

Persuasion Models For Intelligent Interfaces

Marco Guerini, Oliviero Stock and Massimo Zancanaro

ITC-irst, Istituto per la Ricerca Scientifica e Tecnologica
I-38050 Trento, ITALY
{guerini/stock/zancana}@itc.it

Abstract

Future intelligent interfaces will have contextual goals to pursue. As opposed to more traditional scenarios of Human Computer Interaction, the user interface may also aim at inducing the user, or in general the audience, to perform some actions in the real world. Some scenarios of application are dynamic advertisement, preventive medicine, social action and edutainment.

In this prospect we are investigating persuasion mechanisms and how these are connected to other related concepts such as natural argumentation. In modelling persuasion we distinguish argumentation as a subpart of it, because persuasion is also concerned with arational elements. We take a cognitive approach, considering the state of the participants on the basis of their beliefs-desires-intentions, but also their social relations, their emotions and the context of interaction. In this paper we propose a taxonomy of persuasive strategies and a meta-reasoning model that works on this taxonomy. In this paper the focus is on the high level planning of the proposed system: how it is structured and how it is combined with the adoption of appropriate rhetorical strategies (and other elements such as lexical choice) for producing an effective and context-adapted message. The approach is also at the basis of multimodal developments.

1 Introduction

Future intelligent interfaces will have contextual goals to pursue. They may aim at inducing the user, or in general the audience, to perform some actions in the real world. They will have to take into account the “social environment”, exploit the situational context, and value emotional aspects in communication.

Some foreseeable scenarios of this kind are: dynamic advertisement, preventive medicine, social action, edutainment. In all these scenarios rational reasoning is not enough. For intention adoption what often really matters is not only the content but the overall impact of the communication.

The aim of our research is to provide interfaces with the capability of reasoning on the effectiveness of the

message, as well as on the high-level goals and content. To this end, we have focused on persuasion mechanisms.

According to Perelman [Perelman & Olbrechts-Tyteca, 69], persuasion is a skill that human beings use in order to make their partners perform certain actions or collaborate in various activities, see also [Moulin *et al.*, 02]. Argumentation has often been considered as addressing similar points. In our view persuasion is a wider concept: argumentation can be regarded as a resource for persuasion, while negotiation puts the accent on interactivity in argumentation.

In the first place, persuasion is a “superset” of argumentation: while argumentation is concerned with the goal of making the receiver believe a certain proposition (goal to induce a belief), persuasion is concerned with the goal of making the receiver perform a certain action (goal to induce an action). The link relies on the fact that, apart from coercion, the only way to make someone doing something is to change his beliefs [Castelfranchi, 96]. That is to say: if our goal is to induce an action then we must also have the goal to induce a belief. In this prospect argumentation is a resource for persuasion.

The statement that there is more than argumentation in persuasion refers to the fact that persuasion is also concerned with a-rational elements. Examples are inducing emotions as a mean for obtaining a given result, or the use of specific language for threatening or promising. They can all be regarded as resources for inducing the receiver to act in a desired way.

Natural argumentation comes closer to persuasion, as it is also concerned, for example, with the problem of the adequacy - effectiveness - of the message [Fiedler & Horacek, 02]. Even in professional settings, such as juridical argumentation, extra-rational elements can play a major role [Lodder, 99].

Recent works have studied applications of natural argumentation [Walton & Reed, 02; Das, 02]; argumentation-based text generation has been proposed by Zukerman [Zukerman *et al.*, 00], relying on a Bayesian approach. Negotiation has also been widely investigated and modelled in a computational framework; see for instance [Kraus *et al.*, 98; Parson & Jennings, 96].

2 Modelling Persuasion

In modelling persuasion we adopt the BDI model [Cohen & Levesque, 90] as a reference framework, focusing our attention on the four following aspects:

1. The cognitive state of the participants (beliefs and goals of both the user and the interface)
2. Their social relations (social power, shared goals, etc.)
3. Their emotional state (both the emotional state of the user and the one expressed by the system)
4. The context in which the interaction takes place.

In the following some key aspects are briefly reviewed. Examples are taken from a museum visit domain. In this scenario an intelligent interface agent assists the visitor through the visit.

