A Persuasive Chatbot using a Crowd-Sourced Argument Graph and Concerns

Lisa A. CHALAGUINE [a] and Anthony HUNTER [a]
[a] *Department of Computer Science, University College London, London, UK*

Abstract. Chatbots are versatile tools that have the potential of being used for computational persuasion where the chatbot acts as the persuader and the human agent as the persuadee. To allow the user to type his or her arguments, as opposed to selecting them from a menu, the chatbot needs a sufficiently large knowledge base of arguments and counterarguments. And in order to make the user change their current stance on a subject, the chatbot needs a method to select persuasive counterarguments. To address this, we present a chatbot that is equipped with an argument graph and the ability to identify the concerns of the user argument in order to select appropriate counterarguments. We evaluate the bot in a study with participants and show how using our method can make the chatbot more persuasive.

Keywords. chatbots, argumentative persuasion systems, computational persuasion, natural language argumentation, concerns, argument graphs

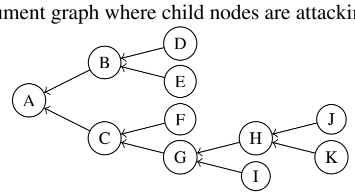
1. Introduction

Chatbots have the potential of being used as agents in argumentative persuasion systems that can engage in argumentative dialogues with users where the chatbot acts as the persuader and the user as the persuadee. Argument graphs as proposed by Dung [11] can be used as a knowledge base for the chatbot. Graphs are a useful representation to study attack and support relationships of a given set of arguments. Different kinds of semantics can be applied in order to identify the “winning” and “losing” arguments in a graph. This, however, assumes that all the possible arguments and their relationships are present in the graph.

Acquiring an argument graph raises several issues: most importantly *where* to obtain the relevant arguments for the argument graph, but also, which arguments to include in the knowledge base and how to justify the inclusion of some and exclusion of others (e.g. noise and repetition of arguments), and how to establish relations between arguments (the arcs of the graph). In our previous work [7] we presented a method and its evaluation for the acquisition of a large argument graph with over 1200 arguments via crowd-sourcing.

An argumentative chatbot could use such a graph in order to persuade a human agent to accept the bot’s stance by presenting arguments from the graph that support its stance and counter user arguments that do not. One way to utilise such a graph is by using a *menu-based* approach where the chatbot, after presenting an argument, gives the user a choice of counterarguments that the user can select from a menu [13]. Taking the argument graph shown in Figure 1 as an example, the chatbot would give argument A and then give the user arguments B and C to choose from. Suppose the user prefers

Figure 1. Argument graph where child nodes are attacking parent nodes.



argument C and selects that one. The chatbot selects a counterargument based on some criteria (or randomly) and replies with argument G and gives the user arguments H and I as countering choices, and so on. This way, the chatbot and the user would follow the arcs of the graph until (depending on the type of graph) all the arguments are used, or the user chooses an argument that has no counterarguments in the graph.

The drawback of the menu-based approach is, of course, that the user is limited to the choice of possible counterarguments presented by the chatbot, which might not include the user's preferred choice. This might limit the persuasive effect of the argumentative dialogue, as well as deny the chatbot the opportunity to acquire novel arguments on that topic which were not collected during the acquisition phase of the graph. The user arguments from the chats could then be used to extend the existing argument graph.

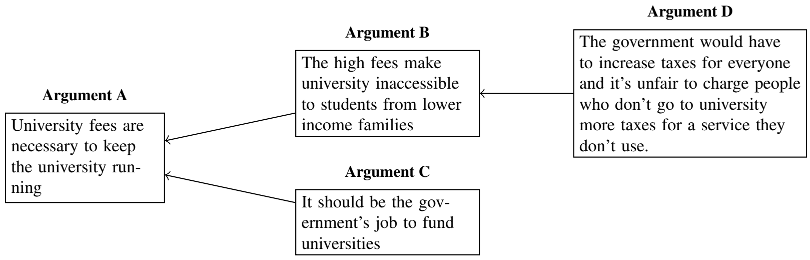
An alternative to the menu-based approach would be a *free-text* approach that allows using a similarity measure to find an argument in the graph that is similar to the user argument. If an argument similar to the one used by the user is present in the graph, the chatbot could simply reply with a counterargument from the graph. Taking the graph from Figure 1 again as an example, the chatbot would present argument A and allow the user to reply via free-text input. The user would counter with an argument that is similar to H. Suppose the chatbot counters it with K and the user replies with an argument that is similar to B. The chatbot could counter it with D or E and so on. In this case, the chatbot can jump around the graph rather than just following a single branch.

