

# Towards a framework for computational persuasion with applications in behaviour change<sup>1</sup>

Anthony Hunter

*Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK*  
E-mail: [anthony.hunter@ucl.ac.uk](mailto:anthony.hunter@ucl.ac.uk)

**Abstract.** Persuasion is an activity that involves one party trying to induce another party to believe something or to do something. It is an important and multifaceted human facility. Obviously, sales and marketing is heavily dependent on persuasion. But many other activities involve persuasion such as a doctor persuading a patient to drink less alcohol, a road safety expert persuading drivers to not text while driving, or an online safety expert persuading users of social media sites to not reveal too much personal information online. As computing becomes involved in every sphere of life, so too is persuasion a target for applying computer-based solutions. An automated persuasion system (APS) is a system that can engage in a dialogue with a user (the persuadee) in order to persuade the persuadee to do (or not do) some action or to believe (or not believe) something. To do this, an APS aims to use convincing arguments in order to persuade the persuadee. Computational persuasion is the study of formal models of dialogues involving arguments and counterarguments, of user models, and strategies, for APSs. A promising application area for computational persuasion is in behaviour change. Within healthcare organizations, government agencies, and non-governmental agencies, there is much interest in changing behaviour of particular groups of people away from actions that are harmful to themselves and/or to others around them.

**Keywords:** Computational persuasion, persuasion dialogues, persuasive arguments, dialogical argumentation, computational models of argument, probabilistic argumentation, argumentation strategies

## 1. Introduction

Persuasion is an activity that involves one party trying to get another party to do (or not do) some action or to believe (or not believe) something. It is an important and multifaceted human facility. Obviously it is very important in commercial activities, but it is also important in professional life, and indeed, in everyday life. In whatever we do, we frequently find ourselves trying to persuade other people with regard to something that is important to us, and/or them.

In this paper, I discuss some aspects of the notion of persuasion, and explain how this leads to the idea of computational persuasion. Computational models of argument are central to the development of computational persuasion. I briefly review some key aspects of computational models of argument, and highlight some topics that need further development. I then briefly cover behaviour change as a field in which we can apply methods from computational persuasion, and evaluate the performance in the field.

---

<sup>1</sup>This paper is an extended version of Anthony Hunter (2016) Computational Persuasion with Applications in Behaviour Change, Computational Models of Argument (COMMA'16), pages 5–18, IOS Press.

## 2. What is persuasion?

The aim of persuasion is for the persuader to change the mind of the persuadee. By persuasion, the persuader directs the persuadee to believe (or disbelieve) something that the persuader would like the persuadee to believe (or disbelieve), and this in turn may result in them doing (or not doing) something that the persuader would like to be brought about.

- In **unidirectional persuasion**, the persuader presents a message to the persuadee with the aim of persuading the persuadee, and there is no facility for the persuadee to respond to the message. Examples of unidirectional persuasion include product advertising (e.g. T.V. or magazine advert), political speech (e.g. prior to an election), and government advisory messages (e.g. recycle paper, metal, etc).
- In **bidirectional persuasion**, the persuader and persuadee enter into a dialogue, with the persuader aiming to persuade the persuadee. The fact that the persuader is trying to persuade the persuadee is often clear from the types of interaction, but sometimes, the persuader tries to hide the fact from the persuadee. Examples of bidirectional persuasion include sales meetings (e.g. in car showroom), some kinds of medical counselling (e.g. on drug abuse), and everyday discussions with a goal (e.g. an employee asking for a payrise).

Some kinds of interaction surrounding persuasion include: The persuader collecting information, preferences, etc from the persuadee; The persuader providing information, offers, etc to the persuadee; The persuader winning favour (e.g. by flattering the persuadee, by making small talk, by being humorous, etc); But importantly, arguments are the essential structures for presenting the claims (and counter claims) in persuasion.

An argument-centric focus on persuasion leads to a number of inter-related aspects (see Section 3) that need to be taken into account, any of which can be important in bringing about successful persuasion. These dimensions can significantly affect the success of argumentation, and therefore go some way to delineating what constitutes a good approach to persuasion. But, before considering these dimensions, I would like to make the following claim.

- **Persuasion is not normative** There are no underlying rules or principles to the use of argumentation in persuasion. This means for instance that arguments can be inconsistent, irrational, untrue, etc. if they persuade. Though inconsistent, irrational, untrue arguments may be counter-productive with some audiences, as well as being potentially problematic from moral, ethical, and regulatory perspectives. Furthermore, persuasion takes on many forms (such as coax, cajole, inveigle, wheedle, lure, coerce, prevail on, proselytize, and urge), and arguably, some of these are not normative.

A corollary of the above claim is that how convincing an argument is does not equal how correct it is. For example, arguments like *homeopathy focuses on processes of health and illness rather than states, and therefore it is better than regular medicine* and *the sheer weight of anecdotal evidence gives rise to the common-sense notion that there must be some basis for homeopathic therapies by virtue of the fact that they have lasted this long* can be convincing for some audiences.

## 3. Levers for persuasion

We now consider some key dimensions that can affect the success of a persuader in persuading a persuadee. We focus on the dimensions that the persuader may have some control over such as rationality

of argumentation, persuasion techniques, argumentation style, framing of the arguments, and emotions invoked by arguments. We consider each of them in the following subsections.

### 3.1. Rationality of argumentation

The study of argumentation has largely focused on what constitute good arguments in a normative sense. This involves identifying the features of good arguments that would be appropriate for honest rational agents to present or accept. For an excellent review, see [146]. The careful presentation of premises, and the use of logical reasoning, is much espoused for good quality argumentation. In order to make a case in professional life (such as politics, academia, business, journalism, etc), it is seen as an essential ability. This has led to techniques for constructing and deconstructing good arguments (see for example [48,66]).

In parallel with elucidating what constitutes a good argument, the study of argumentation has also identified types of poor or inappropriate argumentation such as argumentation fallacies (as well as [146], see reviews in [58,145]). Some fallacies are described as formal fallacies since they violate rules of logic (e.g. claiming the antecedent of an implication is true because its consequent is true), or violate rules of probability theory (e.g. the gambler's fallacy), but many fallacies are informal since they constitute what could be described as unacceptable arguments, such as begging the question (which is providing a version of the claim as a premise, and using that to derive the claim), *argumentum ad hominem* (arguing against the arguer rather than their arguments), and appeal to authority (arguing that an argument is true because of the position of the arguer).

Given that the study of argumentation has developed such a comprehensive understanding of how to differentiate good from bad argumentation, if we are dealing with rational agents when undertaking persuasion, it appears important to use good arguments and argumentation, and avoid poor arguments and argumentation.

Furthermore, the overall *quality* of the argumentation is important from a psychological point of view (see for example [64,67]). If a persuader wants to convince the persuadee of an argument (a persuasion argument), then this includes acceptability of the persuasion argument (against counterarguments), believing the premises of the persuasion argument, fit of persuasion argument with agenda, goals, preferences, etc, quality of constellation of arguments considered (balance, depth, breadth, understandability, etc). Comprehensibility of persuasive arguments has also been shown to be an important determinant of successful persuasion [45].

Despite the importance of being rational in argumentation, there is a tendency in the computational models of argument community to overemphasize the need to be rational. As we will cover later, the role of emotional arguments have been found from psychological studies to be important in persuasion. Furthermore, even for rational argumentation, some widely accepted principles such as not deploying *ad hominem* arguments can be over-restrictive when for example the speaker has a poor reputation or is unqualified in the topic that they are speaking on, though use of argument schema with critical questions goes someway to redressing the balance in this (for a review of argument schema, see [148]).

### 3.2. Persuasion techniques

Psychological studies have identified persuasion techniques that are seen in human interactions in general. For instance, Cialdini has identified the following six principle of influence [34].

- **Reciprocation** Creating an obligation to give or to receive or to repay (e.g. doing a small favour for someone is more likely to result in a big favour being obtained in return).

- **Consistency** People like to be seen to be consistent with their values and identity, and so getting a general commitment first can be advantageous (e.g. getting expressed support for a cause, prior to asking for material support is more likely to be successful).
- **Social proof** People like to conform with those around them (e.g. treating dog phobia in children by showing videos of children playing happily with dogs).
- **Liking** People prefer to comply with people they like (e.g. finding common ground, pleasantries, etc, prior to attempting to persuade, is more likely to be successful).
- **Authority** Persuasion by an expert or with evidence can be more effective than persuasion by a non-expert or without evidence (e.g. a doctor persuading someone to use high factor sunblock is more likely to be successful than a lawyer).
- **Scarcity** People are attracted to something that is scarce (e.g. increasing the price of a product can increase the demand for the product since it appears to be more desirable).

The above principles are supported by empirical studies in psychology. Other studies suggest further principles that appear to be useful in various domains such as law (e.g. [52]), healthcare (e.g. [133]), and ecommerce (e.g. [105]).

### 3.3. Argumentation style

Argumentation style concerns who is presenting the arguments, the language used in those arguments, and the way the dialogue is structured. The issues raised touch on broader aspects of the personality of the persuadee, and the context of the persuasion (see for example [132] for further discussion of these issues).

- **Nature of persuader** The nature of the persuader can be important. From a rational perspective, seemingly good features of a persuader are that s/he has relevant authority, expertise, or knowledge, and seemingly poor features of a persuader are that s/he is attractive, witty, or a celebrity. However, in practice, different persuadees respond to different features. For instance, a teenager is unlikely to be convinced by a government safety expert to wear a helmet when on a bike, but may be influenced by a celebrity to do so. So it is clear that for successful persuasion with many people, it can be more important who the persuader is than what the arguments are. This observation is borne out by how often celebrities are used in advertising products, or taken seriously in politics. Psychological studies show that the degree of attitude change is influenced by the degree of positive feelings towards the proponent [86].
- **Language of arguments** The choice of language in argumentation can be important. This goes from choice of words (e.g. use of *freedom fighter* versus *terrorist*), to choice of metaphor (e.g. using *the whole world's a stage* when persuading someone to do something bold), the use of metonymy (e.g. Oscar Wilde on fox hunting *The unspeakable in pursuit of the uneatable*) or use of irony [35]. Use of language in advertising and in politics is extensively studied, and there are guidebooks that elucidate some of the well-established principles (see for example [90]). The use of persuasive language has also been studied in the media. Newspapers will deploy particular kinds of phraseology and language construction to put forward their viewpoint. For example, critical discourse analysis of war reporting shows how the side that the newspaper supports (respectively does not support) will use terms such as suppress (respectively destroy), eliminate or neutralize (respectively kill), reporting restrictions (respectively censorship), press briefings (propaganda), preemptive (respectively without provocation), etc., and furthermore, critical discourse analysis can reveal differences

in the sentence construction, use of modality, presupposition, rhetorical style, and overall narrative [125].

- **Selectivity of argumentation** Persuasion does not involve exhaustive presentation of all possible arguments [16]. Rather it requires careful selection of arguments that are most likely to be efficacious in changing the mind of the persuadee. Deciding on which arguments to select depends on diverse features of the arguments and the persuadee such as the nature of the persuader, the language of the arguments, use of psychological techniques, personality of the persuadee, use of rational and/or emotional argumentation, etc. Being selective does not mean that argumentation needs to be constrained in any way other than being the most efficacious for persuasion.