2.1 Cognitive Elements.

Beliefs and goals, of both the user and the system, about the domain of the interaction are essential. They are prerequisites for a persuasive interaction to take place, since persuasion is a type of communication leading to belief adoption, with the overall goal of inducing an action by the user by modifying his pre-existent goals. Goals and beliefs are mainly concerned with the visit (regarding paintings, books on the museum, and so on). Below some examples of significant concepts:

Value of a Goal: goals constitute a hierarchy, defined by their importance. We must consider this aspect if there are contrasting goals involved in our persuasive efforts.

Consequence of an action: every action leads to a change in the state of the world. Every state of the world can be more or less compatible with the goals of an agent. Actions are chosen so that the global value of the resulting goals is the greatest.

Value of an action: every action permits to fulfill some goals and endangers others. The value of an action is directly proportional to the value of the goals it permits to fulfill and is inversely proportional to the value of the goals it endangers.

2.2 Social Elements

Social relations exist between the visitor and the system (the system playing the role of a competent museum guide, for example) and between the visitor and other relevant persons such as experts, parents, friends and so on. These social elements can be used both to induce beliefs (as in the case of experts) and to induce actions (e.g. using references to the social consequence of an action, consequences that concern the relation between the user, such as a pupil, and relevant persons: “what would your teacher think of you if you don’t ...”). Below are some examples of significant concepts [Castelfranchi, 90; Mead, 34; Ainsworth & Bowlby, 91]:

Power on: if we have the possibility to perform an action that jeopardizes or helps achieve a goal of another agent, then we have some power on that agent. In particular the possibility of jeopardizing is defined as *power-on-jeopardize* and the possibility of helping as *power-on-help*. The more important the goal, the

stronger the value of the *power-on*. This concept furnishes the basis for threatening and promising.

Tutorial goal: the goal of an agent x to influence an agent y to (have the intention to) perform actions that are in the interest of y without y ’s explicit awareness of that.

Significant-other: the depositary of y ’s moral standards and values. In the example above: the teacher, a significant-other for the pupil, is used in the definition of - social - consequences of an action.

Attachment figure: defined as the agent that has a tutorial goal on the receiver and has an affective value for him. E.g. parents for children.

2.3 Emotional Elements

Emotional elements can enhance or diminish the effectiveness of the message. There are at least three dimensions to be considered: the current emotional state of the user, the one expressed by the interface, and the one produced on the user by the interface communication. In this work we focus on the first two: the emotional state of the receiver (how it affects strategy selection) and the emotion the system has to convey (express) to maximize the effectiveness of the message.

2.4 Contextual Elements

Persuasion strategies can make use of contextual elements to induce an action, e.g. making reference to a painting the visitor has seen previously (“this painting is by the same author of ...”) can increase the probability the user will stop in front of the current painting.

2.5 An example

Let us consider the situation in which the museum guide has the goal of persuading an absent-minded pupil to pay attention to a specific painting. The system could use a persuasive strategy such as threatening the kid to tell the teacher he did not pay attention during the visit. But it can do so only if it knows that: a) for him the judgment of the teacher is important (it weighs more than the boredom derived from paying attention to something that is not of his interest); b) it can do some action to prevent the pupil from reaching his previous, contrasting, goal (to have a good reputation in front of the teacher), and c) it can convey its intention in an appropriate manner (in relation to the emotional state of the pupil). The “deterrence” of the interface must also be credible to the kid.

3 Persuasive Strategies

There is a rich repertoire of persuasive strategies coming from social psychology, see for example [Cialdini, 93; Miceli, 92] that can be accommodated within this framework. This collection is not structured: social, emotional and cognitive aspects interact with each other. Perelman [Perelman & Olbrechts-Tyteca, 69] proposes a structuring from a philosophical point of view, that cannot be used as is for our computational purposes. We propose (a) a criteria for aggregation and (b) an initial taxonomy to account for similarities among strategies.

Let us consider the example introduced earlier: the system can decide to promise something (instead of

threatening) if the pupil is very tired or depressed. Different emotional states of the user can lead to different courses of communicative actions (threatening vs. promising), even if the initial communicative goal is the same. Thus, in this case there are two different strategies:

- S1. *If the agent has the power to jeopardize a goal of the user, **and** the goal has a higher value than the value of the action the agent wants him to perform, **and** the user is “lively”, **then** the agent can threaten the user to jeopardize the goal if he does not perform the required action.*
- S2. *If the agent has the power to help the user achieve a goal, **and** the goal has a higher value than the value of the action the agent wants him to perform, **and** the user is “tired or depressed”, **then** the agent can promise to help him realize his goal if he performs the required action.*

In order to obtain aggregation we modelled these strategies in BDI form using the concepts introduced above. Below y is the user, x the interface agent, a the expected action, $S1$ and $S2$ denote world states. For the sake of simplicity, everything is meant to be in the mutual belief space of the interface agent x .