However, this poses two questions for the free-text approach: firstly, how to deal with a user argument that is not present in the graph. Not finding a match to the user's argument can be expected to be a common phenomenon given that it cannot be assumed that all arguments on that topic are contained in the graph. The versatility of natural language with its seemingly infinite number of ways to rephrase something, is also likely to limit the ability of the chatbot to find a similar argument in the graph. A second problem is that, even if the user's argument is present in the graph, a counterargument must be chosen so as to increase the persuasive effect of the dialogue.

A potential answer to address the first question is for the chatbot to present an argument that is not necessarily a counterargument to the user's argument. This way the dialogue would resemble argumentation as it would happen in real life between two people: if two human agents engage in an argumentative dialogue, just because one presents an argument the other cannot counter, the dialogue is not necessarily ended prematurely. The other agent might switch topics and present a new argument he or she believes in, without referencing and directly countering the previous argument. Another example would be product reviews where reviewers present a range of pro and con arguments. The judgement is not about whether all counterarguments were answered or not, but whether the pro arguments outweigh the con arguments.

An answer to the second question could come from taking the *concerns* of the user into account [8,14,13], a concern being a matter of interest or importance to the user.

Figure 2. Simple argument graph with arguments *B* and *C* attacking argument *A* and argument *D* attacking argument *B*.



The notion of a concern seems to be similar to the notion of a value. Values have been used in a version of abstract argumentation called value-based argumentation frameworks (VAFs) [3]. For this, when selecting the counterargument, the chatbot could select the counterargument that addresses the more important concerns of the user. In our previous works, however, the concerns of the user were either known in advance [8] or the chatbot did not allow free-text input and the concerns that were addressed by each argument in the graph were known [14,13]. During the chat with a chatbot that allows free-text input, however, the concerns that are addressed by the user arguments need to be classified during the chat in order to choose a suitable counterargument accordingly.

In this paper, we present a free-text chatbot that can engage in an argumentative dialogue in order to persuade the user to accept the bots stance. The chatbot is equipped with a crowd-sourced argument graph with automatically assigned concerns to each argument and a concern classifier that can assign concerns to the user arguments during the chat. With the help of this chatbot, we show that it is not necessary to follow the arcs of a graph during each dialogue move in order to create reasonable and relevant dialogues, and that concerns can be automatically detected and used in order to choose appropriate counterarguments to increase the persuasiveness of the dialogue.

The rest of the paper is structured as follows: Section 2 presents our previous work that the current study builds upon; Section 3 gives the aim of the paper and the hypotheses; Section 4 describes the chatbot architecture that was used for the experiments; Section 5 describes the experiments that were conducted with the chatbot including their results, and in Section 6 we discuss and conclude our findings.

2. Previous work

2.1. Crowd Sourced Argument Graph

The purpose of argumentation is to exchange different viewpoints or opinions, handle conflicting information and make informed decisions. A situation involving argumentation can be represented by a directed graph, as proposed by Dung [11]. Each node represents an argument, and each arc denotes an attack by one argument on another. Such a graph can then be analysed to determine which arguments are acceptable according to some general criteria [4,2]. Figure 2 shows such an argument graph and the attack relationships between the arguments.

Argument graphs are extensively studied in the computational argumentation literature. Their acquisition, however, tends to be neglected. In [7] we present a method of automatically acquiring a large argument graph via crowd-sourcing[1]. We evaluated that method in a case study on the topic of UK university fees. The graph contains 5 levels of depth, starting with the root statement *"University fees in the UK should be kept at 9k pounds"* (Depth 0). The next level of depth (Depth 1) contains arguments that counter the root statement. The arguments in Depth 2 counter the arguments in Depth 1 and so on. Our graph contains 1288 arguments with each argument on average having 3 counterarguments, apart from the last level of depth. Depths 1-4 of the acquired argument graph are used as the knowledge base of the chatbot presented in this paper.