Determining the precise parameters of argumentation style raises some complex issues for effective persuasion in a specific application. Experiential knowledge and psychological studies offer some general guidance but in practice the specific parameters need to be determined for a specific application.

### 3.4. Framing of arguments

The **framing style** of an argument concerns the exact phrasing of the content which can influence how it is perceived. It is increasingly acknowledged in healthcare professions that interactions with patients should be effective communication that is respectful and non-judgmental (see for example [106]). Such a stance would be a natural point for automated persuasion systems. However, this still leaves much latitude in how arguments can be framed.

Consider for example experiments by Tversky and Kahneman where participants are asked to imagine preparations for the outbreak of an unusual disease that is expected to kill 600 people, and to consider two alternative programmes to combat the disease [141]. In the first framing of the study (below), the majority of participants prefer Programme A.

- If Programme A is adopted, 200 people will be saved.
- If Programme B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

In contrast, in the second framing of the programmes (below), the majority of participants prefer Program D. So even though, the versions above and below are providing the same information, participants can change their preferences based on the framing. The explanation is that in the context of gain (above) people tend to be averse to risk and in the context of loss (below) they try to minimize the loss.

- If Programme C is adopted 400 people will die.
- If Programme D is adopted there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

In a user study on the persuasiveness of healthy eating messages [138], positively framed messages (e.g. Most people believe that eating a healthy breakfast contributes to a longer lifespan) were shown to be more efficacious than negatively framed messages (e.g. Most people believe that eating an unhealthy breakfast contributes to a shorter lifespan). Furthermore, Cialdini's principles of persuasion [34] were considered (i.e. reciprocity, commitment, consensus, liking, authority, and scarcity), and it was found that arguments that appeal to authority (e.g. Studies conducted by health experts have shown that eating a healthy breakfast keeps you energized) were the most persuasive.

### 3.5. Personality of persuadee

Another key dimension that can affect the success of persuasion is the personality of the persuadee. Obviously, this is not in the control of the persuadee, but if the persuader knows about the personality of the persuadee, the persuader can make a better choice of strategy to use with the persuadee. Consider for example persuading someone to vote in the national election: If the person “follows the crowd”, then telling them that the majority of the population voted in the last election is more likely to get them out to vote, whereas if the person “follows rules rigorously”, then telling them that it is their duty to vote is more likely to get them out to vote. Mistaking the personality trait can have a negative effect on the chances of successful persuasion. Psychology has developed numerous methods for characterizing personality. An important example is the model based on the OCEAN personality traits which are Openness to experience, Conscientiousness, Extroversion, Agreeableness, and Neuroticism [53].

There are many other ways that knowing more about the persuadee can be used in determining an appropriate strategy for persuasion. Attitudes are important determiners in persuasion, and the psychology of attitudes may provide important insights in a persuasion strategy (for a review see [97]). Various other aspects of psychology are routinely harnessed in commerce, particularly ecommerce, such as the psychology of colour, language, graphics, pricing, and negotiation (for a review see [105]). Cultural information about a persuadee, such as nationality, can give important information about the kinds of interaction that a persuadee might respond positively to (see for example [65]). Other kinds of traits, such as political, ethical, or religious, may be important determiners in some contexts for persuasion. Furthermore, detailed information about multi-dimensional private traits can be predicted from digital records of the user (e.g. from Facebook likes) [88].

### 3.6. Emotion invoked by arguments

Presenting emotional arguments can be important. Emotional arguments are predominantly deployed by the persuader in order to influence the persuadee via emotional devices. For example,

- *You have a good income, and so you should feel guilty if you do not donate money to this emergency appeal by Médecins Sans Frontières.*
- *Your parents will be proud of you if you complete your thesis and get your PhD.*

Note, emotional arguments contrast with evidential/logical arguments, such as for example, the following argument.

- *You will have a much higher chance of getting a highly paid job if you complete your thesis and get your PhD award.*

In certain situations, emotional arguments can be powerful arguments in persuasion. For instance, Lukin *et al* [95] have shown that with some audiences, emotional arguments are more effective in persuasion than factual arguments. For this, they categorized audiences according to the OCEAN personality traits (i.e. openness to experience, extroversion, agreeableness, conscientiousness, and neuroticism), and showed that conscientious, open, and agreeable people are more convinced by emotional arguments.



#### 4. What is computational persuasion?

As developments in artificial intelligence attempt to capture more aspects of human cognition, it is natural to consider how persuasion can be captured as a software process. In the following, I define the notion of an automated persuasion system, and use this to define computational persuasion.

- An **automated persuasion system** (APS), i.e. a persuader, is a system that can engage in a dialogue with a user, i.e. a persuadee, in order to persuade that persuadee to do (or not do) some action or to believe (or not believe) something. To do this, an APS aims to use convincing arguments in order to persuade the persuadee. The dialogue may involve moves including queries, claims, and importantly, arguments and counterarguments, that are presented according to some protocol. Whether an argument is convincing depends on the context, and on the characteristics of the persuadee. An APS maintains a model of the persuadee, and this is harnessed by the strategy of the APS in order to choose good moves to make in the dialogue.
- **Computational persuasion** is the study of formal models of dialogues involving arguments and counterarguments, of persuadee models, and strategies, for APSs. Therefore, developments in computational persuasion build on computational models of argument. Note, the aim of computational persuasion is not to produce models of human persuasion (c.f. [19]), rather it is to produce models of persuasion that can be used by computers to persuade humans, and that they can be shown to have a reasonable success rate in some persuasion task (i.e. that a reasonable proportion of the users are persuaded by the arguments and therefore do the action or accept the belief).

Clearly, persuasion is a complex and fascinating phenomenon. Furthermore, it is very important for humans to be able to persuade, and to be persuaded. However, this does not mean that it would be best for both parties (i.e. persuader and persuadee) if the persuader is always successful in persuasion. Sometimes it will be good for the persuadee but not always, and so the persuadee needs to judge whether or not to agree. This raises some interesting challenges for developing computational persuasion. In the rest of this review, I am only able to touch on some of the issues. I will proceed by considering computational models of argument (i.e. computational argumentation), and discuss features that will be useful for computational persuasion, and highlight some shortcomings.

##### 4.1. What do computational models of argument offer?

Computational persuasion is based on computational models of argument. These models are being developed to reflect aspects of how humans use conflicting information by constructing and analyzing arguments. A number of models have been developed, and some basic principles established. We can group much of this work in four levels as follows (with only examples of relevant citations).

###### 4.1.1. Dialectical level

Dialectics is concerned with determining which arguments “win” in some sense. In abstract argumentation, originally proposed in the seminal work by Dung [42], arguments and counterarguments can be represented by a graph. Each node denotes an argument, and each arc denotes one argument attacking another argument. Dung defined some principled ways to identify extensions of an argument graph. Each extension is a subset of arguments that together act as a coalition against attacks by other arguments. An argument in an extension is, in a sense, acceptable. For a review of Dung’s approach and alternatives, see [8]. Labels can also be assigned to arguments, and Caminada and Gabbay [27] provided a formalization based on three labels (in, out and undecided) that they show is equivalent to Dung’s formalisation.

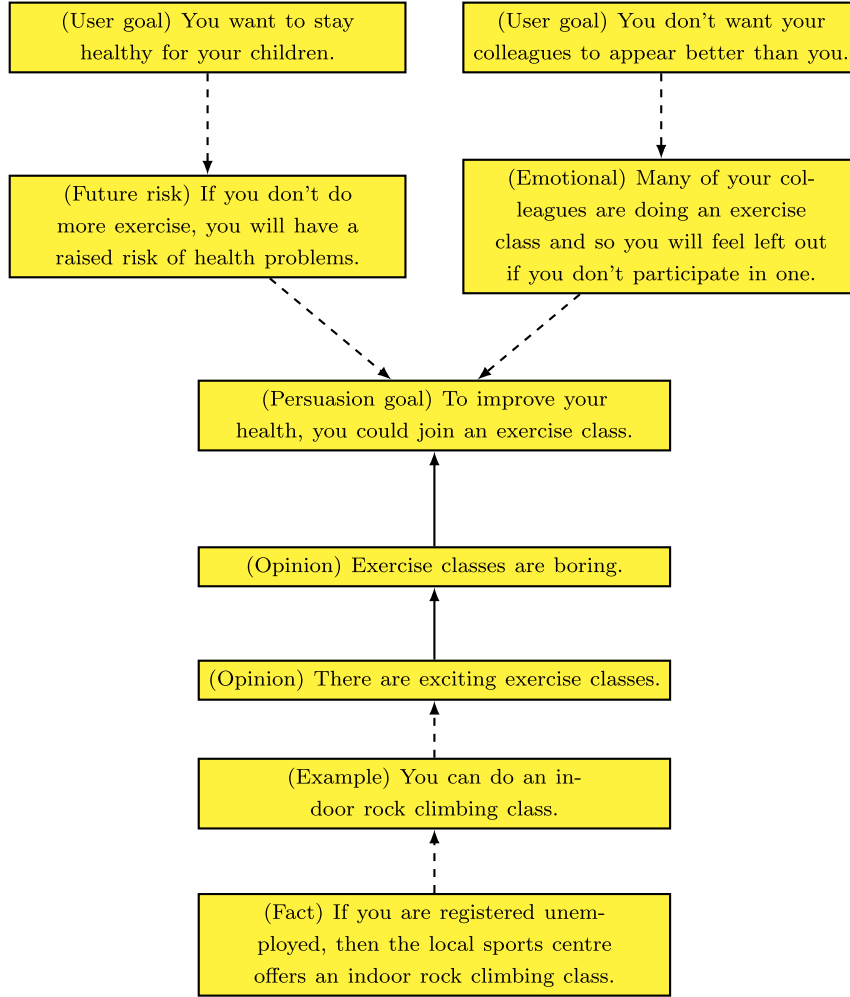


Fig. 1. A simple example of a bipolar argument graph for application for persuading a user to undertake more exercise. A dotted arc from a node A to a node B denotes that argument A supports argument B. A solid arc from a node A to a node B denotes that argument A attacks argument B (i.e. A is a counterargument B).

There have been numerous developments of abstract argumentation that include alternative definitions for extensions such as ranking-based semantics [1,24,120], and introduction of attacks on attacks [7, 103], preferences [2], weights on attacks [44], support relations (see for example [30,31,109] and see Fig. 1 for an example of an argument graph with supporting and attacking arguments), and probabilities [43,72,74,82,92,135]. Furthermore, there has been the development of software solvers for determining extensions (see for example [33,137]), and the application of natural language processing techniques for constructing argument graphs from free text (see for example [93]). In addition, there are methods for argument dynamics to ensure that specific arguments hold in the extensions of the argument graph such as epistemic enforcement in abstract argumentation [9,10,38], revision of argument graphs [36,37], and belief revision in argumentation (e.g. [18,29,41,50]).



