Exciting Avocados and Dull Pears

Combining Behavioural and Argumentative Theory for Producing Effective Advice

Floriana Grasso
Heriot-Watt University - Edinburgh
floriana@cee.hw.ac.uk

Abstract

To produce effective advice several sources of knowledge are needed. Knowledge about the application domain the advice is concerned with is of course necessary, but not sufficient. If the aim of the intervention is inducing people to modify their habits, we also need specific theories of how and why people change behaviour to guide the advising process. In some cases, however, it still does not suffice: when suggesting a change in a well established habit, several factors have to be taken into account, and a good adviser might also need argumentative capabilities, in order to overcome possible personal and environmental barriers to the change. This paper presents a model of advice giving that integrates Artificial Intelligence with concepts and methods coming from different disciplines. The model has been implemented in Daphne, an advice giving system that operates in the nutrition education domain.

Keywords. Advice Giving Systems, Argumentation, Health Promotion, Behaviour Modelling, Dialogue.

Introduction

Advice giving systems are often considered as a sub class of tutoring ones, the difference being that they interact with more expert users, who, rather than being instructed in performing a task, only need some suggestions about how to perform it more efficiently (Forslund 1995). A common hypothesis in such systems is that the user's beliefs are *fewer* than the system's, both in qualitative and in quantitative terms. The system being the *expert* in the conversation domain, a disagreement between system's and user's knowledge is considered as a user's misconception. Also, all the concepts that the user is modelled to have on the domain form a subset of system's ones.

In this work we would like to analyze a wider concept of "advice", that is the provision of opinions about what to do, or how to behave, not necessarily with respect to a technical task, but also concerning more general issues, involving personal opinions and values.

Giving advice, in these situations, does not mean only providing knowledge, and, more importantly, does not mean *owning* knowledge. As Hustadt (1994) pointed out, the hypothesis of system's greater expertise should be relaxed when the personal attitude towards the topic is an important subject of discussion. In these cases both the participants are *experts*: the system's beliefs can be seen as opinions, more or less warranted, to be communicated to the user, and argued about, if necessary, and the user will not always eventually share its knowledge. The aim of such a system is then to produce persuasive messages having a more effective impact on the addressee.

In order to design a system whose ultimate aim is to try and influence the user's behaviour very diverse sources of knowledge have to be integrated. Knowledge about the specific domain, knowledge about how individual behaviour is influenced by a variety of beliefs, attitudes and subjective feelings, and knowledge about how argumentation techniques can be used, all have a crucial role in producing effective messages.

In this work, all these sources of knowledge are combined with Artificial Intelligence methods and techniques in a computer system, Daphne, able to argue with a user. The resulting domain independent model is illustrated in the domain of nutrition education.

Nutrition: an Argumentative Scenario

It is generally agreed (Department of Health 1992) that a healthier nutrition habit can have a crucial role in decreasing diet related diseases. Nevertheless, as Fries & Croyle (1993) pointed out, "far too often, what an educator regards as healthful virtues may be regarded as social vices by clients". Enlightening studies showed that several barriers and stereotypes influence people's nutrition habits. A sample of college students interviewed by Fries & Croyle (1993), for example, found that *persons eating low-fat food are picky* and *people with a high-fat diet attend parties*. Similar stereotypes can be found in Sadalla & Burroughs (1981). Ziebland *et al.* (1998) identified some of the external and internal barriers to change in diet, such as *I enjoy eating foods that are bad for me*. As Fries & Croyle (1993) suggest, the kind of barrier is similar to those faced in advertisement: so an educator should show similar argumentative skills.

Moreover nutrition is a topic in which every person feels in a sense to be an expert: we all know, at least roughly, which are the *bad* and the *good* foods. And we certainly are the experts when speaking about what we like to eat!

