Exploiting Uncertainty and Incomplete Knowledge in Deceptive Argumentation

Valeria Carofiglio and Fiorella de Rosis

Department of Informatics, University of Bari, Italy
{carofiglio, derosis}@di.uniba.it
WWW home page: *http://aos2.uniba.it:8080/IntInt.html*

1 Introduction

Argumentation is not always sincere. This is evident in competitive domains like politics or trading but occurs, as well, in domains in which debates are commonly considered to be governed by 'purely rational', non malicious goals and forms of reasoning, like science [3–5, 7]. When people argue for or against a claim, the differences of arguments they employ are due, in part, to differences in the data they know or in the importance they attach to each argument. However, these differences may originate, as well, from the use of arguments that are not fully sincere: it is therefore worth reflecting on whether and how the various forms of deception that may be employed in argumentation might be formalised.
In this paper, we distinguish between 'uncertain' and 'incomplete' knowledge and we prove how both of them might be exploited by the arguing agent to introduce several forms of deception in argumentation, with a different balance between 'impact' and 'safety' of the deception attempt.

2 Argumentation forms

Let us start from the classical Toulmin's example that is described in http://www.mtsu.edu/mkrueger/toulmin.html): see Table 1. Argumentation may employ purely logical arguments, as in Table 2. However, in the majority of cases (as in Table 1), uncertainty is introduced in the qualifier (and therefore the warrant), the backing of warrant or the influence of the rebuttal.

Various systems have been proposed and prototyped, to show how logical and uncertain argumentation may be formalised: belief networks, in particular, proved to be a powerful tool to model uncertain argumentation, as they enable representing the various warrants that concur to supporting some claim from different data, each with its weight and by appropriately considering the dependency among the data when they exist [10]. Once a model has been built, it may be employed to reason about the effect of evidence available on the claim or, in a sort of 'hypothetical' reasoning, to assess which data combinations produce a desired impact on the claim. For instance, in BIAS, arguments and rebuttals are generated from two belief networks, representing the arguing Agent's first and second order beliefs in the domain [7].

Since Russia has violated 50 of 52 intl agreements.	DATA
Therefore, probably	QUALIFIER
Russia would violate the proposed ban on nuclear weapon testing,	CLAIM
Since	
Past violations are symptomatic of probable future violations	WARRANT
Unless	
The ban on nuclear weapons testing	REBUTTAL
is significantly different from the violated agreements.	
Expert X states that nations that have been chronic violators	BAKING OF WARRANT
nearly always continue such acts.	

Table 1. An example from Stephen Toulmin: The uses of Arguments [11].

DATA	Russia has violated *all* intl agreements in the past.
QUALIFIER	Therefore, certainly,
CLAIM	Russia would violate the proposed ban on nuclear weapon testing,
REBUTTAL	The ban on nuclear weapons testing is significantly different from
	the violated agreements.
WARRANT	Past violations are always symptomatic of future violations
BACKING OF WARRANT	Expert Y states that nations that have been chronic violators
	always continue such acts.

Table 2. Logical argumentation.

The limit of belief networks is, though, in the amount of knowledge required to model the reasoning process: the probability of data have to be known a priori, as well as the links between data and claim, that must be expressed in terms of conditional probability distributions: these parameters may be evaluated subjectively or learned from a training set. The problem of what to do when this knowledge is incomplete, that is, of how argumentation in conditions of 'incomplete knowledge' may be modelled, is still open. We try to imagine, in Table 3, how Toulmin's example would be modified, in this case. The three examined cases (argumentation in a situation of logical, uncertain and incomplete knowledge) are summarised in Table 4: in this table, we employ the following ***Notations:***