2.2. *Concerns*

We have confirmed the long-held view that taking the concerns of the user into consideration increases the persuasiveness of the dialogue in our previous works [8,14,13]. Arguments can raise or address various concerns for the persuadee that need to be accounted for. A persuader might present a perfectly valid argument to a university student (persuadee), e.g. *"If someone decides to go into higher education, the general public should not be expected to pay for it via taxes."*. The persuadee might not even disagree with this argument, however, she is very likely to be concerned about her finances due to her personal debt and therefore this argument may have no impact on her stance. If, however, the persuader presents an argument that addresses her concern like *"If you have a student loan in the UK, it will not appear on your credit report. So, when you are applying for a credit card, loan or mortgage your student loan will not make an appearance."* it is more likely to change her stance. This is not surprising, however, concerns are often ignored when judging the effectiveness of arguments or choosing a strategy. Some studies that make use of different personality traits of the user attributes in order to evaluate what sort of argument might be more effective for this particular person (for examples see [16,10,21,18]). However, computational argumentation largely focuses on sentimental [9], rhetorical [12] and structural [5] attributes of the argument, rather than attributes about the user.

In the following sections, we outline our hypotheses and describe how we utilise the argument graph and the notion of concerns in order to build a chatbot that can engage in persuasive dialogues, and the experiments conducted with the chatbot.

3. Hypotheses

In this paper, we chose UK university Fees as a case study. We have developed a chatbot that utilises a crowd-sourced argument graph, described in [7], as the knowledge base. The chatbot uses concerns to make strategic moves in order to engage in argumentative dialogues with users to persuade them to accept that chatbot's stance (that university fees should be kept).

Given this setting, we want to test two questions: Firstly, whether the crowd-sourced argument graph can be used as a chatbot knowledge base that allows free-text input. This means that the graph contains at least some common arguments that the user might use,

[1] `https://github.com/lisanka93/Argument_Graph_Corpus`

and the resulting dialogues are therefore of an appropriate length and quality, and that the users perceive the chatbot arguments as relevant. And secondly, whether the chatbot can automatically identify the concerns addressed by the user argument and whether replying with counterarguments that address the same concern, increases the persuasiveness of the chat. We summarise these points in the following two hypotheses:

H1 A crowd-sourced argument graph can be used as a knowledge base for a persuasive chatbot allowing free text input by the users. The resulting chats are of appropriate length and quality, and the chatbot arguments perceived as relevant by the users.

H2 A concern raised or addressed by a given user argument can be automatically identified in order to give appropriate counterarguments that address the same concern and thereby increase the persuasiveness of the dialogue.

In the remainder of this paper we describe the design of our chatbot that was used for the argumentative dialogues and explain the experiments conducted with the chatbot in order to test our hypotheses.

4. Chatbot Design

We developed two versions of our chatbot, one that classifies the concern of the user argument and takes it into account when presenting counterarguments (strategic), and one that did not (baseline).

4.1. Argument Graph

The argument graph described in Section 2.1 is used as the chatbot's knowledge base. We only use the depths 1-4, since depth 5 does not have any counterarguments. Depths 1 and 3 contain arguments against keeping university fees, while depth 2 (attacking depth 1 arguments) and 4 (attacking depth 3 arguments) contain arguments that support the stance of keeping university fees.

When the user types in an argument (source argument), the chatbot uses a similarity measure in order to find the closest match of the user argument in the graph (target argument). We used cosine-similarity as a similarity measure [19]. Cosine similarity is a metric used to measure how similar the vector representation of two texts are. It measures the cosine of the angle between two vectors. The smaller the angle, the higher the cosine similarity. We used a threshold of 0.9 for measuring the similarity of two arguments. If the chatbot finds an argument in the graph that has a similarity of 0.9 or above compared with the source argument, the chatbot chooses one of the counterarguments that attack the target argument in the graph as a response. This happens at every dialogue turn, meaning that the target argument can be either in depth 1 or depth 3 of the graph.