#### 4.1.2. Logical level

At the dialectical level, arguments are atomic. They are assumed to exist, but there is no mechanism for constructing them. Furthermore, they cannot be divided or combined. To address this, the logical level provides a way to construct arguments from knowledge. At the logical level, an argument is normally defined as a pair  $\langle \Phi, \alpha \rangle$  where  $\Phi$  is a minimal consistent subset of the knowledgebase (a set of formulae) that entails  $\alpha$  (a formula). Here,  $\Phi$  is called the support, and  $\alpha$  is the claim, of the argument. Hence, starting with a set of formulae, arguments and counterarguments can be generated, where a counterargument (an argument that attacks another argument) either rebuts (i.e. negates the claim of the argument) or undercuts (i.e. negates the support of the argument). A range of options for structured argumentation at the logic level have been investigated (see [17,51,104,140] for tutorial reviews of some of the key proposals. Whilst most proposals for structured argumentation involve simple rule-based reasoning, there is some investigation of the role of classical logic in argumentation (see for example [16]), and of how probabilistic reasoning can be incorporated in structured arguments (see for example [74,139,144]). Patterns of reasoning that arise in persuasion can be captured in rule-based reasoning [134]. Furthermore, argument schemas have been proposed that capture common patterns of argument that can be formalized in structured arguments (for a review see [148]).

#### 4.1.3. Dialogical level

Dialogical argumentation involves agents exchanging arguments in activities such as discussion, debate, persuasion, and negotiation. Starting with [63,96], dialogue games are now a common approach to characterizing argumentation-based agent dialogues (e.g. [3,22,25,39,40,46,87,100,101,114,117,147]). Dialogue games are normally made up of a set of communicative acts called moves, and a protocol specifying which moves can be made at each step of the dialogue. Dialogical argumentation can be viewed as incorporating logic-based argumentation, but in addition, dialogical argumentation involves representing and managing the locutions exchanged between the agents involved in the argumentation. The emphasis of the dialogical view is on the interactions between the agents, and on the process of building up, and analyzing, the set of arguments until the agents reach a conclusion. See [118] for a review of formal models of persuasion dialogues and [23,136] for reviews and analyses of strategies in dialogical argumentation.

By modelling the persuadee, it is possible to update the model during the dialogue based on the persuadee responses. This can then be used to determine whether there is any chance of the dialogue leading to success, and if not, giving up and unsuccessfully terminating the dialogue [23]. Probabilistic models of the opponent have been used in some strategies allowing the selection of moves for an agent based on what it believes the other agent believes [73], selection of moves based on what it believes the other agent is aware of [126], and based on the history of previous dialogues to predict the arguments that an opponent might put forward [59]. In [20], a planning system is used by the persuader to optimize choice of arguments based on belief in premises, and in [21], an automated planning approach is used for persuasion that accounts for the uncertainty of the proponent's model of the opponent by finding strategies that have a certain probability of guaranteed success no matter which arguments the opponent chooses to assert.

Utility theory has also been considered in argumentation (for example [99,113,122,128]) though none of these represents the uncertainty of moves made by each agent in argumentation. Probability theory and utility theory (using decision theory) has been used in [81] to identify outcomes with maximum expected utility where outcomes are specified as particular arguments being included or excluded from extensions. Strategies in argumentation have also been analyzed using game theory [47,121,123], though these are more concerned with issues of manipulation, rather than persuasion.

#### 4.1.4. Rhetorical level

Normally argumentation is undertaken in some wider context of goals for the agents involved, and hence individual arguments are presented so as to contribute to the wider aim. For instance, if an agent is trying to persuade another agent to do something, then it is likely that some rhetorical device is harnessed and this will affect the nature of the arguments used (e.g. a politician may refer to *investing in the future of the nation's children* as a way of persuading colleagues to vote for an increase in taxation).

Aspects of the rhetorical level include believability of arguments from the perspective of the audience [69], impact of arguments from the perspective of the audience [70], use of threats and rewards [4], appropriateness of advocates [71], and values of the audience [11,12,14,112]. The latter has led to the notion of value-based argumentation, and this has been developed for various applications including group persuasion [112], for persuasion concerning plans [102], for analyzing legal reasoning [13], and for analyzing political argument [6].

The use of emotional arguments could be regarded as a rhetorical device. However, the modelling of emotional aspects of argument has received little attention in the computational argumentation literature. There is a proposal for rules for specifying scenarios where empathy is given or received in negotiation [98], and there is a proposal for specifying argument schemas (rules that specify general patterns of reasoning) for capturing aspects of emotional argument [94]. In contrast, it is interesting to note that affective computing has put emotion at the centre of the relationship between users and computing systems [26].

#### 4.2. Shortcomings in the state of the art

So computational models of argument offer a range of formal systems for generating and comparing arguments, and for undertaking this in a dialogue. However there are shortcomings in the state of the art of computational models of argument for application in persuasion. The current state of the literature does not adequately offer the following and hence there are some exciting research challenges to be addressed if we are to deliver computational persuasion.

**Domain knowledge** A formalization of domain knowledge appropriate for constructing arguments concerning persuasion in application such as behaviour change (e.g. a formalism for representing persuadee goals, persuadee preferences, system persuasion goals, and system knowledge concerning actions that can address persuadee goals, etc) though the multiagent communities offer proposals that might be adapted for our needs.

**Persuasion protocols** Since we are not attempting to support free text input from the persuadee, we require protocols that take account of the user's views without recourse to natural language processing. For this, in Section 6, I will discuss how asymmetric dialogues can be used.

**Persuadee models** Persuadee models that allow the persuasion system to construct a model of the persuadee's beliefs and preferences, to qualify the probabilistic uncertainty of that model, and to update that model and the associated uncertainty as the dialogue progresses. There are some promising proposals that could contribute to a solution (e.g. [20,21,59,73,75,127]), and I will discuss the progress we have made on this in Section 7.4. However, if we are to harness some of the other levers of persuasion that I discussed in Section 3, then we will need to broaden the modelling to incorporate aspects of personality and bias.

**Persuasion strategies** Strategies for persuasion that harness the persuadee model to find optimal moves to make at each stage (trading the increase in probability of successfully persuading the persuadee

against the raised risk that the persuadee disengages from the dialogue as it progresses). The strategies may involve the uncertainty in the user's beliefs and awareness of arguments, and it may also include an assessment of the user's personality and/or biases. With this kind of information in the user model, we may be able to harness some of the levers of persuasion discussed in Section 3 such a persuasion techniques, framing style, and argumentation style. I will discuss the progress we have made on this in Section 7.4.

In order to focus research on addressing these shortcomings, we can consider how computational persuasion can be developed and evaluated in the context of behaviour change applications.

#### 4.3. Studies with participants

In order to have well-understood computational models of argument that correspond to human behaviour, there is a need to ground these models with studies with participants. The studies undertaken so far validate some aspects of these models, but also indicate some shortcomings in being able to model human behaviour.

Studies performed by Rahwan *et al* [124] and Cerruti *et al* [32] investigated various forms of reinstatement in argumentation. The users were presented several argument graphs and were asked to explain how acceptable a given argument is in their opinion. The results show that in some cases, the implicit knowledge about domains can substantially affect the given acceptability levels. However, more importantly, the experiments show that the attacked argument's acceptability is lowered, but does not fall to 0, which is what would be predicted by the usual dialectical semantics for abstract argumentation. Additionally, introducing the defense for this argument raises its acceptability. However, typically it does not reach the value of 1, which is the level the usual dialectical semantics would predict.

In a study of argumentation dialogues, Rosenfeld and Kraus [129] undertook an experiment in order to develop a machine learning-based approach to predict the next move a participant would make in a dialogue. This work was further extended in [130,131]. The machine learning models were trained on data that incorporated the sequences of arguments in a dialogue that the participants accept. Once trained, the models were able to predict the acceptance an unseen case would have.

In another machine learning-based approach, Huang and Lin [68] developed a software agent for participating in dialogues with potential customers with the aim of persuading them to offer a higher price for goods. Dialogues were constructed from an argument graph, and training was done on simulated scenarios. In testing with users, the agent was able to persuade the participants to increase the mean price offer.

There are also studies with participants by Masthoff and co-workers that investigate the efficacy of using arguments as a way of persuading people when compared with other counselling methods indicating that argumentation may have disadvantages if used inappropriately [107], and that rather than a confrontational approach, argumentation that is based on appeal to friends, appeal to group, or appeal to fun, may be more efficacious [142,143].

Emotion in argumentation has also been the subject of a study with participants in a debate where the emotional state was estimated from EEG data and automated facial expression analysis. In this study, Benlamine *et al* [15] showed for instance that the number and the strength of arguments, attacks and supports exchanged between a participant could be correlated with particular emotions of the participant.

Table 1

Some examples where people could change their behaviour and for which there would be a substantial quantifiable benefit to themselves, and/or to society

Field	Examples of behaviour change topic
Healthy life-styles	eating fewer calories, eating more fruit and veg, doing more exercise, drinking less alcohol, brushing teeth properly, regular dental check-ups, participating in health screening, participating in vaccination programmes
Addiction management	gambling, smoking, drugs
Treatment compliance	self-management of diabetes, taking vaccines, completing course of antibiotics
Personal finance	borrowing less, saving more
Education	starting or continuing with a course, studying properly
Energy efficiency	reducing electricity consumption, installing home insulation
Citizenship	voting, recycling, contributing to charities, wasting less food
Safe driving	not exceeding speed limits, not texting while driving
Anti-social behaviour	aggression, vandalism, racism, sexism, trolling

## 5. What is behaviour change?

There is a wide variety of problems that are dangerous or unhealthy or unhelpful for an individual, or for those around him/her, and that are expensive to government and/or to society (see Table 1 for examples). For each type of problem, we can conceivably tackle a small proportion of cases with substantial benefit to individuals, government and society using techniques for behaviour change.

Many organizations are involved in behaviour change, and many approaches are used to persuade people to change their behaviour including counselling, information resources, and advertising. Many diverse factors can influence how such approaches can be used effectively in practice such as the following.

- **Perceived social norms** (e.g. everyone drives above the speed limit).
- **Social pressure** (e.g. my friends laugh at me if I drive slowly).
- **Emotional issues** (e.g. speeding is cool).
- **Agenda** (e.g. I am always late for everything, and so I have to speed).
- **Perception of an issue** (e.g. I am a good driver even if I speed).
- **Opportunities to change behaviour** (e.g. access to a race track on which to drive fast instead of driving fast on ordinary roads).
- **Attitude to persuader** (e.g. I listen to Lewis Hamilton not a civil servant).
- **Attitude to information** (e.g. I switch off if I am given statistics).

As computing becomes involved in every sphere of life, so too is persuasion a target for applying computer-based solutions. There are **persuasion technologies** that have come out of developments in human-computer interaction research (see for example the influential work by Fogg [49]) with a particular emphasis on addressing the need for systems to help people make positive changes to their behaviour, particularly in healthcare and healthy life-styles.

Over the past 10 years, a wide variety of systems have been developed to help users to control body weight [91], to reduce fizzy drink consumption [89], to increase physical exercise [149], and

to decrease stress-related illness [85]. Many of these persuasion technologies for behaviour change are based on some combination of questionnaires for finding out information from users, provision of information for directing the users to better behaviour, computer games to enable users to explore different scenarios concerning their behaviour, provision of diaries for getting users to record ongoing behaviour, and messages to remind the persuadee to continue with the better behaviour.