- *x denotes a nation*
- *y denotes a member of the set of intl agreements Y*
- *R denotes 'Russia'*
- *N denotes 'the proposed ban on nuclear weapon testing'*
- *$D(x, y)$: "Nation x violated intl agreement y in the past"*
- *$\forall y D(R, y)$: "Russia violated all intl agreements y in the past"*
- *$C(x, y)$: "Nation x will violate the intl agreement y in the future"*
- *$C(R, N)$: "Russia will violate the proposed ban on nuclear weapon testing N in the future"*
- *$W(x, y)$: "Past violations of intl agreements by any nation x imply future violations of intl agreements y in the set Y": $\forall x \forall y (D(x, y) \rightarrow C(x, y))$*
- *$B(e, x, y)$: "Expert e says that past violations of intl agreements by any nation x imply future violations of intl agreements y": $Say(E, W(x, y))$*
- *$R(y, N)$: "The ban on nuclear weapon testing N is significantly different from the violated agreements":$(N \notin Y)$*
- *All the values $(v_1, v_2, ...v_h, ...v_k, ...)$ denote numbers in the (0, 1) interval.*

In uncertain argumentation (third column), the warrant's qualifier provides a measure of the impact that a considered data produces on the claim; an argument

DATA	We don't know exactly how *frequently* Russia violated intl agreements.
CLAIM	It *is rather credible* that Russia would violate the proposed ban on nuclear weapon testing, but it is also plausible that it will not. We can only make an interval estimate for this event.
WARRANT	*To some extent*, past violations *are* symptomatic of future violations. However, they may be not. There is a number of cases in which we don't know which was the relationship between the two events.
BACKING OF WARRANT	Expert Z states that a large proportion of the nations that had such a record of violations continued such actions.However, there are also nations that did not. And, unfortunately, there are, as well, several cases in which Z can't say whether violations occurred or not.

Table 3. Argumentation in a situation of incomplete knowledge.

is said to be "good" if it induces a desired probability value on the claim. This desired value may be obtained by focusing on a data that influences the claim in the desired direction or by introducing some rebuttal that, in a way, strenghtens or reduces the impact of the considered warrant. It should be noticed that warrants, as well as their backings, are not necessarily unique. For instance, several experts might exist, with different viewpoints, that attach different qualifiers to the same warrant; in this case, some criteria has to be applied for deciding which of them should be evoked to achieve the desired effect on the claim. As we said, bayesian reasoning may be applied when all probability parameters attached to the different components of the current argument are known: $P(C(x,y)|D(x,y))$ and $P(C(x,y)|(D(x,y))$. When only incomplete knowledge about these items is available, instead, an interval value for the probability of data and claims may be estimated (fourth column of Table 3). This means that, when an argument is evaluated, a lower and an upper bound for the probability of the claim may be calculated; the convenience of the warrant is linked to the width of this interval and to its lower value: the more the extremes of the interval are near to the probability of the claim the Arguing agent wants to achieve, the more the argument will be effective; the more the interval is large, the more doubt will be left, by the argument, in the Interlocutor's mind.As a consequence, 'uncertain' and 'incomplete-knowledge' argumentation strategies differ, in our opinion, in the way the warrant and its backing are examined, when selecting an 'appropriate' argument for a claim: in uncertain argumentation, to obtain a desired probability value for the claim, the strength of the link between data and claim (the warrant's qualifier) has to be considered.

In incomplete-knowledge argumentation, the type and level of ignorance about this qualifier and its effect on the width of the interval for the probability of the claim have to be considered instead. This opens the possibility of a wide range of enforceable (and, as we will see, even deceptive) argumentation forms.