4.2. Default Arguments

In case no target argument is found, we also acquired arguments for keeping university fees, where the root statement is the opposite to our main argument graph *"University fees in the UK should be abolished"*. It is therefore a very shallow graph with only one level of depth where the arguments that attack the root argument are for keeping the

Table 1. Types of concern for the topic of charging university tuition fees

Concern	**Description of what concern deals with**
Student Finance	Finances of students, including tuition fees, student debts, life costs etc.
Government Finance	Government finances, including general taxation, government spending etc.
Employment	Careers and employability of students and the general job market.
Free Education	Whether higher education is a human right and should be free or not.
Fairness	Whether something is fair or not (using a general understanding of fairness), including equal and just treatment of individuals.

fees. We also used crowd-sourcing for the acquisition and voting in order to select the best arguments. The best 7 arguments were used as *default* arguments, which the chatbot can use if no match is found. These arguments are therefore not counterarguments in the traditional sense, as they do not refer to or address the source argument but instead "change topic" and present a new issue in the debate. We also added phrases like *"Ok but", "I still think"* and *"Don't you think that"* to the beginning of the default arguments to indicate a deviation from the topic occurs.

4.3. Concern Labelling and Classification

The baseline chatbot uses the argument graph and default arguments during the chat with the user and does not make use of concerns. The strategic chatbot, however, classifies the concern of the source argument and chooses one of the attackers of the target argument that addresses the same concern.

During the acquisition of the argument graph described in [7], only arguments were included in the graph that contained *topic words*. These are words that we considered meaningful in the given context. The choice of suitable topic words depends entirely on the domain and their choice is left to the researchers' discretion and their knowledge of the domain. The topic words in the argument graph were: *loan, debt, job, tax, free, accessible, affordable, government, scholarship, interest, career* and *background*. We grouped topic words that address the same or similar issues into 5 concerns: **Student Finance** (loan, debt, scholarship, interest), **Government Finance** (government, tax), **Employment** (job, career), **Free Education** (free) and **Fairness** (affordable, accessible, background). Apart from the concern *free*, the concerns were taken from [14]. The definitions are given in Table 1.

We took the arguments from the argument graph, as well as the user arguments from the chats with the baseline chatbot that contained any of the topic words, to train a concern classifier using the Python Scikit-learn library[2]. The classifier uses logistic regression and a tf-idf feature representation in order to predict the concern of the incoming user argument. We extract the top two concern predictions. If the top prediction is over 0.7 the argument is labeled with one concern, otherwise with two. If a target argument in the graph is found, the chatbot chooses one of the attackers of the target argument that addresses the same concern as counterargument. If a user argument is labeled with two concerns, an attacker is chosen that addresses one of the concerns, with priority given to the concern with the higher predicted value.

[2]`https://scikit-learn.org`

It could be argued that since the arguments in the graph are labelled with concerns, the source argument addresses the same concerns as the target argument in the graph and hence no classifier is needed as one could take the concerns of the target argument. However, the concerns of the target argument are not necessarily the same as the user's free-text argument, despite being similar. For example, the target argument in the graph *"Universities should be accessible to all, not just those that can afford it, or are not scared away from the high debt after their studies"* would be labeled with both concerns *fairness* and *student finance*. A similar source argument *"Universities should be accessible to everyone who wants a higher education, not just those that can afford it"* does not address the concern *student finance* and would be labeled with *fairness* only by the classifier.

If no match in the graph is found or none of the counterarguments of the target argument address the same concern, the chatbot replies with a default argument.

5. Evaluation of the Chatbot

The purpose of the chatbots was to test both of our hypotheses. The chatbots were deployed on Facebook via the Messenger Send/Receive API. For more on the implementation of such chatbots see [6]. For each chatbot we recruited 50 participants via Prolific[3], which is an online recruiting platform for scientific research studies. Before the chat the users were directed to a Google Form and asked whether they *strongly disagreed, disagreed, neutral, agreed* or *strongly agreed* that university fees should be kept[4].

After submitting their answers they were redirected to the Facebook page where they could begin the chat. The chatbot started the chat by asking why the user believed that university fees should be abolished. The user, therefore, presented their first argument. The chatbot then replied with either a counterargument from the argument graph or a default counterargument, depending on whether a similar argument was found in the graph or not. If a similar match was found, the baseline chatbot replied with a randomly selected counterargument from the direct attackers of the target argument in the graph. The strategic chatbot, however, selected an attacker from the graph that addressed the same concern as the user argument (if such an argument exists). If no match was found, both chatbots replied with a default argument.