Interestingly, argumentation is not central to the current manifestations of persuasion technologies. The arguments for good behaviour seem either to be assumed before the persuadee accesses the persuasion technology (e.g. when using diaries, or receiving email reminders), or arguments are provided implicitly in the persuasion technology (e.g. through provision of information, or through game playing). So explicit consideration of arguments and counterarguments are not supported with existing persuasion technologies. This creates interesting opportunities for computational persuasion to develop APSs for behaviour change where arguments are central.

Argument-based persuasion technology could complement other technologies by helping users when they contemplate change. This fits the technology into the Stages of Change model [119] which comprises the following phases that someone might go through (examples taken from [110]).

- **Pre-contemplation** “I am happy being a smoker and intend to continue smoking”
- **Contemplation** “I have been coughing a lot recently, perhaps I should think about stopping smoking”
- **Preparation** “I will buy lower tar cigarettes”
- **Action** “I have stopped smoking”
- **Maintenance** “I have stopped smoking for four months now”

By acting at the contemplation stage, the user might be prepared to enter into a dialogue with an APS. The role of the APS would then be to provide context-specific (personalized) information to the user through arguments, and to handle the doubts and issues that the user might have in the form of counterarguments. In the next section, we consider the potential of this approach in more detail.

A strategy for an APS needs to find the best choice of move at each stage where best is determined in terms of some combination of the need to increase the likelihood that the persuadee is persuaded by the goal of the persuasion, and the need to decrease the likelihood that the persuadee disengages from the dialogue. For instance, at a certain point in the dialogue, the APS might have a choice of two arguments *A* and *B* to present. Suppose *A* involves further moves to be made (e.g. supporting arguments) whereas *B* is a single posit. So choosing *A* requires a longer dialogue (and higher probability of disengagement) than *B*. Also suppose that if the persuadee engages to the end of each dialogue, then it is more likely that the persuadee believes *A* than *B*. So if the APS is to make the best choice of move, it needs to consider both the risk and potential benefit from each of them.

An APS should present arguments and counterarguments that are informative, relevant, and believable, to the persuadee. If the APS presents uninformative, irrelevant, or unbelievable arguments (from the perspective of the persuadee), the probability of successful persuasion is reduced, and it may alienate the persuadee. A choice of strategy depends on the protocol, and on the kind of dynamic persuadee model. Various parameters can be considered in the strategy such as the preferences of the persuadee, the agenda of the persuadee, etc.

So argument-based persuasion for behaviour change offers a challenging and worthwhile field for developing and evaluating computational persuasion. As indicated by the review of computational models of argument in Section 4.1, there are some promising developments that could form the basis of APSs for behaviour change, as I discuss in the next section. Furthermore, there have already been some promising studies using dialogue games for health promotion [28,54–56], embodied conversational agents for encouraging exercise [108], dialogue management for persuasion [5], and tailored assistive living systems for encouraging exercise [57], that indicate the potential for APSs.

## 6. How can computational persuasion be applied?

Computational models of argument drawing on ideas of abstract argumentation, logical argumentation, dialogical argumentation, together with techniques for argument dynamics and for rhetorics, offer an excellent starting point for developing computational persuasion for applications in behaviour change.

I assume that an APS for behaviour change is a software application running on a website or mobile device. Some difficult challenges to automate persuasion via an app are the following.

1. **Need asymmetric dialogues without natural language interface:** Since we cannot assume that we will have natural language interfaces that can cope with the diversity of language that might arise in the arguments and counterarguments in a behaviour change application, we need to develop alternative ways of getting information from the user. This means that we will have asymmetric dialogues where the choice of moves available to the APS is different to the user. This approach was also taken in dialogue games for health promotion [28].
2. **Need short dialogues to keep engagement:** If an APS uses too many moves in a dialogue, there is a substantial risk that the user will disengage from the dialogue. They will become bored or frustrated by the interaction, or they will run out of time. It is therefore imperative that the APS does try to keep the dialogue short and focused.
3. **Need well-chosen arguments to maximize impact:** If an APS uses arguments that are inappropriate for a particular user, we are at risk of alienating the user, and thereby losing in the attempt to persuade the user.
4. **Need to model the user in order to be able to optimize the dialogue:** So the APS needs to determine at each stage of the dialogue what the best choice of move should be.
5. **Need to learn from previous interactions with the agent or similar agents:** If the APS is to have a useful model of a user, it needs to learn from previous interactions with the agent or similar agents.
6. **Need to model the domain to generate arguments/counterarguments:** If an APS is to undertake a dialogue, it needs access to a range of arguments that it can posit. Furthermore, if it is to anticipate the arguments the user might be entertaining at any stage of the dialogue, it needs to be aware of possible counterarguments.

The dialogue may involve steps where the system finds out more about the persuadee's beliefs, intentions and desires, and where the system offers arguments with the aim of changing the persuadee's beliefs, intentions and desires. The system also needs to handle objections or doubts (represented by counterarguments) with the aim of providing a dialectically winning position. To illustrate how a dialogue can lead to the presentation of an appropriate context-sensitive argument consider the example in Table 2. In this, only the APS presents arguments, and when it is the user's turn s/he can only answer



Table 2

Simple example of an asymmetric dialogue between a user and an APS. As no natural language processing is assumed, the arguments posted by the user are actually selected by the user from a menu provided by the APS

Step	Who	Move
1	APS	To improve your health, you could join an exercise class
2	User	Exercise classes are boring
3	APS	For exciting exercise, you could do an indoor climbing course
4	User	It is too expensive
5	APS	Do you work?
6	User	No
7	APS	If you are registered unemployed, then the local sports centre offers a free indoor climbing course
8	APS	Would you try this?
9	User	Yes

Since you do little exercise, you should do a regular exercise class

When I do exercise, I get very hungry and I put on weight
 

Strongly agree

Agree

Neither agree nor disagree

Disagree

Strongly disagree

How often do you have face-to-face contact with patients?

Less than once per week

Between 1 and 5 per week

Between 6 and 30 per week

Between 30 and 100 per week

More than 100 per week

Which are goals for you?

Staying healthy

Living a long life

Doing interesting activities

Relaxing

Doing as little as possible

Being with the family

Being healthy for the family

Looking good

Being fit for my sport

Saving time

(a) Asking for argument belief.
(b) Asking for a fact.
(c) Asking for goals.

Fig. 2. Interface for an asymmetric dialogue move for asking the user's belief in an argument. (a) The top argument is by the APS, and the second argument is a counterargument presented by the APS. The user uses the menu to give his/her belief in the counterargument. (b) A query is asked that may be used in a user model and menu of answers is provided. (c) A query is asked by the system to determine the goals of the user. Here the user may select any number of the items on the list.

questions (e.g. yes/no questions) or select arguments from a menu. In Fig. 2, a dialogue step is illustrated where a user can state the degree of agreement or disagreement in an argument.

Arguments can be automatically generated from a knowledgebase. For this, we can build a knowledgebase for each domain, though there are many commonalities in the knowledge required for each

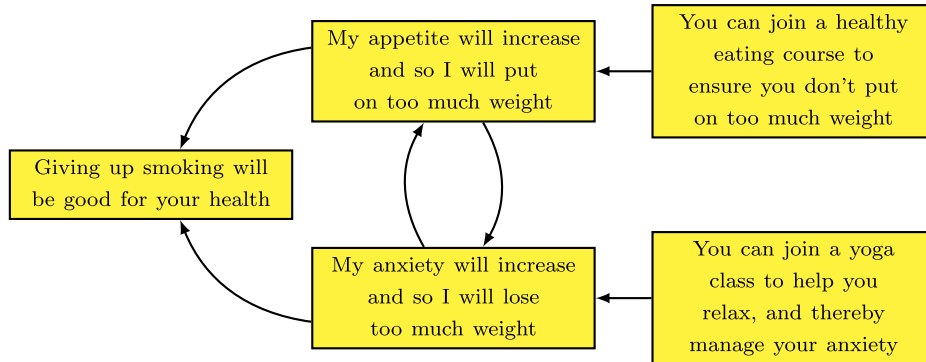


Fig. 3. Example of an argument graph for persuasion.

behaviour change application.

- **Persuadee beliefs** (e.g. cakes give a sugar rush).
- **Persuadee preferences** (e.g. burgers are preferred to apples).
- **Behavioural states** (e.g. persuadee's weight, exercise regime, etc.).
- **Behavioural actions** (e.g. eat a piece of fruit, eat a piece of cake, walk 1 km).
- **Behavioural goals** (e.g. lose 10 kg by Christmas, reduce sugar intake).

To represent and reason with the domain knowledge, we could harness a form of Belief-Desire-Intention (BDI) calculus in predicate logic for relating beliefs, behavioural goals, and behavioural states, to possible actions. We could then use the calculus with logical argumentation to generate arguments for persuasion. A small example of an argument graph that we might want to generate by this process is given in Fig. 3 including the persuasion goal *giving up smoking will be good for your health*.

To support the selection of arguments, we require persuadee models. For this, we can establish the probabilistic uncertainty associated with the APS model of the persuadee's beliefs, behavioural state, behavioural goals, preferences, and tendencies etc by asking the persuadee appropriate questions, by considering previous usage of the APS by the persuadee, and by the type of the persuadee (i.e. by assignment to a built-in model learned from a class of similar users).

## 7. Towards a framework for computational persuasion

In this section, I outline a framework for computational persuasion that is being developed in an ongoing project (for more information, see the project website<sup>2</sup>).

### 7.1. A brief review of probabilistic argumentation

In our framework, we focus on the uncertainty surrounding the user's awareness of arguments, and the user's belief in the arguments s/he is aware of. For this, we have harnessed probabilistic argumentation. Two main approaches to probabilistic argumentation are the constellations and the epistemic approaches [74].

<sup>2</sup>[www.computationalpersuasion.com](http://www.computationalpersuasion.com)

- In the constellations approach, the uncertainty is in the topology of the graph (see for example [43,72,92]). As an example, this approach is useful when one agent is not sure what arguments and attacks another agent is aware of, and so this can be captured by a probability distribution over the space of possible argument graphs. The usual definition for extensions (grounded, preferred, stable, etc) can be applied to each subgraph, and then for each subset of arguments  $X$ , the probability that  $X$  is an extension for the grounded (respectively preferred, stable, etc) extension is the sum of the probability of each subgraph that has  $X$  as a grounded (respectively preferred, stable, etc) extension.
- In the epistemic approach, the topology of the argument graph is fixed, but there is uncertainty about whether an argument is believed [74,82,84,135]. This is formalized by a probability distribution over the subsets of the set of arguments in the graph. In addition, postulates have been proposed to capture intuitive constraints such as the *rational postulate* which states that if an attacker has a probability greater than 0.5 (i.e. it is believed), then any attackee has a belief less than or equal to 0.5 (i.e. it is not believed). The epistemic approach can give a finer grained version of Dung's approach, and it can be used to give a valuable alternative to Dung's approach. For example, for a graph containing arguments  $A$  and  $B$  where  $B$  attacks  $A$ , it might be the case that a user believes  $A$  and not  $B$ , and if so the epistemic extension (the set of believed arguments) would be  $\{A\}$  which is in contrast the Dung's approach where the only extension is  $\{B\}$ .