LOGICAL ARG.	UNCERTAIN ARG.	INCOMPLETE ARG.
	CLAIM $C(x, y)$	
Goal: $C(R, N)$	**Goal**: $Prob(C(R, N)) = v_g$	**Goal**: $Bel(C(R, N)) = v_1$ $Plau(C(R, N)) = v_2$
	DATA $D(x, y)$	
$\forall y D(R, y)$	$\forall y Prob(D(R, y)) = v_1$	$\forall y Prob(D(R, y)) \in [v_3, v_4]$ with $[v_3, v_4] \subseteq [0, 1]$ or $Bel(D(R, y)) = v_3$ and $Plau(D(R, y)) = v_4$
	ON WARRANT $W(x, y)$	
$W^l(x, y)$: $\forall x, \forall y$ $((D(x, y) \rightarrow C(x, y)))$ (which , together with $\forall y D(R, y, inducesC(R, y))$	$W^u(x, y)$: $\forall x, \forall y$ $Prob(C(x, y)\|D(x, y)) = v_h$ $Prob(C(x, y)\|D(x, y)) = v_h$ This together with $Prob(D(R, y)) = v_1$ induces $Prob(C(R, y)) = v_2$ with $v_2 = v_g$	$W^i(x, y)$: $\forall x, \forall y$ $Prob(D(x, y) \rightarrow C(x, y)) = v_h$ $Prob(D(x, y) \rightarrow \neg C(x, y)) = v_k$ $Prob(D(x, y) \rightarrow ?C(x, y)) = v_m$ This together with $Prob(D(R, y)) \in [v_3, v_4]$, induces $Bel(C(R, y)) = v_1$, and $Plau(C(R, y)) = v_2$
	ON REBUTTAL $R(y, N)$ with $N \notin Y$	
$\neg Instance - of(y, N)$ (which induces $W^l(x, N) = unknown$	$Prob(C(x, y)\|D(x, y), R(y, N)) = v_h$ $Prob(C(x, y)\|\neg D(x, y), R(y, N)) = v_k$ etc for the other combinations; this, together with $Prob(D(R, y)) = v_1$ and $Prob(R(y, N)) = v_i$, induces $Prob(C(R, N)) = v_j = v_g$	
	ON BACKING OF WARRANT $B(E, x, y)$	
$\exists e Say(e, W^l(x, y)))$ which together with $Believable(e)$, induces $W^l(x, y)$	$\exists e Say(e, W^u(x, y)))$ which together with $Believable(e)$, induces $W^u(x, y)$	$\exists e Say(e, W^i(x, y)))$ which together with $Believable(e)$, induces $W^i(x, y)$

Table 4. A summary formalisation of logical, uncertain and incomplete-knowledge argumentation forms.

3 Belief and plausibility measures in Dempster&Shafer Theory

To show how Dempster & Shafer's (D&S) theory may be employed in modeling incomplete-knowledge argumentation, we first briefly introduce the main concepts behind this theory (for more details, see [1, 9]). We will denote with D a generic data and with C a generic claim connected to D.
Definitions: let $T = \{t_1, ..., t_j, ..., t_n\}$ be the set of mutually exclusive and exhaustive hypotheses we wish to test (the *frame of discernment*, in Dempster&Shafer' s terms); in our scenario, as we deal with boolean variables, $T = \{C, \neg C\}$; let Θ be the set of all subsets of T; again, in our scenario: $\Theta = \{C, \neg C, ?C\}$, where $?C$ denotes '$Cor\neg C$'. Let $S = \{s_1, ..., s_i, ..., s_m\}$ be a set of possible answers to 'a question related to T'; in our case, for instance:

- s_1: $D \rightarrow C$ denotes the proposition: 'An evidence about the data D implies an uncertain evidence about the claim C'
- s_2: $D \rightarrow \neg C$ denotes the proposition: 'An evidence about the data D implies an uncertain evidence about the claim $\neg C$'
- s_3: $D \rightarrow ?C$ denotes the proposition: 'An evidence about the data D implies an uncertain evidence about the claim C'

We write '$s_i C t_j$' to denote that s_i and t_j are 'compatible'.
We start from a probability distribution P over S, defined according to some subjective or objective evaluation method; in the example: $P(s_1)$, $P(s_2)$ and $P(s_3)$. The meaning of these probability values may be seen as follows: "*As far as I know, data D implies claim C, with a probability $P(s_1)$; it implies $\neg C$ with a probability $P(s_2)$; in the rest of the cases, I cannot say anything about the relation between D and C* ".
The distribution of probability on the elements of S enables us to compute a belief function $Bel(\Theta_h)$ on subsets of T, as follows:

- $Bel(\Theta_h)$=P$\{s_i \mid if\ s_i C t_j\ then\ t_j\ is\ in\ \Theta_h\ \}$.

