Dessalles, J-L. (2008). A computational model of argumentation in everyday conversation: a problem-centred approach.
In P. Besnard, S. Doutre & A. Hunter (Eds.), *Computational Models of Argument*
Proceedings of COMMA 2008, 128-133. Amsterdam: IOS Press.
http://www.enst.fr/~jld/papiers/pap.conv/Dessalles_07111264.pdf

A Computational Model of Argumentation in Everyday Conversation: A Problem-Centred Approach

Jean-Louis DESSALLES
TELECOM ParisTech
46 rue Barrault - 75013 Paris - France
dessalles@enst.fr

Abstract. Human beings share a common competence for generating relevant arguments. We therefore hypothesize the existence of a cognitive procedure that enables them to determine the content of their arguments. The originality of the present approach is to analyse spontaneous argument generation as a process in which arguments either signal problems or aim at solving previously acknowledged problems.

Keywords. Argumentation, conversation, abduction, cognition.

1. Modelling Spontaneous Argumentation

Human beings are all expert in argument generation. Modelling natural argumentation, as it occurs in spontaneous conversation, is important for two reasons at least. First, it is a source of inspiration for argumentation system design, which can choose to imitate it to some extent. Second, as automatically generated arguments are intended to human hearers or readers, it is important to understand the way the latter generate relevant utterances for each other. This paper offers a new model of spontaneous argument generation.

Argumentation represents the major part of spontaneous speech. Our own measures (Table 1) show a distribution of conversational genres, with figures assessed through a sampling method. The corpus we used is composed of 17 hours of family conversation. Conversation proper, which excludes utilitarian (more or less ritualized) speech, occupies more than 70% of the time, and argumentation amounts to 74% of conversation time.

There are strong constraints on every argumentative move in daily conversation, as ill-formed arguments are promptly regarded as pathological and their author as socially inept [1]. We therefore hypothesize the existence of a Human Argumentative Procedure (HAP). This paper is aimed as an attempt to characterize central aspects of the HAP.

Most current approaches to argumentation involve computations over pre-existing ‘arguments’. The problem is then to find out the best argument among a set of more or less adequate moves. However, conversational topics are so varied that finding *one single* acceptable argument is often quite a difficult task. Finding a model for HAP consists in explaining how arguments are *computed*, not how they are merely selected from memory. Human knowledge is potentially huge and hardly circumscribable. We must suppose that it is content-addressed only. This assumption contrasts with most models, in which memory can be randomly searched or scanned for available arguments, to check consistency or detect cycles.

Table 1: Distribution of conversational genres in a corpus of family conversations

Conversation	60 %	Argumentative discussions	74 %
		Narratives	19 %
		Immediate events	7 %
Conversation (inaudible)	13 %		
Other (child screaming, songs)	5 %		
Utilitarian (mainly negotiation about food offer)	11 %		
Empty	11 %		

Natural arguments are generated using only domain knowledge and common sense knowledge. This knowledge-to-argument (K2A) perspective excludes that arguments be manipulated as such or through relations like ‘attacks’ [2]. In a K2A approach, the very notion of argument emerges from the functioning of the argumentative procedure.

The HAP described here essentially consists in (1) detecting some local incompatibility in the participants’ current beliefs and desires, and (2) attempting to resolve this incompatibility. Several authors have noticed that reasoning and argumentation are governed by conflicting beliefs and desires [3] and that argumentation aims at lowering the internal conflict [4]. Our approach differs mainly in the fact that all computations are supposed to be *local*. For instance, Pasquier *et al.* [4] carry out global operations like summing over all constraining relations. We consider such operations to be unrealistic as human knowledge is content-addressed and cannot be scanned.

Our enterprise is to define a minimal argumentative procedure to *generate* (and not merely select) arguments using a K2A approach. The difficulty is to reconstruct real conversations using only domain knowledge. This problem still constitutes a challenge for A.I. and cognitive modelling.

We first introduce a few basic concepts underlying the model: cognitive conflict, strength, abduction. Then we outline our model of HAP, and illustrate how it works on examples. We conclude by mentioning how the model has been implemented.

2. Cognitive Conflicts and Abduction

The argument generation procedure starts with the detection of a cognitive conflict, and stops when this conflict, or the subsequent ones that may come out during the resolution process, have disappeared, or when no solution can be found. Resolution relies on the central mechanism of abduction.

A *cognitive conflict* is detected whenever a given proposition is assigned two opposite *strengths*. We call *strength* of a state of affairs the intensity with which this state of affairs is believed or wished by participants. Strengths are negative in case of disbelief or avoidance. At each step t of the planning procedure, a function $\nu_t(T)$ is supposed to provide the strength of any proposition T on demand. When the strength of T is neither given nor inferred, $\nu_t(T) = 0$. A cognitive conflict is a situation in which a same proposition T is successively given two opposite strengths: $\nu_t(T) \cdot \nu_{t+1}(T) < 0$. The conflict, noted $(T, n_1) \uparrow (T, n_2)$, has an intensity, given by the product $I = -n_1 n_2$. Note that cognitive conflicts are internal to the agents; they are not supposed to be objective. More important, *cognitive conflicts do not oppose persons*, but representations.