These characteristics make the nutrition education scenario very appropriate for our study: it involves personal attitudes and values, the addressees have some expertise on the domain, such to allow them to counter-argue with the educator's advice, and a *dialectical* ability is then involved in order

Schema	Definition	Example
Pragmatic Argument	Evaluates acts or events in terms of their consequences	*Eating fruit will make you slimmer*
Arguing by Model	Presents specific case as a model to be imitated. The model must be admirable or have authority or prestige	*Italians know all about healthy eating*
Arguing by Dissociation	Breaks connected links, reorganizing the addressee's concept of reality	*You said that people who are concerned about diet are self-centred, but I prefer to consider them just responsible persons*

Figure 1: Three New Rhetoric's schemas

to overcome the lack of willingness to accept educator's point of view.

The Theoretical Basis

An objective of this research was to appeal to well established theories for each of the fields of exploration: health promotion and argumentation. This section briefly details the characteristics of the theories which contributed to the realization of the system.

Health Promotion

Some of the theoretical models in health research were examined, in order to find the appropriate framework for our formalization. We found two of them which, with minor changes, could be easily formalized into a computer system.

The transtheoretical model of change, or **Stages of Change Model**, (Prochaska & DiClemente 1992) suggests that individuals progress through very distinct stages of change on their way to change their behaviour. In a first *precontemplation* stage, people see no problem with their behavior and have no intention of changing it. A *contemplation* stage comes when people come to understand their problem, and start thinking about solving it, but have no immediate plans. In a following *preparation* stage, people are planning to take an action in the immediate future, and have already made some small changes in this directions. The *action* stage involves people who are actually in the process of actively making behavior changes. Finally, the *maintenance* stage characterizes a period of behavior continued on a regular basis; 6 months is generally agreed to be a good measure of time. The process of change, however, is not linear and a *relapse* is always possible at each stage.

The **Health Belief Model** (Becker 1974) assumes that for people to actually take an action to avoid a negative situation they need to believe that: they are personally *susceptible* to it; its occurrence would have at least moderate *severity* on some components of their life; taking a particular action would in fact *reduce their susceptibility* to it or, if it occurs, would *reduce its severity*; taking a particular action would not cause them overcoming important (for them) *barriers* (eg. cost, convenience, pain, embarrassment and so forth).

The two models are successfully used in many healthy diet promotion interventions (Roe *et al.* 1997).

A Theory of Argumentation

In looking for a theory able to express everyday arguments, we excluded those dealing only with *demonstrative* reasoning, being more concerned about the effectiveness of arguments than their validity.

Perelman's **New Rhetoric** (Perelman 1979) met our requirements. It is a theory of argumentation which deals with dialectic reasoning, defining an argument as more or less *successful* when its premises make the conclusion plausible to a greater or lesser degree. The concept of audience is crucial in this scenario, as the same argument can produce very different results when addressed to different kinds of people.

According to Perelman's definitions, the orator can build an argument on *premises*, belonging to two wide classes: premises relating to the *real*, that is statements in which the claim can be recognized by the universal audience (for instance *apples are fruit*), and premises relating to the *preferable*, that are statements which have to do with the preference of a particular audience, consisting of personal *Values* (for instance *natural products have to be preferred*), and the ways in which an audience arranges them in *Hierarchies*.

Especially useful for our purposes is the New Rhetoric's concept of *schema*, as a mechanism used by the arguer to arrange premises in such a way that the audience will be most influenced. Perelman proposed several classes of schemas. Warnick & Kline (1992) elaborated the classification, proposing an accurate collection of coding guidelines for each schema. Examples of schemas are given in Fig. 1.

Daphne: Dialectical Argumentation for Providing Healthy Nutrition Education

The design of Daphne was meant to follow two main guidelines: first, it should show a high modularity, that is each single component should be as much independent as possible from the others, in order to allow an easy upgrade. Second, it should employ simple and practical techniques. Figure 2 depicts the system architecture. The theories presented influenced our decisions in particular on the definition of system's reasoning ability and on the establishment of the domain language. Its main components are:

Reasoning unit: it is responsible for managing the factual knowledge of the system, about the domain and the user. It consists of three classes, the third one being split into two sub-classes:

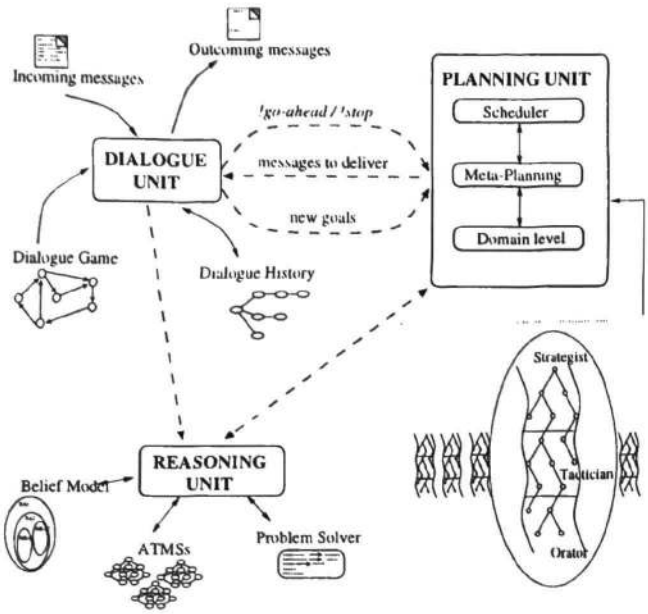

Figure 2: Daphne's Architecture

class 1: the system's own beliefs (B_S);

class 2: the system's beliefs about user's private beliefs ($B_S B_U$);

class 3.1: the system's beliefs about what the user believes be mutually believed about the system's beliefs ($B_S B_U M B_S$);

class 3.2: the system's beliefs about what the user believes be mutually believed about the user's beliefs ($B_S B_U M B_U$).

And each other type of nested belief is collapsed into one of the classes at level 3. This classification was chosen to allow the system to be not fully sincere. Some sort of deception is in a sense inherent in the definition of dialectic argumentation. The fact that the arguer bases his or her justification on the audience's beliefs and not on what he or she really thinks on the matter can be seen as a subtle kind of deception. In particular, the arguer may like a claim to pass as his or her own when it is only a projection of the audience's mind (for example, the claim *Eating fruit is important because it helps slimming* can be made by an arguer who knows the audience gives a high value to being slim, even if he or she does not). With this model it is possible to represent sincere communication, with the possibility for an agent to misinterpret it, and insincere communication about one agent's own beliefs (and the detection or not of it by the other agent). However, it is not possible to deceive about one's beliefs about the other one (I cannot say *I believe you believe X* if I think you disbelieve X). Figure 3 represents how the various parts of the belief models of the speaker and the hearer are updated after different sorts of communication, showing that the four sets proposed are necessary and sufficient to represent them.

A problem solver, using Horn clauses, and an ATMS (de Kleer 1986) to guarantee coherence among beliefs are used in this unit.

Planner unit: The system was designed to have a planning component, whose aim is to decide about built-in goals and trigger follow-up actions when solicited. The presence of a neat separation of the sources of knowledge used, suggested the same neat separation of the competences of the planning operators into conceptual abstraction spaces (Sacerdoti 1974). Three spaces are defined: the high level space, or *strategist*, is concerned about the general goal of the system about the promotion of a given behavior; this space formalizes the knowledge about the Stages of Change and the Health Belief models. The middle level space, or *tactician*, is concerned about the argumentative skills of the system; it is in a sense independent from the subject of the argumentation, and embeds knowledge from the New Rhetoric domain. The bottom level space, or *orator*, has strictly linguistic knowledge: it is concerned about how to express, in a natural language, the argument proposed by the tactician. Moreover, the planner has a structural hierarchy of planning layers (Stefik 1981): a *domain space* layer, a *meta-planning* layer, whose aim is to control the applicability of domain operators, and a *scheduler* layer, which is responsible for controlling the overall process, by activating new goals etc. The output of the planning process consist of communicative actions, that is messages to be delivered to the user. A plan operator has the following structure:

Effect: the goal it aims to achieve;

Prerequisites: the conditions that have to be verified before the operator can be applied;

Decomposition: the description of how the effect is decomposed into subgoals.