Here is a formalization of rule S1 above:

LIVELY (y) GOAL (y S1) GOAL (x a) POWER-ON-JEOPARDIZE(x y S1) VALUE (y S1) > VALUE (y a)
CONDITIONAL-THREATEN (x y S1 a)

Here y is the pupil, a is y paying attention to the visit and $S1$ is a world state where the pupil has good reputation (social image) in front of the teacher. This rule can yield a sentence like: “If you don’t pay attention during the visit I will tell the teacher”.

Here is a formalization of rule S2 above:

TIRED (y) GOAL (y S2) GOAL (x a) POWER-ON-HELP (x y S2) VALUE (y S2) > VALUE (y a)
CONDITIONAL-PROMISE (x y S2 a)

Here y is the pupil, a is y paying attention to the visit and $S2$ is a world state where the pupil has no homework to do. This rule can yield a sentence like: “If you pay attention to the visit I will automatically make a report of the visit so you do not have to prepare your homework”.

In a similar way many other strategies can be specified and some may avoid any argumentative aspect. An example of a strategy completely based on emotion triggering is the one expressed by the sentence: “You know your mother cares for you, would you upset her?” This strategy uses the social concepts of tutorial goal and of attachment figure to trigger the guilty-feeling of the receiver [Ortony *et al.*, 88].

3.1 Rule Classification and an Initial Taxonomy

Rules have been aggregated along two different dimensions: goal-to-induce-belief and goal-to-induce-action. There is a mono-directional relation between the two groups; the second can exploit the first one, but not vice-versa.

Strategies such as *appeal to expert opinion*, *appeal to popular opinion*, and *appeal to empirical evidence* [Walton & Reed, 02] are all concerned with the need for x to enhance the probability of y accepting the content conveyed. They all belong to the first group.

The second group includes strategies concerned with the goal of x to enhance the probability that y performs the required action. These strategies can be used both to show the positive and the negative consequences of an action (one of the main instrument of persuasion). To this group belong strategies such as *appeal to significant-other*, *appeal to social-image*, *artificial-consequences* (negative artificial consequences are called *threatening*, positive artificial consequences are called *promises*).

As remarked above, strategies in the second group can use strategies in the first group but not vice-versa. It is impossible to persuade y to believe something by showing him the negative consequences deriving from not doing it, or to threat y to induce him to believe it. On the contrary, we can make appeal to an expert opinion to motivate our claim that if y does not perform an action he will encounter negative consequences.

3.2 Rules for Abstract-Reasoning

Some persuasive strategies involve application of “abstract-reasoning” on others strategies. These are concerned, for example, with the conveying of a particular mood. The following is used to enhance the motivation of the other (*motivate abstract-strategy*):

*AS1. if the user is “tired” **and** there are applicable strategies showing positive consequences deriving from performing the required action **then** use a “happy” mood.*

3.3 Rules For Meta-Reasoning

Some persuasive strategies involve application of “meta-reasoning” on others strategies. This meta-reasoning includes three different operations that can be made on the other strategies: a selection operation, a modification operation, and an ordering operation.

Selection of strategies: turning back to the example of the museum guide, we can interpret the sentence “if the user is tired it is better not to threaten him” as a condition of application of the threatening-strategy. But, exploiting taxonomy of strategies we can define it as a decision, at a more abstract level, relative to the group of strategies concerned with the showing negative consequences of an action. The structure of the strategy that we may call *don’t upset the user* is:

*MS1. if the user is “tired” **and** there are applicable strategies showing positive consequences **and** there are applicable strategies showing nega-*

ive consequences, about the required action **then** the agent should avoid using strategies showing negative consequences.

For example this meta-rule allows us not only to give account for the decision between promising something to the pupil (positive artificial consequence) and threatening him (negative artificial consequence) when he is tired. It accounts for an entire class of interactions between the cognitive and the emotional level.

Another meta-rule that accounts for the interaction of the context and the other aspects of persuasion is *avoid repetition*:

*MS2. if a strategy (S2) has just been used **and** there is another strategy (S1) available **then** the agent should use S1.*

Modification of strategies: meta-strategies concerned with the modification of the features of strategies.