$Bel(\Theta_h)$ is a measure of the belief we commit to Θ_h, based on P. This belief function assigns, in particular, a belief value to all elements t_j of our set of hypotheses T. It enables, as well, computing the Dempster&Shafer's plausibility of these elements as follows: $Plau(t_j) = 1 - Bel(\neg t_j)$. This upper bound is a measure of how much belief we commit to $T\ \Theta_h$.
The range of variation of the probability we may attach to t_j varies in the interval $[Bel(t_j), Plau(t_j)]$; the width of this interval is a measure of the 'level of doubt' in our beliefs. As a consequence, if $Bel(t_j) = Plau(t_j) = c$ we may say that we believe t_j to a degree c and $\neg t_j$ to a degree $(1-c)$, that our degree of uncertainty about t_j is equal to $Bel(t_j)$while our doubt about this uncertainty value is null.
D&S's belief combination rule may be applied to combine incomplete knowledge from different information sources, as well; that is, in argumentation, to estimate the degree of belief of elements in T when several warrents on the same claim exist.

4 How to deceive?

Deceptive argumentation differs from sincere argumentation in the assumption that A may cite facts that do not correspond to his own beliefs. This entails, to A, the risk of being discovered, by I, in a deception attempt, that he needs to avoid. Therefore, in deceptional argumentation, the convenience of an argument depends, at the same time, on its *impact* on the claim and on its *safety* (probability of not being discovered by I in a deception attempt). Deception may be applied to a combination of data, warrant, its backing and rebuttal. The selection of the deception form depends of what A considers to be the most 'convenient' means to achieve his goal about the claim.
In another paper, we analyse the forms of deception that may be applied when

uncertainty is measured in bayesian terms [6].In this short paper, we will focus our discussion on a comparison between bayesian and D&S reasoning. We wish to prove, on one side, that uncertainty may serve to achieve the (possibly deceptive) argumentation goal in several ways and, on the other side, that, by deceptively simulating incomplete knowledge, the safety of the deception attempt may be increased, though the impact of the argument is reduced.

4.1 Deception in uncertain reasoning

The example in Table 1 might be formalised by assigning the following parameters to the warrant:

Example 1:

$Prob(C(x,y)|D(x,y)) = .9999;$

$Prob(C(x,y)|D(x,y)) = .01;$

In this case, if $P(D(R,y)) = .96$, then $Prob(C(R,y)) = .96$

If, instead, $D(R,y)$ is false, the probability of $C(R,y)$ lowers to .01

Let us now assume that A believes that I is uncertain about the claim (that is, that $Prob(C(R,y)) = .5$ in I's mind). Let us suppose, as well, that A believes that this claim is true, but that his goal is to keep I in doubt about it: that is, he wants to obtain, through his argumentation, that the probability that I assigns to the claim remains invaried ($= .5$).

Deception form 1: manipulation of uncertainty parameters.
If A reasons in bayesian terms, he may deceptively operate on the uncertainty of the warrant, by pretending, in his conversation with I, that the conditional probability values are such, that the data have a very low impact on the claim:

For instance, A might say:

"There is no relation between past and future violations of intl agreements"

$(Prob(C(x,y)|D(x,y)) = Prob(C(x,y)|(D(x,y)) = .5);$

Alternatively, he might operate directly on uncertainty of the data:

"Russia violated about half of the agreements, in the past" $(P(D(R,y)) = .5)$.

In both cases, he will obtain that $P(C(R,y)) = .5$. Notice that, due to the 'extreme' values of conditional probabilities, example 1 is not much different from a case of 'purely logical' reasoning: deception form 1 may therefore be considered as a form of deception that consists in introducing uncertainty in logical argumentation, with the advantage of having the possibility to deceive without lying.