Dialogue unit: it is responsible for the actual communication with the user; messages are exchanged between the two conversants, and a track of the dialogue so far is kept. The dialogue mechanism is mostly inspired by the metaphor of a dialogue game: the two conversants are viewed as players, each of them having a certain number of allowed moves in every situation of the game (Carlson 1983). An important task of the Dialogue unit is to suggest following up goals to the planner, on the basis of the message received. An example of this ability will be shown later. Daphne and its user communicate by means of a mailbox: Daphne reads messages the user sent in an incoming mailbox, and put its messages to the user in an outcoming mailbox.

Agent central unit: the aim of this unit is to activate and coordinate all the others; there is no explicit mention of it in Fig. 2, but its role can be represented by the dashed lines between the various components.

Knowledge base definition: the definition of the language of the domain was also affected by the theories used. Val-

ues are implemented as the attribution of a certain *degree of importance* to an object. As different values can be associated with the same object for different reasons, the specification of the *perspective* or point of view from which the attribution is considered is included (for example *Having a slim figure is good from the social life perspective*). Other predicates in the language were defined after the concepts of the behavioural models: we needed to express the *successfullness* of actions with respect to a given purpose, the association of the user's stage of change towards an action and so on. Finally, to realize the New Rhetoric's schemas, we needed to represent concepts such as *being a model* or *having consequences* etc.

A Detailed Example

This section will show Daphne's behaviour with respect to the following short dialogue (system and user's turns are numbered and denoted with S and U respectively).

S1:	*Have you considered eating more fruit? It is good for your health, as it helps you preventing cancer.*
U1:	*Fruit is boring! It's granny's stuff!*
S2:	*Maybe pears are. But what about avocados? Did you know it's Californians' favourite?*

First System's Turn: S1

The system is trying to move the user from the precontemplation to the contemplation stage with respect to eating fruit. The following strategic operator applies:

Effect	PrecToCont (EatingFruit)
Prereq	$B_S B_U$ (Precontemplator(EatingFruit)) &
	B_S (Successful (EatingFruit,Prevent(Cancer)))
Decomp	PragmaticArgueNeg(EatingFruit, Cancer)

The decomposition refers to a tactician operator, so the control is passed to the next abstraction level. The evoked operator involves a pragmatic argumentation, labelled as "negative" because it stresses on what the proposed action avoids rather than achieves.

Effect	PragmaticArgueNeg(EatingFruit, Cancer)
Prereq	$B_S B_U$ HasVal(Cancer, Health, Bad) &
	B_S (Successful (EatingFruit,Prevent(Cancer)))
Decomp	Claim (HasVal(EatingFruit, Health,Good)) &
	Support (Successful (EatingFruit,Prevent(Cancer)))

The application of the operator also bound the variable indicating the perspective (Health) from which the effect that the action prevents has a Bad value, so passing a Good value from the same perspective to the action itself.

The Claim and Support operators refer to the orator abstraction level, producing the phrase in the dialogue above.

User's response: U1

The user communicated that:

B_U Is(EatingFruit, Boring)
B_U Habit(EatingFruit, Granny)

With the assumption that boredom is not a positive emotion:

$B_S B_U$ HasVal(Boring,Emotional,Bad)

the system can infer that the same value is passed to EatingFruit; it is also inferred that in user's hierarchy the Emotional perspective has a higher position than the Health one; moreover, as it is supported by the fact that EatingFruit is presented as a Granny's habit, it is inferred that Granny is an anti-model for the user from the Emotional perspective:

$B_S B_U$ HasVal (EatingFruit,Emotional,Bad)
$B_S B_U$ MoreValuable (Emotional, Health)
$B_S B_U$ AntiModel (Granny, Emotional)

Second System's Turn: S2

The dialogue manager, after analyzing the input message, can propose to the planner some counter-moves. The main argument of the user involves a member/set relation (EatingFruit is a member of the Boring set). A response to this argument can be a Dissociation, if two instances of the set EatingFruit can be found with different values with respect to the Emotional perspective. The dissociation should emphasize a positive value. The tactician operator to try will be:

Effect	DissociatePos(EatingFruit, Emotional)
Prereq	$B_S B_U M B_U$ HasVal(EatingFruit, Emotional, Bad) &
	$B_S B_U M B_U$ Is(EatingPears, EatingFruit) &
	$B_S B_U$ Is(EatingAvocados, EatingFruit) &
	$B_S B_U$ HasVal(EatingAvocados, Emotional, Good)
Decomp	Concede (HasVal(EatingPears, Emotional, Bad)) &
	Distance (EatingAvocados, EatingPears) &
	Claim HasVal(EatingAvocados, Emotional, Good)

Notice that it is not necessary to assume that EatingPears has a specific negative value: as a common instance of EatingFruit, it is assumed that the user will pass to it the negative emotional value of EatingFruit.

Let us suppose, however, that the system cannot assume that the user will believe that EatingAvocados has a good emotional value, that is the prerequisite is not satisfied and needs to be planned for. The most general tactician operators have the attribution of a value as their effect, and one of the argumentative techniques as their decomposition. For instance, there will be the operator:

Effect	$B_S B_U$ HasVal(EatingAvocados, Emotional, Good)
Prereq	
Decomp	ArgueByModel (EatingAvocados, Emotional)

It will be fired, as the following operator can be fired:

Effect	ArgueByModel (EatingAvocados, Emotional)
Prereq	B_S Habit(EatingAvocados, Californians) &
	$B_S B_U M B_U$ Model(Californians, Emotional)
Decomp	Claim Habit(EatingAvocados, Californians)

Unexplored Alternatives

After receiving the user's response, the system had several choices. A repair to the original plan could have been done at the strategic level: the user was not yet ready to move to the contemplator stage, and it could have been induced by increasing the user's perceived susceptibility to cancer, or by finding another effect that EatingFruit can achieve or prevent.