The epistemic approach has been extended with a probability distribution over subsets of the set of attacks [116]. This can be used to represent an agent's belief in each attack. This is potentially useful for handling enthymemes. Since most arguments are presented in natural language, different agents may interpret them differently, and hence some agents may be belief that an attack holds between a pair of arguments whereas other agents might not.

## 7.2. Studies with participants

We have undertaken studies with participants to evaluate how they deal with arguments arising in a dialogue [115]. We asked each participant for their belief in the arguments at each stage of the dialogue, and whether they saw a negative (i.e. attacking) or positive (i.e. supporting) relationship between the latest argument added in the dialogue and the previous arguments. For this study, we were able to make a number of observations including the following.

- **Use of constellations approach.** People may interpret statements and relations between them differently, and not necessarily in the intended manner, and hence a probability distribution over the possible subgraphs can capture this uncertainty. Furthermore, people may explicitly declare that two statements are connected, however, they might not be sure of the exact nature of the relation between them. We therefore need to express the uncertainty that a person has about his own views, and this can be addressed with the constellations approach.
- **Use of epistemic approach.** People may assign levels of agreement to statements going beyond the 3-valued approach of Dung; The epistemic postulates (e.g. if the belief in an attacking argument is greater than 0.5, then the belief in the attackee is less than 0.5), in contrast to the standard semantics, can be highly adhered to; And the extended epistemic postulates (which also model uncertainty in attacks) allow us to model situations where the perceived "strength" of a relation might not necessarily be tightly related to the level of agreement assigned to its source argument. These points suggest the need for the epistemic approach to probabilistic argumentation.

- **Use of bipolar argumentation.** The notion of defense (i.e. when  $A$  attacks  $B$  and  $B$  attacks  $C$ , then  $A$  defends  $C$ ) does not account for all of the positive relations between the statements viewed by the participants. This suggests the need for bipolar argumentation frameworks.

We have also developed methods for acquiring crowd-sourced opinions on arguments, and shown how they can be used for predicting opinions on arguments [79]. We evaluated our approach by crowd-sourcing opinions from 50 participants about 30 arguments. This work shows how it is viable to acquire data from a number of contributors to construct classifiers, and that these classifiers can then be deployed to substantially decrease the number of questions that need to be asked of any particular user in a persuasion dialogue.

### 7.3. Strategic argumentation

Given the potential for probabilistic argumentation to capture key aspects of uncertainty in the user model, I indicate below how strategic argumentation can be developed to harness the user model.

- **Awareness of arguments** For considering the uncertainty about the structure of the graph in the persuadee mind, we use the constellations approach. We can update the model with each argument/attack presented, and we can use expected utility to identify best choice of argument/attack to present [81,83].
- **Belief in arguments** For considering the uncertainty about the beliefs of the persuadee, we use the epistemic approach. The epistemic approach is useful for asymmetric dialogues where the user is not allowed to posit arguments or counterarguments [76]. So the only way the user can treat arguments that s/he does not accept is by disbelieving them. In contrast, in symmetric dialogues, the user could be allowed to posit counterarguments to an argument that s/he does not accept. The distribution can be updated in response to moves made (posits, answers to queries, etc) using different assumptions about the persuadee (credulous, skeptical, rational, etc). The aim is to choose moves that will increase belief in positive persuasion goals or decrease belief in negative persuasion goals. We have proposed methods for updating beliefs during a dialogue [76,77,80], for efficient representation and reasoning with the probabilistic user model [61], for representing uncertainty in the user model [78], and for harnessing decision-theoretic decision rules for optimizing the choice of arguments based on the user model [62].
- **Moves in a dialogue** For considering the possible dialogues that might be generated by a pair of agents, a probabilistic finite state machine can represent the possible moves that each agent can make in each state of the dialogue assuming a set of arguments that each agent is aware of [75]. Each state is composed of the public state of the dialogue (e.g. what has been said) and the private state of each participant (e.g. the arguments they believe). We can find optimal sequences of moves by handling uncertainty concerning the persuadee using partially observable Markov decision processes (POMDPs) when there is uncertainty about the private state of the persuader [60].
- **Disengagement** For considering disengagement, we have investigated two options. The first is a simple Markov model that increases the probability of disengagement with each step of the dialogue [78]. The second is to use a weighting factor on the quality of a dialogue that is taken into account when we applying decision rules to construct an optimal policy [62].

Key possible dimensions for modelling uncertainty are summarized in Table 3. These developments offer a framework with a well-understood theoretical methodology, and implementations that are computationally viable for strategic argumentation.

Table 3  
Possible dimensions of uncertainty in models of persuadee

Type of uncertainty	Modelling technique
Beliefs of persuadee	Epistemic approach
Arguments/attacks known by persuadee	Constellations approach
Moves that persuadee makes	PFSMs/POMDPs
Risk of disengagement	Markov models/Discount factors

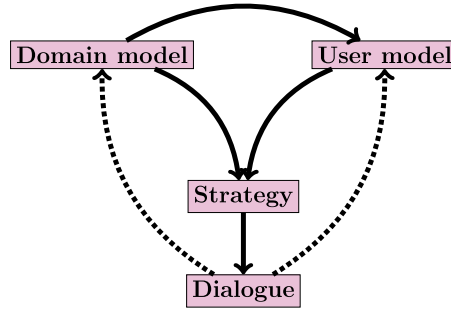


Fig. 4. Key aspects of our framework for computational persuasion. A solid arrow indicates a necessary flow of information whereas a dotted arrow indicates an optional flow of information.

#### 7.4. An integrated theory for computational persuasion

The key research issues for our project are summarized in Fig. 4, and described below. The aim is to address these issues in order to provide an integrated theory for applications in behaviour change.

- **Domain model** The content of the dialogues come from the domain model, and given that these are argumentation dialogues, the domain model needs to have the capacity to provide appropriate arguments and counterarguments as required. In the simplest case, the domain model may be based on an argument graph, perhaps with meta-level information such as typing of arguments. A more sophisticated domain model might be based on a knowledgebase of logical formulae that can be used to construct arguments and counterarguments.
- **User model** The user model incorporates information about which arguments in the domain model that the user is aware of and information about the belief the user has in these arguments. So information in the domain model is drawn on by the user model, and the user model uses the constellations and epistemic approaches to probabilistic argumentation to model these aspects. A user model can be acquired or predicted at the start of the dialogue, and it may be updated during the dialogue, and there are a number of options for this.
  - **Interrogative** The user is asked questions, and so the answers from the user are used for constructing or updating the user model. This may give accurate information but it may be boring to the user to be asked too many questions.
  - **Observational** The user's contributions to the dialogue imply information that can be used for constructing or updating the user model (e.g. if the user selects an argument *A* from a menu, then we might assume that the user believes *A*, and hence we can update the user model with this information).

- **Historical** The user’s answers and contributions to previous dialogues imply information that can be used for constructing or updating the user model. (e.g. if the user has previously expressed belief in a specific argument).
- **External** Other sources (e.g. social media) can be used for acquiring information about the user and/or predicting information about the user (see for example [88]).
- **Population** Data from similar users (e.g. crowd-sourced data) can be used to predict information about the user. For instance, we have developed methods for acquiring beliefs in arguments, and shown how naive Bayes classifiers can be trained to predict the belief in arguments for a given user [115].
- **Strategy and dialogue** There are options for the strategy as discussed in Section 7.3. The strategy draws on the domain model for arguments to use as content in the dialogue model, and it draws on the user model to determine the best choice of move at each stage of the dialogue. The net result of the strategy is the system contribution to the dialogue. Optionally, the moves by the user in the dialogue can inform the user model as suggested above, and if the user is able to present his/her arguments in the dialogue, these may inform the domain model.

We have put uncertainty in arguments, in particular in belief in arguments, at the core of our framework for computational persuasion, as we believe this is a minimum necessary for persuasive behaviour. However, we believe that other dimensions are highly desirable for a more comprehensive framework for computational persuasion. In particular, some of the dimensions considered in Section 3 are potentially valuable including rationality of arguments (in particular quality of arguments and of argumentation), persuasion techniques, argumentation style, framing of arguments, and emotion of arguments.

## 8. Discussion

Computational persuasion, being based on computational models of argument, is a promising approach to technology for behaviour change applications. Advantages of dialogical persuasion over uni-directional persuasion for behaviour change include:

- **Personalization** A dialogue can be tailored to the needs of the user.
- **Context-sensitivity** A dialogue can take into account the context of the user.
- **Interactivity** A user can provide input to help guide the content and approach to the dialogue, and thereby making them more engaging and useful for the user.

Developing an automated persuasion system (APS) involves research challenges including: undertaking the dialogue without using natural language processing; having an appropriate model of the domain in order to identify arguments; having an appropriate dynamic model of the persuadee; and having a strategy that increases the probability of persuading the persuadee. Furthermore, with even a modest set of arguments, the set of possible dialogues can be enormous, and so the protocols, persuadee models, and strategies need to be computationally viable.

Persuasion can be described as a process for overcoming barriers to behaviour change. There are many kinds of barriers. A simple dichotomy is that of informational barrier and psychological barrier. An **informational barrier** concerns the information that the user may have and/or lack. So a user may hold incorrect information (e.g. soft drinks are healthy because they are made from vegetable extracts) or lack necessary information (e.g. someone might have an unhealthy diet because they have never learned about what a healthy diet is and why it is important). Such informational barriers to behaviour



change are the primary focus of our approach to computational persuasion so far. Through dialogical argumentation, an APS can find out more about the user's beliefs and priorities, and then use that to present arguments tailored to that user, and to counter any incorrect beliefs that the user may have. A second kind of barrier could be described as a **psychological barrier** which concerns psychological reasons for why someone has a particular behaviour and/or finds it difficult to see that they should change behaviour [111]. This may include debilitating emotions, dysfunctional attitudes, and perceived invulnerabilities, and therefore addressing psychological barriers calls for more sophisticated approaches to the development computational persuasion.

In the short-term, we may envisage that the dialogues between an APS and a user involve limited kinds of interaction. For example, the APS manages the dialogue by asking queries of the persuadee, where the allowed answers are given by a menu or are of restricted types (e.g. age), and by positing arguments, and the persuadee may present arguments that are selected from a menu presented by the APS. Obviously richer natural language interaction would be desirable, but it is not feasible in the short-term. Even with such restricted asymmetric dialogues, it may be possible that effective persuasion can be undertaken, and furthermore, we need to investigate this conjecture empirically with participants. In the longer-term, there are likely to be exciting opportunities for combining computational models of argument with computational linguistics for much more involved and convincing dialogical argumentation for persuasion.

## Acknowledgements

I am grateful to the anonymous reviewers for helpful feedback for improving the paper. This research is part-funded by EPSRC grant EP/N008294/1 *Framework for Computational Persuasion*.