Deception form 2: introduction of a rebuttal.
A might evoke, as well, a rebuttal $R(y,N)$, by manipulating the conditional probability table so that the impact of the combination of $D(R,y)$ and $R(y,N)$ on $C(R,y)$ is close to 0.

A might, for instance, say:

"Nothing may be said about the probability that a nation will violate a particular future

agreement N, if it violated the previous intl agreements but N is significantly different from these agreements; on the contrary, if N is similar to them, it is very likely that this nation will violate it as well; finally, the probability of a violation is very low, when agreements have not been violated in the past".

The previous sentence equates to setting the following conditional probability distribution for the mentioned events:

$Prob(C(x,N)|D(x,y), R(y,N)) = .5$

$Prob(C(x,N)|D(x,y), \neg R(y,N)) = .9999$

$Prob(C(x,N)|(D(x,y), R(y,N)) = .01$

$Prob(C(x,N)|(D(x,y), \neg R(y,N)) = .01$

Notice that this table is compatible with the values in Example 1. Notice also that, in this case, if $R(y,N)$ and $D(x,y)$ are both true, their combined effect is that the probability of $C(x,N)$ is still equal to .5. Therefore, the declaration that "*It is true that Russia violated intl agreements in the past; however, the ban on nuclear weapon testing is significantly different from the violated agreements*" might help A to leave I in a state of uncertainty about $C(x,N)$, while accepting the data $D(x,y)$ as true.

4.2 Deception in incomplete-knowledge reasoning

We now consider the case in which only incomplete knowledge about the relationship between data and claim is available.

Example 2: backing of warrant by Expert E1

Expert E1 says the following:

"*In 80 % of cases, it was observed that, when a nation was known to systematically violate intl agreements in the past, it continued doing the same*": $P_1(D(x,y) \rightarrow C(x,y)) = .8$

"*It almost never happened that a nation which was known to have systematically violated intl agreements in the past stopped doing such action* ": $P_1(D(x,y) \rightarrow (C(x,y)) = .01$

"*There is a 20 % of cases in which the relation between past and future violations is unknown*": $P_1(D(x,y) \rightarrow ?C(x,y)) = .19$, where $?C(x,y) = C(x,y) or \neg C(x,y)$.

Notice that the difference between Examples 1 and 2 is that, in the second case, knowledge about the effect of the data on the claim is incomplete. According to D&S's theory, if $D(x,y)$ is true, the probability of $C(x,y)$ varies in the (belief, plausibility) interval $(.8, 1)$, while the probability of $\neg C(x,y)$ varies in the interval $(0, .2)$. Notice that the belief and the plausibility of $C(x,y)$ are both high: the width of the interval between the two values is not very large, due to the presence of a single backing of warrant and of a not too incomplete knowledge. Translation into argument: "*Let us suppose that Russia violated all intl agreements in the past. The degree of belief in the fact that Russia will violate future agreements, based on the present state of knowledge, is equal to .8. However, the plausibility of this fact is equal 1: this means that the degree of belief of this claim might go up to 1, should all information that is not available at present be so as to confirm the warrant*".

Deception form 3: simulation of incomplete knowledge.

Let us suppose that A has a complete, although uncertain, knowledge about the domain in question: that is, he may estimate, either subjectively or objectively, all uncertainty parameters, as in Example 1. If we compare Example 2 with Example 1, we may notice that A might *simulate* an incomplete knowledge to increase, in I's mind, the level of doubt about the claim: he might indirectly deceive I about the claim by deceiving her, in fact, about his own level of knowledge about the warrant. In fact, if $D(x,y)$ is true, $P(C(x,y)) = .99$ in the first case, while, in the second case, it lies in the interval $(.8, 1)$.

Deception form 4: introduction of an 'uninformed' information source. Let us now assume that A introduces some different backing about the same warrant, for instance by citing the opinion of an Expert E2 whose knowledge is different and more incomplete than knowledge of E1. In this example, the probabilities of evidences might be settled as follows: the relationship between data and claim is available.

