content has increased the attention of many authors. The task is a bit challenging since the informal nature of user-generated content makes the task difficult. Recently, however, the focus has also shifted to argumentation mining from social media texts, such as online debates [2], or product reviews [6].

On the other hand, nowadays there are many argumentation tools have been created to support the users in online social discussions. However, the main drawback of these tools is that they do not cope with the automatic generation of the arguments from the natural language discussions of the users [3]

[3] have proposed to use the textual entailment to automatically generate the abstract arguments from the dialogues. The abstract arguments as well as their relationships are then structured in an argumentation graph to evaluate the dialogue as a whole. The main goal of this approach is that offer the dynamics of the dialogues among users allowing to find the use of argumentation natural enough to be really adopted.

V. CONCLUSION

From the technical perspective, the main contribution consists of incorporating domain knowledge into the textual entailment machinery. As most of the available knowledge bases comes in form of OWL ontologies, we focused on translating description logic axioms into lexical rules suitable for entailment inference.

From the conceptual perspective, out contributions is twofold. First, the entailment machinery is heavily based on machine learning, by providing entailment and nonentailment pairs of free text for training. In this line, we managed to enrich the machine learning with domain knowledge. Second, entailment machinery extensively searches for various replacements within the T or H in order to minimise the semantic distance between them. Here, the domain ontology acts as a heuristic which guides searching by given preference to lexical rules obtained from the ontology.

The methodology proposed here can be easily adapted to different application domains. Simple, one has to provide a different OWL ontology in order to improve the difficult task of recognizing textual entailments.

However, the impact of domain knowledge on textual entailment accuracy depends on how well the ontology covers the domain encapsulated within the corpora used for training.

Acknowledgements

We thank the reviewers for their valuable comments. This work was supported by the by the Department of Computer Science from Technical University of Cluj-Napoca (TUCN), Romania through grant ”WSN in precision agriculture: hybrid models for recommendation and learning, 2017-2018.

References

[1] F. Baader, *The description logic handbook: theory, implementation, and applications*. Cambridge university press, 2003.

[2] E. Cabrio and S. Villata, “Combining textual entailment and argumentation theory for supporting online debates interactions,” in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2*. Association for Computational Linguistics, 2012, pp. 208–212.

[3] ——, “Generating abstract arguments: A natural language approach.” in *COMMA*, 2012, pp. 454–461.

[4] T. Chklovski and P. Pantel, “Verbocean: Mining the web for fine-grained semantic verb relations.” in *EMNLP*, vol. 4, 2004, pp. 33–40.

[5] M. Denkowski and A. Lavie, “Meteor-next and the meteor paraphrase tables: Improved evaluation support for five target languages,” in *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR*. Association for Computational Linguistics, 2010, pp. 339–342.

[6] D. Ghosh, S. Muresan, N. Wacholder, M. Aakhus, and M. Mitsui, “Analyzing argumentative discourse units in online interactions,” in *Proceedings of the First Workshop on Argumentation Mining*, 2014, pp. 39–48.

[7] A. Groza, P. Ozturk, R. R. Slavescu, R. Prasath, and A. Marginean, “Arguing on climate change in social media: an intelligent system for argument aggregation and debate analysis,” in *arXiv*, 2017.

[8] I. Gurevych, M. Mühlhäuser, C. Müller, J. Steimle, M. Weimer, and T. Zesch, “Darmstadt knowledge processing repository based on uima,” in *Proceedings of the First Workshop on Unstructured Information Management Architecture at Biannual Conference of the Society for Computational Linguistics and Language Technology, Tübingen, Germany*, 2007, p. 89.

[9] M. Kouylekov and B. Magnini, “Recognizing textual entailment with tree edit distance algorithms,” in *Proceedings of the First Challenge Workshop Recognising Textual Entailment*, 2005, pp. 17–20.

[10] P. Lendvai, I. Augenstein, K. Bontcheva, and T. Declerck, “Monolingual social media datasets for detecting contradiction and entailment,” in *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC), Portoroz, Slovenia. European Language Resources Association (ELRA)*, 2016.

[11] G. Liobikien and M. Butkus, “The european union possibilities to achieve targets of europe 2020 and paris agreement climate policy,” *Renewable Energy*, vol. 106, pp. 298 – 309, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960148117300447

[12] B. Magnini, R. Zanoli, I. Dagan, K. Eichler, G. Neumann, T.-G. Noh, S. Pado, A. Stern, and O. Levy, “The excitement open platform for textual inferences,” *ACL 2014*, p. 43, 2014.

[13] S. Padó, T.-G. Noh, A. Stern, R. Wang, and R. Zanoli, “Design and realization of a modular architecture for textual entailment,” *Natural Language Engineering*, vol. 21, no. 02, pp. 167–200, 2015.

[14] A. the Science of Climate Change (ASCC), “Americas climate choices: Panel on advancing the science of climate change,” 2010.