If the user response was shorter than 6 words, the chatbot queried the user to expand on their answer. However, if the user agreed with an argument the chatbot gave, for example by sending *"I agree"*, the chatbot would not ask to expand despite the message being shorter than 6 words, and instead replied with a default argument.

The chatbot would eventually end the chat as soon as all default arguments were used up and no match in the graph was found. The users were, however, advised that they could end the chat anytime by sending the word *"stop"*. At the end of the chat the chatbot presented the user with a link that redirected them to a second Google Form where they were asked a series of questions[5]:

---

[3] `https://prolific.co`

[4] For the baseline chatbot only 2 people selected *agree* and none for the strategic one. 98% of participants therefore did not share the chatbots stance before the chat.

[5] Further questions were asked but analysis of the answers is left to future work

Table 2. Answers to first three questions for baseline and strategic groups

Chatbot	Understood (Q1)			Relevant Args (Q2)			Points addresses (Q3)		
	Yes	No	Sometimes	Yes	No	Some	Yes	No	Some
Baseline	16	4	30	21	3	26	13	15	22
Strategic	15	6	29	31	1	18	10	14	26

1. Did you feel understood by the chatbot? (Yes/No/Sometimes)
2. Did you feel that the chatbot's arguments were relevant? (Yes/No/Sometimes)
3. Do you feel like all your points were addressed? (Yes/No/Some of them)
4. How much do you agree that fees in the UK should be kept as they are? (Strongly disagree - strongly agree)

Questions 1-3 were used to test our first hypothesis and judge the relevance, length and quality of the chats, and question 4 was to test our second hypothesis and compare the persuasiveness of the baseline chatbot to the strategic chatbot. Table 2 shows the results for the first three questions for the baseline and the strategic groups. One can see that the majority of the participants considered the chatbot's arguments as relevant in most cases, and answered the first three questions with either *yes* or *sometimes*. Interestingly there is a 50% increase in the perception of relevance for the strategic chatbot, while the numbers for questions 1 and 3 remained almost the same. This is a statistically significant difference with a p-value of 0.045 using Chi-Square. Using concerns, therefore, makes the arguments more relevant.

Regarding questions 1 and 3, given that the chatbot does not use natural language generation and only relies on finding matches in an argument graph, it is not able to address novel arguments or expand on existing ones by giving more information. It is therefore not surprising that the numbers for these questions remained approximately the same. The given results, however, are promising given that the chatbot solely relies on a similarity measure in order to pick counterarguments from a crowd-sourced graph or to pick default arguments in case no match is found.

Regarding the length of the chats, chats with the chatbots on average lasted 24 turns (disregarding the chatbot querying the user to expand). This means that the chatbot, on average, gave 12 arguments, 7 of which were default arguments and 5 from the graph. This supports our first hypothesis that a crowd-sourced argument graph can be used as a chatbot knowledge base and that the resulting argumentation dialogues are of satisfactory length and quality, with perceived relevance of the arguments being 50% higher during chats with the strategic chatbot. An example of a chat can be seen in Figure 3. All chatlogs and the code for the chatbot can be found in [1].

Using the replies for question 4 we calculated the change in stance for both groups in order to evaluate the persuasiveness of the strategic chatbot. The persuasion goal of the chatbot was that UK university fees should be kept. Since participants were given the choice of 5 stances ("strongly disagree" to "strongly agree") before and after the chat, they could either change their stance to the better, to the worse or not at all. The change in stance is the final stance minus the original stance. We call the units of this measure *change in stance* (CS) points. For example, "strongly disagree" being 1 and "strongly agree" being 5, if a participant changed her stance from "disagree" (2) to "neutral" (3) after chatting with the bot this counts as 1 CS point, whereas changing from "neutral" to "disagree" counts as -1 CS point. Persuasion means that a person who was against

Figure 3. Example chat between a participant and the strategic chatbot. Chatbot arguments are in the dark boxes and user arguments in the light boxes. Default arguments are indicated with an *. The chat begins in the left column and continues in the right one.

So tell me Why do you think university fees in the UK should be abolished?