Example 3

Expert E2 says the following:

"*In 30 % of the cases, when a nation was known to systematically violate intl agreements, it continued doing the same*": $Say(E2, P_2(D(x,y) \rightarrow C(x,y)) = .3)$.

"*It never happened that a nation which was known to have systematically violated intl agreements in the past stopped doing such action* ": $Say(E2, P_2(D(x,y) \rightarrow \neg C(x,y)) = 0)$.

"*There is a 70 % of cases in which the relation between past and future violations is unknown*": $Say(E2, P_2(D(x,y) \rightarrow ?C(x,y)) = .7)$.

This knowledge leads to a $(.3, 1)$ interval estimate for the probability of the claim. We already saw how the first interval estimate could be interpreted. In a similar way, the interpretation of this second interval is the following: "*If Expert E2 was the only information source, in the present state of knowledge the claim could not be believed, as this information has a low degree of belief*". Notice that in this case, while the degree of belief in the claim is low, the width of the interval is very large, due to the high degree of incompleteness in knowledge of E2. Evoking this expert rather than expert E1 might be employed deceptively, by A, to induce a high level of doubt in I's mind. Notice also that, as the plausibility of the claim is very high, the risk, to A, of being discovered, by I, in a deception attempt is rather controlled: if discovered, A might always say: "*Sorry, I didn't know!*",or "*Sorry, E2 said it!*": this way, he will appear less guilty than if he had to confess a lie.

Deception form 5: introduction of a 'confounding' backing.
As a last example, let us now look at how the two separate backings of warrants in examples 2 and 3 may be integrated with D&S's rule of combination of evidences and how this integration may be exploited deceitfully. According to this rule, if both E1 and E2 are cited and $D(x,y)$ is true, the range of variation for the degree of belief in $C(x,y)$ varies in the interval $(.63, .1)$: we will not go into mathematical details. Notice that, in this case, the degree of belief is lower than in the case in which only knowledge of E1 was available, while nothing changed about the plausibility. The effect of introducing, in the argumentation,

expert E2's viewpoint is that the range of doubt about the claim (the width of the belief-plausibility interval) increased. This example shows that, to insinuate, in the Interlocutor's mind, a doubt about the claim, the arguing Agent may deceptively introduce the viewpoint of another Expert: even in the presence of a relatively informed previous opinion, if the second expert's knowledge is limited or wrong, the final effect will be to 'confound' the interlocutor. This application of D&S's combination rule appears to simulate in a rather effective way some particular types of deceptive argumentation, in which confounding arguments are introduced to divert the interlocutor's attention from the truth, to insinuate doubt in her mind or to increase her level of ignorance.