## References

- [1] L. Amgoud and J. Ben-Naim, Ranking-based semantics for argumentation frameworks, in: *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'13)*, LNAI, Vol. 8078, Springer, 2013.
- [2] L. Amgoud and C. Cayrol, Inferring from inconsistency in preference-based argumentation frameworks, *Journal of Automated Reasoning* **29** (2002), 125–169. doi:[10.1023/A:1021603608656](https://doi.org/10.1023/A:1021603608656).
- [3] L. Amgoud, N. Maudet and S. Parsons, Arguments, dialogue and negotiation, in: *Proceedings of the European Conference on Artificial Intelligence (ECAI'00)*, IOS Press, 2000, pp. 338–342.
- [4] L. Amgoud and H. Prade, Formal handling of threats and rewards in a negotiation dialogue, in: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS'05)*, IFAAMAS, 2005, pp. 529–536.
- [5] P. Andrews, S. Manandhar and M. De Boni, Argumentative human computer dialogue for automated persuasion, in: *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*, ACL, 2008, pp. 138–147.
- [6] K. Atkinson, T. Bench-Capon and P. McBurney, Persuasive political argument, in: *Proceedings of the International Workshop on Computational Models of Natural Argument (CMNA 2005)*, IJCAI, 2005, pp. 44–51.
- [7] P. Baroni, M. Caminada and M. Giacomin, An introduction to argumentation semantics, *Knowledge Engineering Review* **26** (2011), 365–410. doi:[10.1017/S0269888911000166](https://doi.org/10.1017/S0269888911000166).
- [8] P. Baroni, F. Cerutti, M. Giacomin and G. Guida, AFRA: Argumentation framework with recursive attacks, *International Journal of Approximate Reasoning* **52** (2011).
- [9] R. Baumann, What does it take to enforce an argument? Minimal change in abstract argumentation, in: *Proceedings of the European Conference on Artificial Intelligence (ECAI'12)*, IOS Press, 2012, pp. 127–132.
- [10] R. Baumann and G. Brewka, Expanding argumentation frameworks: Enforcing and monotonicity results, in: *Computational Models of Argument (COMMA'10)*, IOS Press, 2010, pp. 75–86.
- [11] T. Bench-Capon, Persuasion in practical argument using value based argumentation frameworks, *Journal of Logic and Computation* **13**(3) (2003), 429–448. doi:[10.1093/logcom/13.3.429](https://doi.org/10.1093/logcom/13.3.429).

- [12] T. Bench-Capon, Open texture and argumentation: What makes an argument persuasive? in: *Sergot Festschrift*, LNAI, Vol. 7360, Springer, 2012, pp. 220–233.
- [13] T. Bench-Capon, K. Atkinson and A. Chorley, Persuasion and value in legal argument, *Journal of Logic and Computation*, **15** (2005), 1075–1097. doi:[10.1093/logcom/exi058](https://doi.org/10.1093/logcom/exi058).
- [14] T. Bench-Capon, S. Doutre and P. Dunne, Audiences in argumentation frameworks, *Artificial Intelligence* **171**(1) (2007), 42–71. doi:[10.1016/j.artint.2006.10.013](https://doi.org/10.1016/j.artint.2006.10.013).
- [15] S. Benlamine, M. Chaouachi, S. Villata, E. Cabrio, C. Frasson and F. Gandon, Emotions in argumentation: An empirical evaluation, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'15)*, IJCAI, 2015, pp. 156–163.
- [16] P. Besnard and A. Hunter, *Elements of Argumentation*, MIT Press, 2008. doi:[10.7551/mitpress/9780262026437.001.0001](https://doi.org/10.7551/mitpress/9780262026437.001.0001).
- [17] P. Besnard and A. Hunter, Constructing argument graphs with deductive arguments: A tutorial, *Argument and Computation* **5**(1) (2014), 5–30. doi:[10.1080/19462166.2013.869765](https://doi.org/10.1080/19462166.2013.869765).
- [18] P. Bisquert, C. Cayrol, F.D. de Saint-Cyr and M. Lagasquie-Schiex, Enforcement in argumentation is a kind of update, in: *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'13)*, LNCS, Vol. 8078, Springer, 2013, pp. 30–42.
- [19] P. Bisquert, M. Croitoru and F. Dupin de Saint-Cyr, Four ways to evaluate arguments according to agent engagement, in: *Proceedings of International Conference on Brain Informatics and Health (BIH'15)*, LNCS, Vol. 9250, Springer, 2015.
- [20] E. Black, A. Coles and S. Bernardini, Automated planning of simple persuasion dialogues, in: *Proceedings of the International Workshop on Computational Logic in Multi-Agent Systems (CLIMA'14)*, LNCS, Vol. 8624, Springer, 2014, pp. 87–104.
- [21] E. Black, A. Coles and C. Hampson, Planning for persuasion, in: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS'17)*, ACM, 2017.
- [22] E. Black and A. Hunter, An inquiry dialogue system, *Autonomous Agents and Multi-Agent Systems* **19**(2) (2009), 173–209. doi:[10.1007/s10458-008-9074-5](https://doi.org/10.1007/s10458-008-9074-5).
- [23] E. Black and A. Hunter, Reasons and options for updating an opponent model in persuasion dialogues, in: *Theory and Applications of Formal Argumentation*, Vol. 9524, Springer, 2016, pp. 21–39.
- [24] E. Bonzon, J. Delobelle, S. Konieczny and N. Maudet, A comparative study of ranking-based semantics for abstract argumentation, in: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'16)*, AAAI Press, 2016.
- [25] E. Bonzon and N. Maudet, On the outcomes of multiparty persuasion, in: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS'11)*, IFAAMAS, 2011, pp. 47–54.
- [26] R. Calvo and S. D'Mello, Affect detection: An interdisciplinary review of models, methods, and their applications, *IEEE Transactions on Affective Computing* **1**(1) (2010), 18–37. doi:[10.1109/T-AFFC.2010.1](https://doi.org/10.1109/T-AFFC.2010.1).
- [27] M. Caminada and D. Gabbay, A logical account of formal argumentation, *Studia Logica* **93**(2–3) (2009), 109–145. doi:[10.1007/s11225-009-9218-x](https://doi.org/10.1007/s11225-009-9218-x).
- [28] A. Cawsey, F. Grasso and R. Jones, A conversational model for health promotion on the world wide web, in: *Proceedings of Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM'99)*, LNAI, Vol. 1620, Springer, 1999, pp. 379–388.
- [29] C. Cayrol, F.D. de Saint-Cyr and M.-C. Lagasquie-Schiex, Change in abstract argumentation frameworks: Adding an argument, *Journal of Artificial Intelligence Research* **38** (2010), 49–84.
- [30] C. Cayrol and M. Lagasquie-Schiex, On the acceptability of arguments in bipolar argumentation frameworks., in: *Proceedings of the 8th Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'05)*, LNCS, Vol. 3571, Springer, 2005, pp. 378–389.
- [31] C. Cayrol and M. Lagasquie-Schiex, Bipolarity in argumentation graphs: Towards a better understanding, *International Journal of Approximate Reasoning* **54**(7) (2013), 876–899. doi:[10.1016/j.ijar.2013.03.001](https://doi.org/10.1016/j.ijar.2013.03.001).
- [32] F. Cerutti, N. Tintarev and N. Oren, Formal arguments, preferences, and natural language interfaces to humans: An empirical evaluation, in: *Proceedings of the European Conference on Artificial Intelligence (ECAI'14)*, IOS Press, 2014, pp. 207–212.
- [33] G. Charwat, W. Dvorak, S. Gaggl, J. Wallner and S. Woltran, Reasoning problems in abstract argumentation – A survey, *Artificial Intelligence* **220** (2015), 28–63. doi:[10.1016/j.artint.2014.11.008](https://doi.org/10.1016/j.artint.2014.11.008).
- [34] R. Cialdini, *Influence: The Psychology of Persuasion*, HarperCollins, 1984.
- [35] R. Cockcroft and S. Cockcroft, *Persuading People*, Macmillan, 1992. doi:[10.1007/978-1-349-22254-4](https://doi.org/10.1007/978-1-349-22254-4).
- [36] S. Coste-Marquis, S. Konieczny and J.-G. Maily, On the revision of argumentation systems: Minimal change of argument statuses, in: *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR'14)*, AAAI Press, 2014, pp. 72–81.
- [37] S. Coste-Marquis, S. Konieczny and J.-G. Maily, A translation-based approach for revision of argumentation frameworks, in: *Proceedings of the European Conference on Logics in Artificial Intelligence (JELIA'14)*, LNCS, Vol. 8761, Springer, 2014, pp. 77–85.