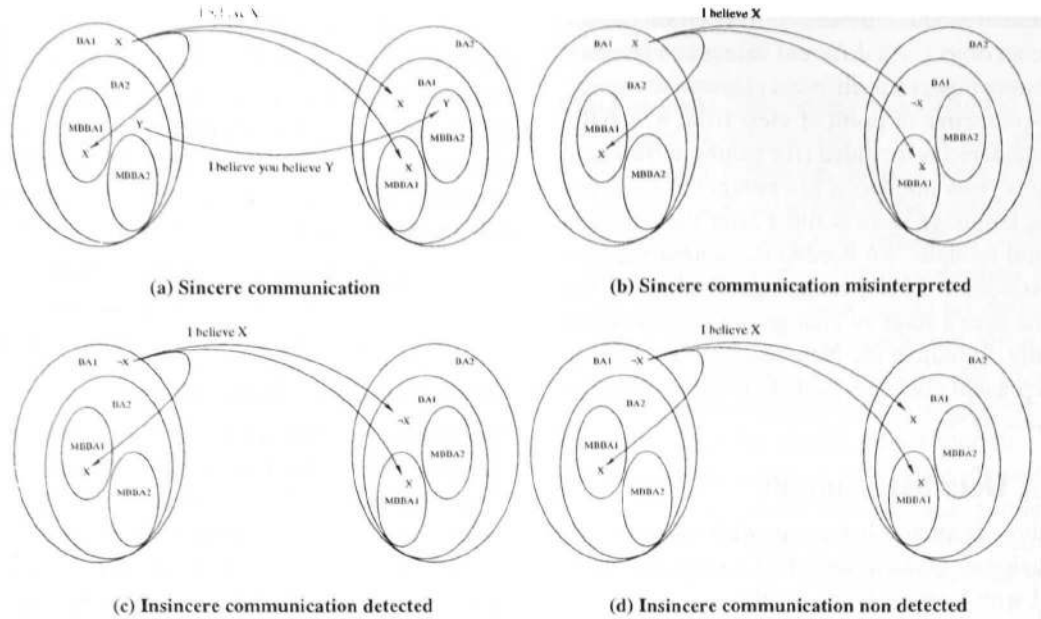

Figure 3: Communication types

At the tactician level, an alternative response to the member/set relation used by the user could have been another member/set (rather than a Dissociation), that is finding another set of which EatingFruit is member that has a good value from the Emotional perspective. Or, the user's supporting argument could have been attacked: an anti-model can be counter argued by a model (that is by finding a model whose habit is EatingFruit) or another anti-model (that is by finding an anti-model whose habit is *not* EatingFruit). The system could also have chosen not to follow the dialogue manager's suggestion, and keep on arguing about EatingFruit, by applying the generic operator to pass a positive value to the action, either from the same emotional perspective or from a brand new one.

Further Developments

After Daphne's prototype was completely implemented in CaML, in a Unix environment, an evaluation stage began, consisting of three phases.

The first phase, just completed, consisted of the collection of *real* dialogues with nutritionists. Five researchers from the Department of Human Nutritions of the University of Glasgow were contacted by e-mail, and asked to engage into a conversation about dietary habits with a "user" (whose responses were constructed by hand on the base of five different "characters"). These dialogues were analyzed in term both of the behavioural models and the argumentative techniques. When asked to comment about the e-mail chat, two of them explicitly stated they had the Stages of Change model in mind to guide the advising process, one used a *motivational interviewing* approach, and the remaining two did not appeal to a particular model, but stated they tried to assess the user's health beliefs. All of them did not just provide knowledge about nutrition, but used the user's responses to establish his/her priorities and give advice from different perspectives, and all their responses could be catalogued according to the New Rhetoric's classification of argumentative schematas.

In the second phase, in progress, the situation is the opposite: some "users" were contacted via e-mail and solicited to engage into a conversation about diet. Daphne's responses are generated and then sent by hand to the user, again via e-mail. This phase will hopefully assess the appropriateness of Daphne's responses, and will help collecting a corpus of "real" arguments to augment system's expertise.

The third phase has just began, and will consist of preparing a simplified Web version of Daphne's prototype, in order to let a wider population to have access to it via a Web site. A more accurate evaluation of the argument produced will then be possible, together with the assessment of whether this style of health promotion intervention is actually acceptable/useful.

Conclusions

The aim of this system was to generate effective advice on a controversial subject, involving personal attitudes and values, by appealing to a well established theoretical basis to produce more effective advice.

The idea is not entirely novel. The Stages of Change Model has already attracted interest from the artificial intelligence community (Marcu 1996; Reiter *et al.* 1997). Also, many argumentative systems exist, although they are either focused on presenting "valid" arguments (Fox, Krause, & Elvang-Goransson 1993; Karacapidilis 1996; Vreeswijk 1997), or stress the generation of arguments in natural language (Elhadad 1995; Maybury 1993; Reed & Long 1997). Zukerman, Korb, & McConachy (1996) describe a system able to generate *nice* arguments tailored to the addressee, even though validity is again the evaluating measure used, an argument

being 'nice' when it can persuade though having some steps missing in its logic chain. A system very close to ours is PERSUADER, by Sycara (1990): she also maintains that the argument generation process is always guided by argumentation goals, changing the importance a person attaches to things, and argumentation strategies to achieve them. The theoretical basis on which the strategies are defined are, however, not made explicit.

Daphne's strength was to combine both behavioural and argumentative research for its purpose, in a well formalized framework. The first results from the evaluation phase seem to confirm the validity of the framework.

References

Becker, H., ed. 1974. *The Health Belief Model and Personal Health Behavior*. Thorofare, N.J: C.B. Slack.

Carlson, L. 1983. *Dialogue Games : an Approach to Discourse Analysis*. Synthese Language Library. Dordrecht: D. Reidel.

de Kleer, J. 1986. An Assumption-Based Truth Maintenance System. *Artificial Intelligence* 28:127–162.

Department of Health. 1992. The Health of the Nation: a Strategy for Health in England. London: HMSO.