Abduction is central to problem-solving [5], to diagnosis reasoning [6] and more generally to human intelligence [7]. It is also essential to the argumentative procedure proposed here. For the sake of simplicity, causal links are supposed to be explicitly given to the model, and abduction is performed by using causal links backwards. Abduction from E using the causal clause $(C_1 \& C_2 \& \ldots C_n) \rightarrow E$ returns the weakest cause in the clause, *i.e. Argmin* $\nu_t(C_i)$. This is, of course, a gross simplification. Further developments of the model could involve procedures to perform Bayesian abduction or sub-symbolic simulations to make the abduction part more plausible. We distinguish *diagnostic abduction* from *creative abduction*. The former returns only actual (*i.e.* observed) causes, whereas the latter may return any cause from the chosen clause.

3. Resolving Cognitive Conflicts

The argumentative procedure is inherently problem-based: It is launched as soon as the current proposition T creates a conflict (we may consider T as the last event observed or the last input in working memory).

(a) Conflict: Consider the conflict $(T, -n_1) \uparrow (T, n_2)$, with $n_1 > 0$ and $n_2 > 0$. There may be a certain threshold I_0, depending on the context, below which the conflict is ignored. If $I = n_1 n_2 > I_0$, the resolving procedure goes as follows.

(b) Propagation: Perform diagnostic abduction from T (T is unbelieved or unwanted with strength n_1). If successful, it returns an actual cause C_i of T. If $0 < \nu_t(C_i) < n_1$, the cognitive conflicts propagates to its cause: Make $\nu_{t+1}(C_i) = -n_1$, and go through step (b) anew with the cognitive conflict $(C_i, -n_1) \uparrow (C_i, \nu_t(C_i))$. However, if $\nu_t(C_i) \leq 0$, the conflict is solved through negation by suggesting $\neg C_i$.

In the following conversation, adapted from [8], we see how cognitive conflict propagation leads participants to produce arguments.

C- How did you get – I mean how did you find that side of it, because...
A- Marvellous
C- You know some people say that... that driving a car across a ferry is the devil of a job [. . .] *well I'll tell you the sort of thing I've heard, I mean every summer, you see stories of tremendous queues at the...*
D- But they're people who haven't booked

The initial cognitive conflict is about driving a car across the Channel, which is presented as ‘marvellous’ by A and D and ‘the devil of a job’ by C. At some point, C propagates the conflict onto its actual cause: the mention of ‘tremendous queues’. D did not have to wait in these queues, so he propagates the new conflict onto an actual cause for being in such queues: not having booked, which happens to have a negative strength in A and D's case. The conflict thus vanishes, as the strength inherited from 'marvellous' is negative too. We can see how the content of the three arguments ('driving a car across the ferry is the devil of a job', 'you see stories of tremendous queues', 'but they're people who haven't booked') results from conflict detection and propagation.

(c) Reparation: If propagation fails, negate T to form the *counterfactual* $\neg T$ ($\neg T$ is believed or wanted with strength n_1) and perform creative abduction. If successful, it returns a (possible) cause C_i of $\neg T$. If $-n_1 < \nu_t(C_i) < 0$, make $\nu_{t+1}(C_i) = n_1$ and go to step (b) with the cognitive conflict $(\neg C_i, -n_1) \uparrow (\neg C_i, -\nu_t(C_i))$. If $\nu_t(C_i) \geq 0$, suggest C_i; if C_i is an action and is feasible, simulate its execution by making its consequences actual and reset $\nu_{t+1}(C_i)$ to 0; then observe the resulting situation and restart the whole procedure.

Consider the following conversation (original in French). R, S and their friends want to project slides on a white door, as they have no screen.

[The projector would be ideally placed on the shelves, but it is unstable]
R- Can't you put the projector there [on the desk]*?*
S- […] *it will project on the handle. That will be nice!*
R- Put books underneath. But can't you tilt it?
S- It will distort the image

R initial suggestion (put the projector on the desk) is motivated by the instability of the projector, which creates a cognitive conflict. The conflict propagates to its cause. Then reparation occurs: the problematic term (projector on shelves) is negated, and an action is found that realises this counterfactual: remove the projector from the shelves. The procedure goes on, with an alternation of conflict detection, propagation and reparation (Table 2).

Table 2. Covert and overt argumentative moves

Argumentative move	**Procedure phase**
projector unstable	*Conflict detection*
projector on the shelves	*Propagation*
remove the projector from the shelves	*Repair*
image no longer on the door	*Conflict detection*
"Can't you put the projector there?" [on the desk]	*Repair*
"I'll project on the handle."	*Conflict detection*
the projector is horizontal	*Propagation*
"Put books underneath."	*Repair*
"But can't you tilt it?"	*Repair*
"It will distort the image."	*Conflict detection*

(d) Failure: When reparation fails, make $\nu_{t+1}(T) = n_1$ (T is thus marked as resisting resolution with strength n_1) and redo the whole procedure.