Ordering of strategies: meta-strategies mainly concerned with the order in which the strategies have to be presented. Let us consider the *show defect* strategy, used to catch the trust of an “expert” user:

*MS3. if the user is “expert” **and** there are applicable strategies showing negative consequences deriving from performing the required action **and** there are applicable strategies showing positive consequences deriving from performing the required action **then** the agent can catch the trust of the user by putting the strategies showing negative consequences first.*

4 General Architecture

The overall architecture of our system is based on three different levels of processing (figure 1). The strategies introduced above belong to the first level, and the focus of this work is mainly devoted to this level.

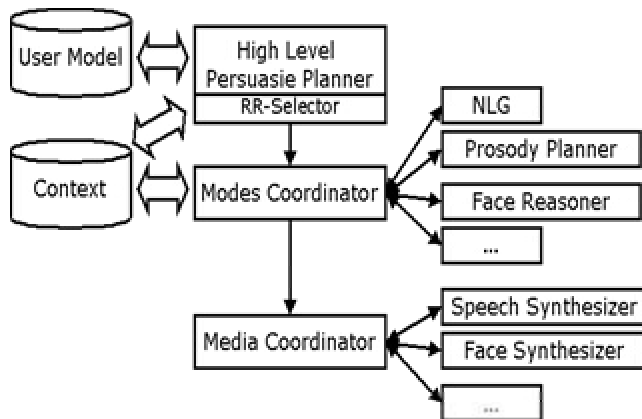


Figure 1

- The first level selects, on the basis of the cognitive, social and emotional states of the participants, the possible strategies to pursue for persuading (like threatening or promising). It results in a perlocutionary act and/or an illocutionary act specification. In addition it can specify the emotional state the interface agent has to display, and provide other information to pass on to the next level.

- At the second level, realization strategies are defined involving the available modes. For instance a natural language generator plans the rhetoric structure of the text. Realizers take into account information passed on by the first level. Selection is performed along the indication of the first level, but it is independent from this one. For example a severe mood can be conveyed by a harsh intonation. Or it can be conveyed by a text whose words are selected taking into account their emotional values (for example, selecting “kick the bucket” instead of “die”). Or it can be conveyed even by font selection.
- At the third level coordination of the different available media is performed so that different components are integrated and synchronized.

This architecture is intrinsically multimodal. Each component performs the realization of a portion of the message in its own way, including characterizing it emotionally.

Coordination takes place at the second and third level of the system. At the second level content is allocated to the different modes and coordinated according to rhetoric rules. For example, the harsh intonation of the synthetic speech can be accompanied by a serious facial expression of an embodied agent. At the third level, a different kind of coordination takes place, involving mainly temporal constraints, for example synchronization between certain facial muscular movements and pitch levels in synthesized speech.

Implementation

The theory introduced here has been implemented in an initial version. The High Level Persuasive Planner (HLPP) is realized within CLOS-actions framework [Strapparava & Zancanaro, 01], an extension of CLOS to allow flexible rule management and meta-object definition.

Persuasive strategies are encoded using a Clos-hierarchy and CLOS-actions-rules, while Meta-reasoning is implemented with CLOS-actions-rules. There are two main hierarchies: one for the persuasive strategies, and another one that contains the social, contextual, cognitive and emotional elements of the “world” necessary for strategy selection.

There are different rules operating on these hierarchies. Rules for *generating strategies*, that operate on the world-hierarchy and produce instances of strategies in the strategy-hierarchy. Rules for *selecting strategies*, operating on the strategy-hierarchy that delete instances of strategies. Rules for *modifying strategies*, operating on the strategy-hierarchy that modify instances of strategies. Rules for *ordering strategies*, operating on the strategy-hierarchy and producing instances of messages containing ordered sequences of strategies.