5 Final Considerations

This short paper is a first contribution to a reflection on how uncertain and incomplete knowledge might be considered in simulating deceptive argumentation. We may conclude that logical, uncertain and incomplete-knowledge reasoning define distinct scenarios: in these scenarios, deceptional arguments decrease of impact on the desired goal about the claim, with the advantage of reducing, at the same time, the negative consequences of being discovered, by the interlocutor, in a deceitful attempt (or, at least, with the advantage of keeping some opportunity of defending oneself). Suspicion in humans is, with no doubt, inherently uncertain: in human communities, it is very uncommon to be absolutely confident in other people's claims; so, as deception implies placing oneself on the interlocutor's viewpoint to reason about the effect of a deceiving attempt, it is reasonable to assume that deception should be simulated in conditions of uncertainty. Situations of 'incomplete knowledge' are very frequent as well, both among humans and in MultiAgent systems, consistently with the idea of an 'open world' with a limited possibility of observing the other Agents' behaviour and of situations whose occurrence and whose relationship with the rest of the world are known only in part. The interest of D&S's theory is that it enables to come to a decision, in these situations of incomplete knowledge. The possibility to induce an interval estimate of the degree of belief in the claim by simulating or exaggerating his ignorance leaves, to the arguing agent, the opportunity of considering, in his decision, what he presumes to be the interlocutor's "personality". 'Pessimistic' or 'optimistic' interlocutors will look at the same (belief, plausibility) interval with different viewpoints:"*I will not get convinced about this claim until I'll have more evidence*" vs "*I accept it until there is some couterevidence*". One might argue that D&S theory does not correspond to a realistic way of modeling human reasoning: this is certainly true. As Tversky and Kahneman proved in their seminal paper, humans tend to apply ad hoc heuristics rather than probability theory, in taking their decisions [12]. A still more curious finding is reported by Barwise, who found that even mathematicians do not necessarily reason logically about mathematical question issues such as 'There are quite a few prime numbers' [2]. So, it's absolutely unlikely that humans apply, in incomplete knowledge reasoning, D&S's theory. However, this type of objection relates to the very long debate on whether computerised decision-making should try to emulate human reasoning as it is or should apply its own reasoning style, to try to come to a 'correct' decision. We tend to support the second solution and tend to prefer mathematically correct and well grounded theories to 'ad hoc' heuristics, like, for instance, those that have been proposed, for years, to handle uncertainty in 'expert systems' ad have now been abandoned: but this is and will be, of course, a matter of debate for still longtime.

Acknowledgements
We are indebted with Cristiano Castelfranchi for involving us in his long-lasting interest towards deception, for his high-value suggestions and critiques to our work and, in particular, for his hints about the role of uncertainty in this, intriguing indeed, form of reasoning.

References

1. Barnett, J. A.: "Calculating Dempster-Shafer Plausibility". IEEE Transactions on Pattern Analysis and Machine Intelligence. 13, 6, 1991.
2. Barwise, J: "Monotone quantifiers and admissible sets". In J E Fenstad, R O Gandy and G E Sacks: Generalized Recursion Theory. North Holland Publ Co, 1978.
3. Buller,D. B., Burgoon, J. K., Buslig, A., and Roiger, J.: "Testing interpersonal deception theory; the language of interpersonal deception". Communication theory, 6, 3, 1996.
4. Castelfranchi, C. and Poggi, I.: Lying as pretending to give information. Pretending to Communicate, H. Parret (Ed), Springer Verlag, 1993.
5. Castelfranchi, C., Poggi, I.: Bugie, finzioni e sotterfugi. Carocci Publ Co, 1998.
6. de Rosis, F., Castelfranchi, C., Carofiglio, V.: Can computers deliberately deceive? A simulation attempt of Turing's Imitation Game. Sumbitted for publication.
7. Jitnah, N., Zukerman, I., McConachy, R. and George, S.,(2000): Towards the Generation of Rebuttals in a Bayesian Argumentation System. In INLG'2000 Proceedings – the First International Natural Language Generation Conference, pp. 39-46, Mitzpe Ramon, Israel.
8. Kashy, D. and DePaulo, B.: "Who lies?" Journal of Personality and Social Psychology. 70, 5, 1996.
9. Shafer, G. and Logan, R.: Implementing Dempster's Rule for Hierarchical Evidence. Artificial Intelligence, 33, 1987.
10. Spiegelhalter, D. J.: "Probabilistic Reasoning in Predictive Expert Systems". Uncertainty in Artificial Intelligence. L. N. Kanal and J. F. Lemmer (Eds), Elsevier Science Publishers B. V. (North Holland), vol. 4, 1986.
11. Toulmin, S.: "The use of argument", Cambridge University Press, Cambridge MA, 1958.
12. Tversky, A. and Kahneman, D.: "Causal schemata in judgments under uncertainty". In progress in social psychology, ed. M. Fishbein. Hillsdale, N.J.: Lawrence Erlbaum, 1977.
13. Zlotkin, G. and Rosenschein, J.: "Incomplete information and deception in multiagent negotiation". XII International Joint Conference on Artificial Intelligence, Sydney, 1991.