At the end of the preceding dialogue, the strength n_1 of the image distortion (D) is inherited from the strength of tilting the projector (through (c)), which is itself inherited from the strength of not having the image projected on the handle. When D is observed, it conflicts with the desire n_2 of having a non-distorted image. The conflict reads: $(D, n_1) \uparrow (D, -n_2)$. If $n_1 > n_2$ and there is no actual cause weaker than n_2 for D, propagation fails. If there is no way to produce $\neg D$, reparation fails as well. One proceeds through step (d) and D is stored with the new strength n_1. This leaves the situation with the unresolved conflict $(D, -n_1) \uparrow (D, n_1)$.

(e) Giving up: The system exits the resolution procedure when it is blocked and the strength value hierarchy does not change.

Figure 1 summarizes the whole argumentation generation procedure.

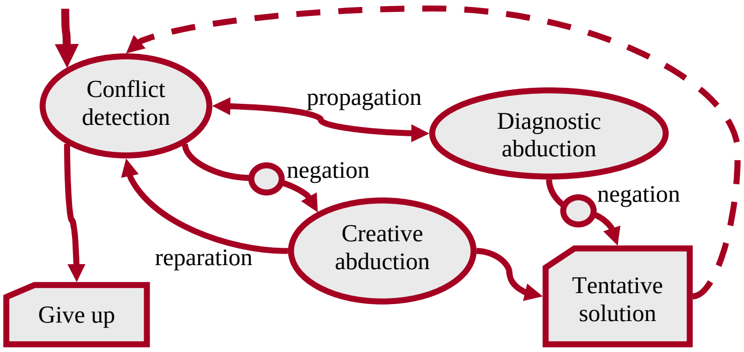

Figure 1. The argument generation procedure

4. Conclusion

We implemented the model in a Prolog programme. For such an implementation to remain plausible, the size of the programme must be kept minimal. The above procedure is realized with less than 130 Prolog goals (15 clauses), excluding display, domain knowledge and state transitions when an action is performed. This amount is five times less than previous attempts [9]. The programme is able to produce the same arguments as those observed in a variety of real dialogues, using a small number of steps. This performance should not be underestimated. Usually, planning procedures consider many useless possibilities, and unlike humans, base their choice on multiple evaluations. The challenge of the approach is not only to produce the correct argumentation, but also to produce it in a reasonable number of steps and with a minimally complex procedure.

The current model of HAP may still be improved. For instance, 'negation' and 'abduction' are still called twice in the procedure. We may think of an even simpler version of the procedure, but it is still to be discovered.

One important result suggested by the model is that human argumentation can be achieved without central embedding. Though argumentative dialogues most often end up as balanced trees of arguments, the procedure that generates them is only rightrecursive. The reason is that the procedure constructs trees of arguments by exploring the web of constraints in a non-deterministic and often redundant way.

The model is conceived to offer a tentative plausible image of cognitive processes underlying argument generation. It does not aim at technical efficiency. If used to process specific task-driven dialogues, it could prove as inefficient as would be a novice in comparison with an expert. However, the model may prove technically helpful when limited knowledge is available to utter arguments that will nevertheless appear relevant. It may be also useful to understand the relevance of users' arguments.

The pursued objective is scientific rather than technical. We consider our approach as a promising step toward better understanding of human spontaneous argumentation. The current limitations of the model are due to the extreme simplification of the knowledge made available to the system, which consists of explicit causal rules. The good side of it is that the argument generation procedure is general, simple and systematic, and offers a plausible, though still partial, image of the human spontaneous argumentative ability.

References

[1] Polanyi, L. (1979). So What's the point?. *Semiotica*, *25* (3), 207-241.

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

[3] Bratman, M. E. (1990). What is intention? In P. R. Cohen, J. L. Morgan & M. E. Pollack (Eds.), *Intentions in communication*, 15-32. Cambridge, MA: MIT press.

[4] Pasquier, P., Rahwan, I., Dignum, F. P. M. & Sonenberg, L. (2006). Argumentation and persuasion in the cognitive coherence theory. In P. Dunne & T. Bench-Capon (Eds.), *First international conference on Computational Models of Argumentation (COMMA)*, 223-234. IOS Press.

[5] Magnani, L. (2001). Abduction, reason and science - Processes of discovery and explanation. New York: Kluwer Academic.

[6] Reiter, R. (1987). A theory of diagnosis from first principles. *Artificial Intelligence*, *32*, 57-95.

[7] Hobbs, J. R., Stickel, M. & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, *63* (1-2), 69-142.

[8] Crystal, D. & Davy, D. (1975). *Advanced Conversational English*. Londres: Longman.

[9] Dessalles, J-L. (1990). The simulation of conversations. In T. Kohonen & F. Fogelman-Soulié (Eds.), *Proceedings of the Cognitiva-90 Symposium*, 483-492. Amsterdam: North Holland. http://www.enst.fr/~jld/papiers/pap.conv/Dessalles_90072003.html