HLPP is triggered by events happening in the real world. Events such as the user passing in front of a painting activate the goal of the system to describe that painting, or when the visit is at the end (the user reaches the last room) then the goal of the system is activated to persuade the user to buy a particular book. This triggering creates two kinds of goals of the system in the world-hierarchy:

goal-to-induce-belief, where the system (x) wants the user (y) to believe something (in the case mentioned above this is realized by presenting the description of the painting)

goal-to-induce-action, where the system (x) wants the user (y) to perform a certain action a (in the case mentioned above it is the action to buy a book with the description of the museum)

The activation of these goals fires the generation-rules that, according to the elements present in the world-hierarchy select the possible persuasive-strategies to use. Abstract and meta-reasoning in our system are accomplished by means of selection-rules, ordering-rules and modification-rules on the strategy-hierarchy. These rules have the form:

- a) *If (S1 & S2 & Cond) then (DELETE S2)*
- b) *If (S1 & S2 & Cond) then (ORDER (S1,S2))*
- c) *If (S1 & Cond) then (MODIFY S1)*

That is to say:

(A) if strategies of the kind S1 and S2 can be used and a particular condition holds, then choose the strategy S1;

(B) if strategies of the kind S1 and S2 are fired and a particular condition holds, then order them by putting S2 after S1.

(C) if a strategy of the kind S1 is fired and a particular condition holds in the world, then modify it in a particular way.

An example for (A) is *avoid repetition*:

EQUAL (TYPE-OF (S1) TYPE-OF (SX)) SUCCESSOR (TURN (S1) TURN (SX)) S2
DELETE (S2)

An example for (B) is *show defect*:

TYPE-OF (S1 negative-consequence) TYPE-OF (S2 positive-consequence) EXPERTISE-PROFILE (user high)
ORDER (S1 S2)

The output of this level of processing is a sequence of strategies with the specification of the elements necessary for their application.

The sequence of strategies is then processed by the **RR-selector** that has the task to map it onto a Rhetorical Relations tree. This is accomplished by means of patterns that match pairs of strategies on Rhetorical Relations taking in consideration the class of the strategies and their topics (a belief or an action).

For example: a *show defect* meta-rule has generated an order object containing a strategy S1, showing a negative consequence of the action a , and two strategies S2 and S3 showing positive consequences of the same action a .

First of all S2 and S3 are matched on a *conjunction* pattern since (a) S2 and S3 belong to the same class (i.e. *positive-consequence*) and (b) S2 and S3 have the same topic (i.e. the same action a). This produces a text span rendered as “this book contains all the comments on this

museum made by your favorite art critic **and** it contains lots of photos of the paintings”.

Then S1 is added to the sub-tree produced above and matched on a *concession* pattern since (a) S1 and the text span have the same topic (i.e. the same action a) and (b) the text span belong to a class (i.e. *positive-consequence*) opposite to S1 (i.e. *negative-consequence*). This produces a text span rendered as “**though** this book is quite expensive, it contains all the comments on this museum made by your favorite art critic **and** it contains lots of photos of the paintings”.

Interaction in the Museum Scenario

In the final museum scenario, the visitor will have a PDA combined with a locating system [Stock & Zancanaro, 02]. The system starts off, at the beginning of the visit, with an initial user profile. This is the initialization of a dynamic user model that is continuously updated, on the basis of the user behaviour (time spent in front of an exhibit, movement, general attitude in the course of the visit, topic selection on the PDA). Of particular importance is the interest model and the attentional-emotional state (e.g. if the visitor is near to a painting the system is describing but not looking at it, the system can infer the user is bored).

Let us consider the visitor passing in front of a painting. Since the system knows the visitor position, it can possibly assume the goal of setting his attention on that particular painting (goal-to-induce-action). Given that the user has spent a long time in the museum the system also assumes that he is probably “tired”. The system therefore plans a persuasive message by instantiating a goal-to-induce-action. This fires the instantiation of a number of strategies that, in turn, trigger rules and meta-rules. In particular the strategies referring to the negative consequences of not performing the action a “set attention on the painting” are discharged because of the *don't upset the user* meta-rule that refers to the user being tired. Those referring to the impact on the social-image of the user are also discharged because already used in the previous interaction. This is done because of the *avoid repetition* meta-rule. At the end, if an art critic's praise of that paint is available, an *appeal to an expert opinion* can be used to motivate the user. This appeal refers to the positive consequences of the action and, being the user “tired”, the *motivation* abstract rule sets the mood for the final message to “happy”: “**Hey look:** Umberto Eco has stated that this is one of the most interesting painting of the Middle Ages!!!”.

Future Work

For the future, we plan to refine the HLPP, extending the taxonomy and the meta-reasoning capabilities. From the point of view of the model we intend to enrich it with more fine-grained details, regarding, for example, the possibility to reason on goals and beliefs with multivalued logic.

Particular attention will also be posed on the study of patterns to link strategy order and rhetorical relations.