It means students do not have to come from a privileged background in order to study. Everyone will have an equal opportunity

They all go to school and get equal knowledge and opportunities, it is got nothing to do with the family background. Whilst at high school, student get all the information they need to go further

But someone who leaves school without going to university will have lower work opportunities as most jobs require a degree or

We have to give everyone equal chance to succeed and end up with the best job they can, naturally there will be those who reach the top, and those who don't.

Students would be able to focus more on their education and not have to worry about the fees and debt they'll get into

There are loans about so no one pays fees up front anyway, only when they have secured a job.

But when you leave university you don't want to have to be burdened with debt for the rest of your lift because they are so large

That's why the government should start giving scholarships to people who really can't pay university fees but still want to get a degree.

But just how much would taxes have to increase to be able to afford everyone having that equal opportunity

After graduation, many people get a much better job so they can afford to pay for loans. This is a normal bird in many European countries

But there is no guarantee that everyone will end up with a decent job with their being an increase in competition for jobs

Better education does not mean it's free. But better education actually leads to better performance and the possibility of better jobs.

But if a student leaves without the debt from university it will mean they spend more upon leaving by having the financial freedom to afford housing etc

You do not need to pay straight away, the University is accessible as never before for everyone; young, mature, from less privileged backgrounds etc.

No but you do have to pay when you leave and that money could be saved from their high paying job and be used to buy a home and benefit the economy in other ways with more disposible income

It gives them an objective to accomplish, it gives value to their degree and future career. getting free education is good and necessary in some cases but that also gives people the option to slack off and reduce their chances of getting a proper work ethic.

It doesn't stop students who wish to pay for a degree from a higher valued institution from doing so though

University is not a right. Most people fall into university because they think they have too. Going to university should be an important decision and taking a loan should be part of this.

No because at school we are not taught the value of money and throwing a student into debt should not be a lesson.

* But don't you think that if someone decides to go into higher education, the general public should not be expected to pay for it via taxes.

But they'd be expected to pay for it as scholarships

* Ok but I think university fees make students take the university seriously. If it would be free, it wouldn't have the same value.

But you are taking away the freedom of choice from someone who wants to go but cannot because they cannot afford it

It is understandable that it is a concern that people need to consider if they can afford it but a student loan is an investment in themselves and their future

So by offering free tuition they wouldn't need to have to consider forgoeing further education

* I believe the fee is a way of selecting people that really want to get in the better universities and invest in their future.

That doesn't stop the 'better' universities from charging a fee though. If someone really wants a prestigious degree they can pay for it and let those who wish to have an education study for free

* I think the fee is a good incentive for students to finish the degree. We see models in several european countries where uni- versity is free, and the result is that many either never finish their studies , or take an obscene number of years to finish.

If they are taught in school about the competition for jobs then more people will take it seriously because a better education could be the difference between a low and a high paid job

You cannot only charge fees for certain degrees as that would have far reaching consequences. People working in lower paid jobs may never have to repay any of the loan anyway if they do not earn above the threshold.

But they always have that debt burden over them and know that they will have to start paying at some point in the future

The argument that not every student will find a financial situation to pay off the loan does not appeal to me, because in adult life you have to manage your finances in such a way as to be a plus.

There is a huge difference in being able to manage your finances and being given a debt worth tens of thousands to start your future

The debt doesn't count against someone's credit score and is largely ignored, if they started earning less they would stop paying it back. Percentage wise there are more people going to university

stop

Table 3. Change of stance measured by number of participants and CS points

Chatbot	Baseline			Strategic		
Change in stance	Negative	No Change	Positive	Negative	No Change	Positive
No. of participants	5	41	4	1	26	23
Change in CS points	-5	0	5	-1	0	32

keeping university fees before the chat changed her stance to the positive and that her CS points score is positive.

Table 3 shows the number of participants who changed their stance to the worse (negative), to the better (positive), and that did not change their stance at all (no change) for both chatbots, as well as the number of total CS points. We can see that 23 people changed their stance to the better when chatting with the strategic chatbot with a total of 32 CS points, meaning that some participants changed their stance by more than 1 CS point (e.g. from *disagree* to *agree*). If counting the total number of CS points, also including the participants that changed their stance to the worse, the strategic chatbot achieved a total change of 31 CS points whereas for the baseline the total number of CS points is 0.