Elhadad, M. 1995. Using Argumentation in Text Generation. *Journal of Pragmatics* 24:189–220.

Forslund, G. 1995. Toward Cooperative Advice-Giving Systems. *IEEE Expert* 10(4):56–62.

Fox, J.; Krause, P.; and Elvang-Goransson, M. 1993. Argumentation as a General Framework for Uncertain Reasoning. In Heckerman, D., and Mamdani, A., eds., *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, 428–434. Morgan Kaufmann Publishers, San Francisco, CA.

Fries, E., and Croyle, R. 1993. Stereotypes Associated with a Low-Fat Diet and their Relevance to Nutrition Education. *Journal of the American Dietetic Association* 93(5):551–555.

Hustadt, U. 1994. A Multi-Modal Logic for Stereotyping. In Goodman, B., ed., *Proceedings of the 4th International Conference on User Modeling*, 87–92. The MITRE Coorporation.

Jokinen, K.; Maybury, M.; Zock, M.; and Zukerman, I., eds. 1996. *Proceedings of the ECAI-96 Workshop on: Gaps and Bridges: New directions in Planning and NLG.*

Karacapidilis, N. 1996. An Argumentation Based Framework for Defeasible and Qualitative Reasoning. In Jokinen et al. (1996), 37–42.

Marcu, D. 1996. The Conceptual and Linguistic Facets of Persuasive Arguments. In Jokinen et al. (1996), 43–46.

Maybury, M. 1993. Communicative Acts for Generating Natural Language Arguments. In *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI93)*, 357–364. AAAI press.

Perelman, C. 1979. *The New Rhetoric and the Humanities: essay on Rhetoric and its application*. Dordrecht: D. Reidel.

Prochaska, J., and DiClemente, C. 1992. Stages of Change in the Modification of Problem Behavior. In Hersen, M.; Eisler, R.; and Miller, P., eds., *Progress in Behavior Modification*, volume 28. Sycamore, IL: Sycamore Publishing Company.

Reed, C., and Long, D. 1997. Content Ordering in the Generation of Persuasive Discourse. In *Proceedings of the 15th International Joint Conference of Artificial Intelligence (IJCAI-97)*.

Reiter, E.; Cawsey, A.; Osman, L.; and Roff, Y. 1997. Knowledge Acquisition for Content Selection. In *Proceedings of the 5th European Workshop on Natural Language Generation*, 117–126.

Roe, L.; Hunt, P.; Bradshaw, H.; and Rayner, M. 1997. *Health Promotion Interventions to Promote Healthy Eating in the General Population: a review*. Health Education Authority - The Cromwell Press.

Sacerdoti, E. 1974. Planning in a Hierarchy of Abstraction Spaces. *Artificial Intelligence* 5:115–135.

Sadalla, E., and Burroughs, J. 1981. Profiles in Eating: Sexy Vegetarians and Other Diet-Based Social Stereotypes. *Psychology Today* 15(10):51–57.

Stefik, M. 1981. Planning and Meta-Planning (MOLGEN: Part 2). *Artificial Intelligence* 16:141–170.

Sycara, K. 1990. Persuasive Argumentation in Negotiation. *Theory and Decision* 28(3):203–242.

Vreeswijk, G. 1997. Abstract Argumentation Systems. *Artificial Intelligence* 90(1-2):225–279.

Warnick, B., and Kline, S. 1992. The New Rhetoric's Argument Schemes: a Rhetorical View of Practical Reasoning. *Argumentation and Advocacy* 29:1–15.

Ziebland, S.; Thorogood, M.; Yudkin, P.; Jones, L.; and Coulter, A. 1998. Lack of Willpower or Lack of Wherewithal? 'Internal' and 'External' Barriers to Changing Diet and Exercise in a Three Year Follow-Up of Participants in a Health Check. *Social Science and Medicine* 46(4-5):461–465.

Zukerman, I.; Korb, K.; and McConachy, R. 1996. Perambulations on the Way to an Architecture for a Nice Argument Generator. In Jokinen et al. (1996), 32–36.