- [38] S. Coste-Marquis, S. Konieczny and J.-G. Maily, Extension enforcement in abstract argumentation as an optimization problem, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'15)*, IJCAI, 2015, pp. 2876–2882.
- [39] J. Devereux and C. Reed, Strategic argumentation in rigorous persuasion dialogue, in: *Proceedings of International Workshop on Argumentation in Multi-Agent Systems (ArgMAS'09)*, LNCS, Vol. 6057, Springer, 2009, pp. 94–113.
- [40] F. Dignum, B. Dunin-Keplicz and R. Verbrugge, Dialogue in team formation, in: *Issues in Agent Communication*, Springer, 2000, pp. 264–280.
- [41] M. Diller, A. Haret, T. Linsbichler, S. Rümmele and S. Woltran, An extension-based approach to belief revision in abstract argumentation, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'15)*, IJCAI, 2015, pp. 2926–2932.
- [42] P. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games, *Artificial Intelligence* **77** (1995), 321–357. doi:[10.1016/0004-3702\(94\)00041-X](https://doi.org/10.1016/0004-3702(94)00041-X).
- [43] P. Dung and P. Thang, Towards (probabilistic) argumentation for jury-based dispute resolution, in: *Computational Models of Argument (COMMA'10)*, IOS Press, 2010, pp. 171–182.
- [44] P. Dunne, A. Hunter, P. McBurney, S. Parsons and M. Wooldridge, Weighted argument systems: Basic definitions, algorithms, and complexity results, *Artificial Intelligence* **175**(2) (2011), 457–486. doi:[10.1016/j.artint.2010.09.005](https://doi.org/10.1016/j.artint.2010.09.005).
- [45] A. Eagly, Comprehensibility of persuasive arguments as a determinant of opinion change, *Journal of Personality and Social Psychology* **29**(6) (1974), 758–773. doi:[10.1037/h0036202](https://doi.org/10.1037/h0036202).
- [46] X. Fan and F. Toni, Assumption-based argumentation dialogues, in: *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'11)*, IJCAI, 2011, pp. 198–203.
- [47] X. Fan and F. Toni, Mechanism design for argumentation-based persuasion, in: *Proceedings of Computational Models of Argument (COMMA'12)*, IOS Press, 2012, pp. 322–333.
- [48] A. Fisher, *The Logic of Real Arguments*, Cambridge University Press, 1988.
- [49] B. Fogg, Persuasive computers, in: *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, CHI, 1998, pp. 225–232.
- [50] D. Gabbay and O. Rodrigues, A numerical approach to the merging of argumentation networks, in: *Proceedings of International Workshop on Computational Logic in Multi-Agent Systems (CLIMA'12)*, LNCS, Vol. 7486, Springer, 2012, pp. 195–212.
- [51] A. Garcia and G. Simari, Defeasible logic programming: DeLP-servers, contextual queries, and explanations for answers, *Argument and Computation* **5**(1) (2014), 63–88. doi:[10.1080/19462166.2013.869767](https://doi.org/10.1080/19462166.2013.869767).
- [52] V. Gold, Covert advocacy: Reflections on the use of psychological persuasion techniques in the courtroom, *North Carolina Law Review* **65**(3) (1987), 481–516.
- [53] L. Goldberg, The structure of phenotype personality traits, *American Psychologist* **48**(1) (1993), 26–34. doi:[10.1037/0003-066X.48.1.26](https://doi.org/10.1037/0003-066X.48.1.26).
- [54] F. Grasso, Exciting avocados and dull pears – combining behavioural and argumentative theory for producing effective advice, in: *Proceedings of the Annual Meeting of the Cognitive Science Society*, Cognitive Science Society, 1998, pp. 436–441.
- [55] F. Grasso, Rhetorical coding of health promotion dialogues, in: *Proceedings of the Conference on Artificial Intelligence in Medicine (AIME'03)*, LNCS, Vol. 2780, Springer, 2003, pp. 179–188.
- [56] F. Grasso, A. Cawsey and R. Jones, Dialectical argumentation to solve conflicts in advice giving: A case study in the promotion of healthy nutrition, *International Journal of Human-Computer Studies* **53**(6) (2000), 1077–1115. doi:[10.1006/ijhc.2000.0429](https://doi.org/10.1006/ijhc.2000.0429).
- [57] E. Guerrero, J. Nieves and H. Lindgren, An activity-centric argumentation framework for assistive technology aimed at improving health, *Argument and Computation* **7** (2016), 5–33.
- [58] R. Gula, *Nonsense: Red Herrings, Straw Men, and Sacred Cows: How We Abuse Logic in Our Everyday Language*, Axios Press, 2006.
- [59] C. Hadjinikolis, Y. Siantos, S. Modgil, E. Black and P. McBurney, Opponent modelling in persuasion dialogues, in: *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'13)*, IJCAI, 2013, pp. 164–170.
- [60] E. Hadoux, A. Beynier, N. Maudet, P. Weng and A. Hunter, Optimization of probabilistic argumentation with Markov decision models, in: *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'15)*, IJCAI, 2015, pp. 2004–2010.
- [61] E. Hadoux and A. Hunter, Computationally viable handling of beliefs in arguments for persuasion, in: *Proceedings of the International Conference on Tools with AI (ICTAI'16)*, IEEE Press, 2016, pp. 319–326.
- [62] E. Hadoux and A. Hunter, Strategic sequences of arguments for persuasion using decision trees, in: *Proceedings of AAAI Conference on Artificial Intelligence (AAAI'17)*, AAAI Press, 2017, pp. 1128–1134.
- [63] C. Hamblin, Mathematical models of dialogue, *Theoria* **37** (1971), 567–583.
- [64] H. Hoeken, R. Timmers and P. Schellens, Arguing about desirable consequences: What constitutes a convincing argument?, *Thinking and Reasoning* **18**(3) (2012), 394–416. doi:[10.1080/13546783.2012.669986](https://doi.org/10.1080/13546783.2012.669986).

- [65] G. Hofstede, *Culture's Consequences: International Differences in Work-Related Values*, 2nd edn, Sage, 1984.
- [66] T. Hollihan and K. Basske, *Arguments and Arguing: The Products and Process of Human Decision Making*, Waveland Press, 2005.
- [67] J. Hornikx and U. Hahn, Reasoning and argumentation: Towards an integrated psychology of argumentation, *Thinking and Reasoning* **18**(3) (2012), 225–243. doi:[10.1080/13546783.2012.674715](https://doi.org/10.1080/13546783.2012.674715).
- [68] S. Huang and F. Lin, The design and evaluation of an intelligent sales agent for online persuasion and negotiation, *Electronic Commerce Research and Applications* (2007), 285–296.
- [69] A. Hunter, Making argumentation more believable, in: *Proceedings of the National Conference on Artificial Intelligence (AAAI'04)*, AAAI Press, 2004, pp. 269–274.
- [70] A. Hunter, Towards higher impact argumentation, in: *Proceedings of the National Conference on Artificial Intelligence (AAAI'04)*, AAAI Press, 2004, pp. 275–280.
- [71] A. Hunter, Reasoning about the appropriateness of proponents for arguments, in: *Proceedings of AAAI Conference on Artificial Intelligence (AAAI'08)*, AAAI Press, 2008, pp. 89–94.
- [72] A. Hunter, Some foundations for probabilistic abstract argumentation, in: *Computational Models of Argument (COMMA'12)*, IOS Press, 2012, pp. 117–128.
- [73] A. Hunter, Modelling uncertainty in persuasion, in: *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'13)*, LNCS, Vol. 8078, Springer, 2013, pp. 57–70.
- [74] A. Hunter, A probabilistic approach to modelling uncertain logical arguments, *International Journal of Approximate Reasoning* **54**(1) (2013), 47–81. doi:[10.1016/j.ijar.2012.08.003](https://doi.org/10.1016/j.ijar.2012.08.003).
- [75] A. Hunter, Probabilistic strategies in dialogical argumentation, in: *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'14)*, LNCS, Vol. 8720, Springer, 2014, pp. 190–202.
- [76] A. Hunter, Modelling the persuadee in asymmetric argumentation dialogues for persuasion, in: *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'15)*, AAAI Press, 2015, pp. 3055–3061.
- [77] A. Hunter, Persuasion dialogues via restricted interfaces using probabilistic argumentation, in: *Proceedings of the International Workshop on Scalable Uncertainty Models (SUM'16)*, LNCS, Vol. 9858, Springer, 2016, pp. 184–198.
- [78] A. Hunter, Two dimensional uncertainty in persuadee modelling in argumentation, in: *Proceedings of the European Conference on Artificial Intelligence (ECAI'16)*, IOS Press, 2016, pp. 150–157.
- [79] A. Hunter and S. Polberg, Empirical methods for modelling persuadees in dialogical argumentation, in: *Proceedings of the International Conference on Tools with Artificial Intelligence (ICTAI'17)*, IEEE Press, 2017, in press.
- [80] A. Hunter and N. Potyka, Updating probabilistic epistemic states in persuasion dialogues, in: *Proceedings of European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, LNCS, Vol. 10369, Springer, 2017, pp. 46–56.
- [81] A. Hunter and M. Thimm, Probabilistic argument graphs for argumentation lotteries, in: *Computational Models of Argument (COMMA'14)*, IOS Press, 2014, pp. 313–324.
- [82] A. Hunter and M. Thimm, Probabilistic argumentation with incomplete information, in: *Proceedings of the European Conference on Artificial Intelligence (ECAI'14)*, IOS Press, 2014, pp. 1033–1034.
- [83] A. Hunter and M. Thimm, Optimization of dialectical outcomes in dialogical argumentation, *International Journal of Approximate Reasoning* **78** (2016), 73–102. doi:[10.1016/j.ijar.2016.06.014](https://doi.org/10.1016/j.ijar.2016.06.014).
- [84] A. Hunter and M. Thimm, Probabilistic reasoning with abstract argumentation frameworks, *Journal of Artificial Intelligence Research* **59** (2017), 565–611.
- [85] K. Kaipainen, E. Mattila, M. Kinnunen and I. Korhonen, Facilitation of goal-setting and follow-up in Internet intervention for health and wellness, in: *Persuasion Technology 2010*, LNCS, Vol. 7822, Springer, 2010, pp. 238–249.
- [86] H. Kelman and A. Eagly, Attitude toward the communicator, perception of communication content, and attitude change, *Journal of Personality and Social Psychology* **1**(1) (1965), 63–78. doi:[10.1037/h0021645](https://doi.org/10.1037/h0021645).
- [87] M. Koit, Influencing the beliefs of a dialogue partner, in: *Proceedings of AIMSA'16*, LNCS, Vol. 9883, Springer, 2016, pp. 216–225.
- [88] M. Kosinski, D. Stillwell and T. Graepel, Private traits and attributes are predictable from digital records of human behavior, *Proceedings of the National Academy of Sciences* **110**(15) (2013), 5802–5805. doi:[10.1073/pnas.1218772110](https://doi.org/10.1073/pnas.1218772110).
- [89] S. Langrid and H. Oinas-Kukkonen, Less fizzy drinks: A multi-method study of persuasive reminders, in: *Persuasive Technology 2012*, LNCS, Vol. 7284, Springer, 2012, pp. 256–261.
- [90] R. Lehrman, *The Political Speechwriter's Companion: A Guide for Writers and Speakers*, CQ Press, 2009.
- [91] T. Lehto and H. Oinas-Kukkonen, Persuasive features in six weight loss websites: A qualitative evaluation, in: *Persuasive Technology 2010*, LNCS, Vol. 6137, Springer, 2010, pp. 162–173.
- [92] H. Li, N. Oren and T.J. Norman, Probabilistic argumentation frameworks, in: *Proceedings of Theory and Applications of Formal Argumentation (TAFIA'11)*, LNCS, Vol. 7132, Springer, 2011, pp. 1–16.
- [93] M. Lippi and P. Torroni, Argumentation mining: State of the art and emerging trends, *ACM Transactions on Internet Technology* **16**(2) (2016), 10:1–10:25. doi:[10.1145/2850417](https://doi.org/10.1145/2850417).