Evaluation will follow as soon as HLPP is integrated with the mobile guide.

Conclusions

Intelligent interfaces will need to be persuasive. This means they must have the capability of reasoning on the effectiveness of the message. In this paper we have proposed a framework that includes four aspects (cognitive, social, emotive, contextual) we believe are fundamental for persuasion mechanisms. Together with this framework an initial computational architecture has been depicted. We have focused especially on the high level planning part of such architecture, proposing a model that exploits meta-reasoning to account for the interaction between the four aspects of persuasion. Finally, this approach has been conceived to support multimodal interactions.

Acknowledgements

This work was supported by the TICCA and the PEACH projects.

References

- [Ainsworth & Bowlby, 91] Ainsworth and Bowlby, An ethological approach to personality development, *American Psychologist* 46:333--341, 1991.
- [Castelfranchi, 90] Castelfranchi C. Social Power, a point Missed in Multi-Agent Systems. Demazeau Y. and Muller J. P. (Eds.), *Decentralized AI*, pages 49--62, Elsevier Science Publisher, 1990.
- [Castelfranchi, 96] Castelfranchi C. Reasons: Beliefs structure and goal dynamics. *Mathware & Soft Computing*, 3(2):233--247, 1996.
- [Cialdini, 93] Cialdini R. B. *Influence. The psychology of persuasion*. William Morrow & Company, Inc., New York, 1993.
- [Cohen & Levesque, 90] Cohen P. and Levesque H. Persistence, Intention, and Commitment. P. Cohen, J. Morgan, and M. Pollack (Eds.), *Intentions in Communication*, pages 33--70, MIT Press, Boston, 1990.
- [Das, 02] Das S. Logic of probabilistic argument. *Proceedings of the ECAI 2002 Workshop on Computational Model of Natural Argument*, pages 9--18, Lyon, 2002.
- [Fiedler & Horacek, 02] Fiedler A., Horacek H. Argumentation within Deductive Reasoning. *Proceedings of the ECAI 2002 Workshop on Computational Model of Natural Argument*, pages 55--64, Lyon, 2002.
- [Kraus et al., 98] Kraus S., Sycara K. and Evenchik A. Reaching agreements through argumentation: a logic model and implementation, *Artificial Intelligence Journal* 104:1--69, 1998.
- [Lodder, 99] Lodder A. R. *DiaLaw: On Legal Justification and Dialogical Models of Argumentation*. Kluwer Academic Publisher, 1999.
- [Mead, 34] Mead G., *Mind, self, and society*, University of Chicago Press, Chicago 1934.
- [Miceli, 92] Miceli M. How to Make someone Feel Guilt: Strategies of Guilt Inducement and Their Goals. *Journal for the theory of social behaviour*. 22(1):81--104, 1992.
- [Moulin et al., 02] Moulin B., Irandoust H., Belanger M. and desbordes G. Explanation and Argumentation Capabilities: Towards the Creation of More Persuasive Agents. *Artificial Intelligence Review* 17:169--222, 2002.
- [Ortony et al., 88] Ortony A., Clore G. L. and Collins A.. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.
- [Parson & Jennings, 96] Parson S. and Jennings N. R. Negotiation through argumentation-a preliminary report. Proceedings of the International Conference on Multi-agent System. Kyoto, 1996.
- [Perelman & Olbrechts-Tyteca, 69] Perelman C. and Olbrechts-Tyteca L. *The new Rhetoric: a treatise on Argumentation*, Notre Dame Press, 1969.
- [Strapparava & Zancanaro, 01] Strapparava C. and Zancanaro M., Firing Rules in an Object World. *Proceedings of AIIA '02 Demo Session*, Bari, 2001.
- [Stock & Zancanaro, 02] Stock O., Zancanaro M., Intelligent Interactive Information Presentation for Cultural Tourism. Invited Talk, *Proceedings of the International Workshop on Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems*, Copenhagen, 2002.
- [Walton & Reed, 02] Walton D. N. and Reed C. A. Argumentation Schemes and Defeasible Inferences. *Proceedings of the ECAI 2002 Workshop on Computational Model of Natural Argument*, Lyon, 2002.
- [Zukerman et al., 00] Zukerman, I., McConachy, R. and Korb, K. Using Argumentation Strategies in Automated Argument Generation. In *INLG'2000 Proceedings -- the First International Natural Language Generation Conference*, pages 55--62, Mitzpe Ramon, Israel, 2000.