It could be argued that a change from *strongly disagree* to *disagree* is not a remarkable change in stance despite resulting in the change of 1 CS point, whereas changing someone's stance from *disagree* to *neutral* or even better, *agree* is a much stronger shift in stance. However, for the strategic chatbot, only 2 participants changed their stance from *strongly disagree* to *disagree*, while the remaining 21 participants changed their stance from disagreement (strongly or not) to neutral (16 participants), from neutral to agreement (3 participants) and from disagreement to agreement (2 participants).

We used the number of participants who changed their stance to the positive in order to calculate the statistical significance of the difference between the control group that chatted with the baseline chatbot and the group that chatted with the strategic chatbot using the Chi-Square test. All results were statistically significant with a p-value of 0.00017. The results support our second hypothesis, that concerns can be automatically classified based on the use of topic key words which can be seen as a good indicator of the concerns being addressed or raised by the arguments. Presenting arguments that address the user's concern is more likely to have a positive impact on their stance, than presenting arguments that ignore the user's concern.

6. Discussion

Our contribution in this paper is twofold. Firstly, we have shown that a crowd-sourced argument graph can be utilised as a knowledge base for a chatbot that engages in argumentative dialogues. The resulting chats are of good length and quality and are perceived as relevant by the users. And secondly, we have shown that concerns can be automatically identified in order to give suitable counterarguments that address the same concern and thereby significantly increase the persuasiveness of the dialogue. Additionally, we have shown that the chatbot can jump around in the graph, without systematically following each arc and only use arguments that are connected via an attack relationship.

To date, at least two arguing chatbots have been presented in the literature: a chatbot Debbie, that uses a similarity algorithm to retrieve counterarguments [17] and Dave that used retrieval- and generative-based models [15]. Debbie's knowledge base consists of a subset of the qualitatively best arguments from the corpus created by Swanson et al [20] which is a combination of online political debates, Internet Argument Corpus (IAC), [22] and dialogues from online debate forums. Dave's knowledge base consists only of the IAC. Our chatbot, however, is different in several ways: firstly, our knowledge base consists of a previously crowd-sourced argument graph. And secondly, the aim of Dave/Debbie was to keep the conversation going, whereas we were interested in persuading the user to accept our chatbot's stance.

This study can be seen as a partial extension of the work in [14] where a chatbot was used to persuade the user to accept the chatbot's stance on the topic of university fees in the UK. The argument graphs that were used as the chatbot's knowledge base were hand-crafted and manually labeled. The chatbot also did not allow free-text input and was strictly following the arcs of the argument graph. The chatbot presented in this paper allows free-text input and uses a similarity measure to extract similar arguments from the graph and does not restrict the selection of arguments to a single path in the graph. If a match is not found, the chatbot replies with an argument that is not contained in the original graph. Our evaluation showed this approach performed well and shows that it is not necessary to and, in fact, often impossible to establish all possible relationships in a big argument graph. Therefore, instead of following a single path through the graph and only allowing the user to choose arguments that are present in the graph, one can search for a similar argument at each dialogue step without relying on a connecting arc between the new user argument and the previously given chatbot argument. And to avoid ending the chat prematurely if no similar user argument is found, default arguments can be introduced to keep the chat going.

We faced the additional challenge of having to automatically identify the concern of the user arguments during the chat. We showed that by grouping the most common meaningful words of the argument graph (topic words) into concerns, one can train a concern classifier on the graph arguments that can be used by the chatbot in order to improve its persuasive effect.

The advantage of using a crowd-sourced argument graph as a knowledge base is that it does not require professional research but solely relies on the input of participants and can be acquired quickly. This method also scales easily which allows obtaining many arguments from different people, and thereby create large and comprehensive argument graphs. There are, however, also potential risks to consider. For example, the spread of invalid arguments which, despite being popular, might contain wrong information. Therefore, in the future, we want to investigate methods on how to utilise the argument graph to improve the quality and persuasive effect of the chats even further. The chatbot could, for example, identify invalid or unpopular arguments and delete them from the graph. The bot could also learn which are the more persuasive arguments and use those more often in the future.

Acknowledgements

The first author is funded by a PhD studentship from the EPSRC. The authors would like to thank Sylwia Polberg for valuable feedback on earlier versions of this paper.