- [94] M. Lloyd-Kelly and A. Wyner, Arguing about emotion, in: *Advances in User Modeling – UMAP 2011 Workshops*, LNCS, Vol. 7138, Springer, 2012, pp. 355–367.
- [95] S. Lukin, P. Anand, M. Walker and S. Whittaker, Argument strength is in the eye of the beholder: Audience effect in persuasion, in: *Proceedings of the European Chapter of Association Computational Linguistics (ACL'17)*, ACL, 2017, pp. 742–753.
- [96] J. Mackenzie, Question begging in non-cumulative systems, *Journal of Philosophical Logic* **8** (1979), 117–133.
- [97] G. Maio and G. Haddock, *The Psychology of Attitudes and Attitude Change*, Sage, 2015.
- [98] B. Martinovski and W. Mao, Emotion as an argumentation engine: Modeling the role of emotion in negotiation, *Group Decision and Negotiation* **18**(3) (2009), 235–259. doi:[10.1007/s10726-008-9153-7](https://doi.org/10.1007/s10726-008-9153-7).
- [99] P. Matt and F. Toni, A game-theoretic measure of argument strength for abstract argumentation, in: *Proceedings of European Conference on Logics in Artificial Intelligence (JELIA'08)*, LNCS, Vol. 5293, 2008, pp. 285–297.
- [100] P. McBurney and S. Parsons, Games that agents play: A formal framework for dialogues between autonomous agents, *Journal of Logic, Language and Information* **11** (2002), 315–334. doi:[10.1023/A:1015586128739](https://doi.org/10.1023/A:1015586128739).
- [101] P. McBurney, R. van Eijk, S. Parsons and L. Amgoud, A dialogue-game protocol for agent purchase negotiations, *Journal of Autonomous Agents and Multi-Agent Systems* **7** (2003), 235–273. doi:[10.1023/A:1024787301515](https://doi.org/10.1023/A:1024787301515).
- [102] R. Medellin-Gasque, K. Atkinson and T. Bench-Capon, Persuasion strategies for argumentation about plans, in: *Computational Models of Argument (COMMA 2012)*, IOS Press, 2012, pp. 334–341.
- [103] S. Modgil, Reasoning about preferences in argumentation frameworks, *Artificial Intelligence* **173** (2009), 901–934. doi:[10.1016/j.artint.2009.02.001](https://doi.org/10.1016/j.artint.2009.02.001).
- [104] S. Modgil and H. Prakken, The ASPIC+ framework for structured argumentation: A tutorial, *Argument and Computation* **5**(1) (2014), 31–62. doi:[10.1080/19462166.2013.869766](https://doi.org/10.1080/19462166.2013.869766).
- [105] N. Nahai, *Webs of Influence: The Psychology of Online Persuasion*, 2nd edn, Pearson, 2017.
- [106] National Institute for Health and Clinical Excellence (eds), *Weight Management: Lifestyle Services for Overweight or Obese Adults*, Public Health Guideline, Vol. 53, NICE, 2014.
- [107] H. Nguyen and J. Masthoff, Designing persuasive dialogue systems: Using argumentation with care, in: *Proceedings of the International Conference on Persuasive Technology (Persuasive'08)*, LNCS, Vol. 5033, Springer, 2008, pp. 201–212.
- [108] H. Nguyen, J. Masthoff and P. Edwards, Persuasive effects of embodied conversational agent teams, in: *Proceedings of the International Conference on Human–Computer Interaction*, LNCS, Vol. 4552, Springer, 2007, pp. 176–185. ISBN 978-3-540-73110-8. doi:[10.1007/978-3-540-73110-8\\_19](https://doi.org/10.1007/978-3-540-73110-8_19).
- [109] F. Nouioua and V. Risch, Argumentation frameworks with necessities, in: *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'11)*, Lecture Notes in Computer Science, Springer, 2011, pp. 163–176.
- [110] J. Ogden, *Health Psychology*, McGraw-Hill, 2011.
- [111] J. Olsen, Psychological barriers to behaviour change, *Canadian Family Physician* **38** (1992), 309–319.
- [112] N. Oren, K. Atkinson and H. Li, Group persuasion through uncertain audience modelling, in: *Computational Models of Argument (COMMA'12)*, IOS Press, 2012, pp. 350–357.
- [113] N. Oren and T. Norman, Arguing using opponent models, in: *Proceedings of International Workshop of Argumentation in Multi-Agent Systems (ArgMAS'09)*, LNCS, Vol. 6057, Springer, 2009, pp. 160–174.
- [114] S. Parsons, M. Wooldridge and L. Amgoud, Properties and complexity of some formal inter-agent dialogues, *Journal of Logic and Computation* **13**(3) (2003), 347–376. doi:[10.1093/logcom/13.3.347](https://doi.org/10.1093/logcom/13.3.347).
- [115] S. Polberg and A. Hunter, Empirical evaluation of abstract argumentation: Supporting the need for bipolar and probabilistic approaches, *International Journal of Approximate Reasoning* (2018), in press.
- [116] S. Polberg, A. Hunter and M. Thimm, Belief in attacks in epistemic probabilistic argumentation, in: *Proceedings of the International Conference on Scalable Uncertainty Management (SUM'17)*, LNCS, Vol. 10564, Springer, 2017, pp. 223–236.
- [117] H. Prakken, Coherence and flexibility in dialogue games for argumentation, *Journal of Logic and Computation* **15**(6) (2005), 1009–1040. doi:[10.1093/logcom/exi046](https://doi.org/10.1093/logcom/exi046).
- [118] H. Prakken, Formal systems for persuasion dialogue, *Knowledge Engineering Review* **21**(2) (2006), 163–188. doi:[10.1017/S0269888906000865](https://doi.org/10.1017/S0269888906000865).
- [119] J. Prochaska and W. Velicer, The transtheoretical model of health behavior change, *American Journal of Health Promotion* **12** (1997), 38–48. doi:[10.4278/0890-1171-12.1.38](https://doi.org/10.4278/0890-1171-12.1.38).
- [120] A. Rago, F. Toni, M. Aurisicchio and P. Baroni, Discontinuity-free decision support with quantitative argumentation debates, in: *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR'16)*, AAAI Press, 2016, pp. 63–73.
- [121] I. Rahwan and K. Larson, Mechanism design for abstract argumentation, in: *Proceedings of the Conference on Autonomous Agents and MultiAgent Systems (AAMAS'08)*, IFAAMAS, 2008, pp. 1031–1038.
- [122] I. Rahwan and K. Larson, Pareto optimality in abstract argumentation, in: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'08)*, AAAI Press, 2008, pp. 150–155.

- [123] I. Rahwan, K. Larson and F. Tohmé, A characterisation of strategy-proofness for grounded argumentation semantics, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'09)*, IJCAI, 2009, pp. 251–256.
- [124] I. Rahwan, M. Madakkatel, J. Bonnefon, R. Awan and S. Abdallah, Behavioural experiments for assessing the abstract argumentation semantics of reinstatement, *Cognitive Science* **34**(8) (2010), 1483–1502. doi:[10.1111/j.1551-6709.2010.01123.x](https://doi.org/10.1111/j.1551-6709.2010.01123.x).
- [125] J. Richardson, *Analysing Newspapers: An Approach from Critical Discourse Analysis*, Palgrave Macmillan, 2007. doi:[10.1007/978-0-230-20968-8](https://doi.org/10.1007/978-0-230-20968-8).
- [126] T. Rienstra, Towards a probabilistic dung-style argumentation system, in: *Proceedings of the International Conference on Agreement Technologies (AT'12)*, CEUR Workshop Proceedings, CEUR-WS.org, 2012, pp. 138–152.
- [127] T. Rienstra, M. Thimm and N. Oren, Opponent models with uncertainty for strategic argumentation, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'13)*, IJCAI, 2013.
- [128] R. Riveret, H. Prakken, A. Rotolo and G. Sartor, Heuristics in argumentation: A game theory investigation, in: *Computational Models of Argument (COMMA'08)*, IOS Press, 2008, pp. 324–335.
- [129] A. Rosenfeld and S. Kraus, Providing arguments in discussions based on the prediction of human argumentative behavior, in: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'15)*, AAAI Press, 2015, pp. 1320–1327.
- [130] A. Rosenfeld and S. Kraus, Providing arguments in discussions on the basis of the prediction of human argumentative behavior, *ACM Transactions on Interactive Intelligent Systems* **6** (2016), 30:1–30:33. doi:[10.1145/2983925](https://doi.org/10.1145/2983925).
- [131] A. Rosenfeld and S. Kraus, Strategic argumentative agent for human persuasion, in: *Proceedings of the European Conference on Artificial Intelligence (ECAI'16)*, IOS Press, 2016, pp. 320–328.
- [132] H. Simons and J. Jones, *Persuasion in Society*, 2nd edn, Routledge, 2011.
- [133] J. Swindell, A. McGuire and S. Halpern, Beneficent persuasion: Techniques and ethical guidelines to improve patient's decisions, *Annals of Family Medicine* **8** (2010), 260–264. doi:[10.1370/afm.1118](https://doi.org/10.1370/afm.1118).
- [134] K. Sycara, Argumentation: Planning other agent's plans, in: *Proceedings of the International Joint Conference in Artificial Intelligence (IJCAI'89)*, Morgan Kaufmann, 1989, pp. 517–523.
- [135] M. Thimm, A probabilistic semantics for abstract argumentation, in: *Proceedings of European Conference on Artificial Intelligence (ECAI'12)*, IOS Press, 2012, pp. 750–755.
- [136] M. Thimm, Strategic argumentation in multi-agent systems, *Künstliche Intelligenz* **28** (2014), 159–168. doi:[10.1007/s13218-014-0307-2](https://doi.org/10.1007/s13218-014-0307-2).
- [137] M. Thimm, S. Villata, F. Cerutti, N. Oren, H. Strass and M. Vallati, Summary report of the first international competition on computational models of argumentation, *AI Magazine* **37** (2016), 102–104. doi:[10.1609/aimag.v37i1.2640](https://doi.org/10.1609/aimag.v37i1.2640).
- [138] R.J. Thomas, J. Masthoff and N. Oren, Adapting healthy eating messages to personality, in: *Proceedings of the International Conference on Persuasive Technology (Persuasive 2017)*, LNCS, Vol. 10171, 2017, pp. 119–132.
- [139] S. Timmer, J. Meyer, H. Prakken, S. Renooij and B. Verheij, A two-phase method for extracting explanatory arguments from Bayesian networks, *International Journal of Approximate Reasoning* **80** (2017), 475–494. doi:[10.1016/j.ijar.2016.09.002](https://doi.org/10.1016/j.ijar.2016.09.002).
- [140] F. Toni, A tutorial on assumption-based argumentation, *Argument and Computation and Computation* **5**(1) (2014), 89–117. doi:[10.1080/19462166.2013.869878](https://doi.org/10.1080/19462166.2013.869878).
- [141] A. Tversky and D. Kahneman, The framing of decisions and the psychology of choice, *Science* **211**(4481) (1981), 453–458. doi:[10.1126/science.7455683](https://doi.org/10.1126/science.7455683).
- [142] J. Vargheese, S. Sripada, J. Masthoff and N. Oren, Persuasive strategies for encouraging social interaction for older adults, *International Journal of Human Computer Interaction* **32**(3) (2016), 190–214. doi:[10.1080/10447318.2016.1136176](https://doi.org/10.1080/10447318.2016.1136176).
- [143] J. Vargheese, S. Sripada, J. Masthoff, N. Oren, P. Schofield and V. Hanson, Persuasive dialogue for older adults: Promoting and encouraging social interaction, in: *Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, 2013, pp. 877–882.
- [144] B. Verheij, Proof with and without probabilities. Correct evidential reasoning with presumptive arguments, coherent hypotheses and degrees of uncertainty, *Artificial Intelligence and Law* **25** (2017), 127–154. doi:[10.1007/s10506-017-9199-4](https://doi.org/10.1007/s10506-017-9199-4).
- [145] D. Walton, *Informal Logic: A Handbook for Critical Argumentation*, Cambridge University Press, 1989.
- [146] D. Walton, *Fundamentals of Critical Argumentation*, Cambridge University Press, 2006.
- [147] D. Walton and E. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, SUNY Press, 1995.
- [148] D. Walton, C. Reed and F. Macagno, *Argumentation Schemes*, Cambridge University Press, 2012.
- [149] M. Zwinderman, A. Shirzad, X. Ma, P. Bajracharya, H. Sandberg and M. Kaptein, Phone row: A smartphone game designed to persuade people to engage in moderate-intensity physical activity, in: *Persuasion Technology 2012*, LNCS, Vol. 7822, Springer, 2012, pp. 55–66.