References

[1] https://github.com/lisanka93/university_fees_bot.

[2] P. Baroni, M. Caminada, and M. Giacomin. An introduction to argumentation semantics. In *Knowledge Engineering Review 26(4)*, pages 365–410, 2011.

[3] T. J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[4] P. Besnard, A. J. García, A. Hunter, S. Modgil, H. Prakken, G. Simari, and F. Toni. Introduction to structured argumentation. *Argument and Computation*, 5(1):1–4, 2014.

[5] F. Boltuzic and J. Snajder. Back up your stance: Recognizing arguments in online discussions. In *Proc. of the First Workshop on Argumentation Mining*, pages 49–58, 2014.

[6] L. A. Chalaguine and A. Hunter. Chatbot design for argument harvesting. In *Proc. of the Workshop on Systems and Algorithms for Formal Argumentation COMMA'18*, pages 457–458, 2018.

[7] L. A. Chalaguine and A. Hunter. Knowledge acquisition and corpus for argumentation-based chatbots. In *Proc. of the 3rd Workshop on Advances In Argumentation In Artificial Intelligence*, pages 1–14, 2019.

[8] L. A. Chalaguine, A. Hunter, F. L. Hamilton, and H. W. W. Potts. Impact of argument type and concerns in argumentation with a chatbot. In *Proc. of the 31st International Conference on Tools with Artificial Intelligence*, pages 1557–1562, 2019.

[9] Y. Choi and C. Cardie. Learning with compositional semantics as structural inference for subsentential sentiment analysis. In *Proc. of the Association of Computational Linguistics*, pages 793–801, 2008.

[10] T. Ding and S. Pan. Personalized emphasis framing for persuasive message generation. In *Proc. of the European Chapter of Computational Linguistics*, pages 1432–1441, 2016.

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

[12] I. Habernal and I. Gurevych. Which argument is more convincing? Analyzing and predicting convincingness of web arguments using bidirectional LSTM. In *Proc. of the Association of Computational Linguistics*, pages 1589–1599, 2016.

[13] E. Hadoux and A. Hunter. Comfort or safety? Gathering and using the concerns of a participant for better persuasion. *Argument & Computation*, pages 1–35, 2019.

[14] A. Hunter, S. Polberg, and E. Hadoux. Strategic argumentation dialogues for persuasion: Framework and experiments based on modelling the beliefs and concerns of the persuadee. Technical report, University College London, 2020.

[15] D. T. Le, C. Nguyen, and K. A. Nguyen. Dave the debater: a retrieval-based and generativeargumentative dialogue agent. In *Proc. of the 5th Workshop on Argument Mining*, pages 121–130, 2018.

[16] S. Lukin, P. Anand, M. Walker, and S. Whittaker. Argument strength is in the eye of the beholder: Audience effect in persuasion. In *Proc. of the European Chapter of Computational Linguistics*, pages 742–753, 2017.

[17] G. Rakshit, K. K. Bowden, L. Reed, A. Misra, and M. Walker. Debbie,the debate bot of the future. In *Advanced Social Interaction with Agents - 8th International Workshop on Spoken Dialog Systems*, pages 45–52, 2017.

[18] R. Santos, G. Marreiros, C. Ramos, J. Neves, and J. Bulas-Cruz. Using personality types to support argumentation. In *Proc. of the Sixth International Workshop on Argumentation in Multi-Agent Systems*, pages 292–304, 2009.

[19] A. Singhal. Modern information retrieval: A brief overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 24:35–43, 2001.

[20] R. Swanson, B-Ecker, and M. Walker. Argument mining: Extracting arguments from online dialogue. In *Proc. of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 217–226, 2015.

[21] S. Villata, E. Cabrio, I. Jraidi, S. Benlamine, M. Chaouachi, C. Frasson, and F. Gandon. Emotions and personality traits in argumentation: An empirical evaluation. *Argument and Computation*, 8(1):61–87, 2017.

[22] M. A. Walker, P. Anand, J. F. Tree, R. Abbot, and J. King. A corpus for research on deliberation and debate. In *Proc. of the 8th International Conference on Language Resources and Evaluation (LREC)*, pages 812–817, 2012.