
Assessing Information Heard on the Radio

Antoni Diller

School of Computer Science, University of Birmingham, B15 2TT, UK

Abstract. Much work in AI is devoted to the problem of how computers can acquire beliefs about the world through perception, but little effort is devoted to the problem of how computers can acquire beliefs through testimony. This paper is part of a continuing project whose ultimate goal is that of constructing an implementable model of how agents acquire knowledge through testimony. In particular, it looks at how agents acquire information from the radio and many factors are identified that may cause an agent to override the defeasible rule to believe what he hears.

1 Introduction

This paper is written as part of the continuing work of the Electric Monk Research Project [3–10]. The Electric Monk is a fictional character, created by the science-fiction humorist Douglas Adams, who ‘believed things for you, thus saving you what was becoming an increasingly onerous task, that of believing all the things the world expected you to believe’ [1, p. 3]. The main goal of this Research Project is to model the human ability to acquire knowledge through testimony in sufficient detail to allow a computer program to be written to emulate this ability. The study of testimony is almost totally ignored in AI and yet AI’s ultimate goal, namely that of constructing an artificial person [2, p. 7], cannot be achieved until we can give androids the ability to acquire beliefs through testimony. As well as having a large amount of in-built knowledge and the ability to acquire beliefs by means of perception, a general-purpose, intelligent and autonomous android would also have to have the ability to extend its knowledge by accepting other people’s assertions. Such an ability is needed if the android is going to be able to interact with human beings in any meaningful way [7,9].

Although we get most of our information about the world from other people by believing what they assert, we do not believe every assertion. I have argued that our way of dealing with assertions is governed by the defeasible rule to believe them [7,9]. This sounds very simple, but it is extremely fruitful because it forces us to look at the reasons why we decide not to accept another person’s word. In other papers I have looked at the factors that cause us to override the defeasible rule to believe other people when we are hearing what another person says to us in the flesh [5], when we are reading a book [6] and when we are reading a journal article [4]. In this short paper I briefly summarise the factors that may cause an agent not to accept the assertions that he hears on the radio. Examples of how these factors work in practice can be found in the long, but unpublished, version of this paper [8].

The ability to evaluate assertions is very different from that of understanding them. My concerns are epistemological and not semantic. Natural-language processing deals with the problem of how language is understood. I am interested in the totally different question of how we decide whether or not to accept an assertion once we have understood it.

Several years ago I introduced a two-stage model of belief-acquisition [3]. Since then the model has been considerably improved. Here I briefly summarise the latest version [7, pp. 4–7]. There are two main ways in which we acquire beliefs directly, namely through perception and through testimony. These constitute the first phase of belief-acquisition. How we acquire knowledge through perception is not my concern here. I have argued that our acceptance or rejection of other people's assertions is governed by the defeasible rule to believe them [9]. This rule, however, is not infallible. Sometimes we acquire beliefs that are in fact false and sometimes we reject assertions that are in fact true. Thus, there is a second phase of belief-acquisition in which we critically examine some of our beliefs in order to weed out the false ones and replace them with better ones. In the second phase we can also re-examine some of the assertions we have rejected in order to check whether or not they are true after all.

The way in which I unpack the defeasible rule to believe other people's assertions is to represent it as an ordered set of rules all of which except the last are conditional ones. The last rule is the non-defeasible rule to believe the assertion in question. I call such a set of rules an *assessment component*. There are many reasons why someone may decide not to accept an encountered assertion and each such reason becomes the antecedent of one of the conditional rules in the assessment component. For example, a radio play is a work of fiction and so we do not normally believe the various assertions that the actors performing it make. This can be captured by adding the following conditional rule to the assessment component, 'If the assertion X is uttered by an actor during the performance of a radio play, then reject X .' (This is a simplification of the actual way in which we treat assertions by actors, since sometimes they are made to say things that are true in the real world.)

So far I have been describing the first stage of belief-acquisition. The evaluations that we carry out concerning the assertions we encounter have to be done in real time. The decision has to be made virtually instantaneously whether or not to accept an assertion. This means that the assessment cannot be very sophisticated. So, people will come to have some false beliefs and they will also reject some assertions which, as a matter of fact, are true. In my model, therefore, there is a second stage of belief-acquisition in which a small number of assertions are subjected to a detailed and possibly time-consuming investigation. A person can either check something that he believes or investigate an assertion that he previously rejected. When someone checks one of his beliefs, he may either decide that he was correct in having that belief in the first place or he may conclude that on this occasion he made a

mistake and reject the belief. Similarly, when someone investigates the truth of an assertion that he rejected previously, he may change his mind about its truth-value and add it to his belief-system or he may decide that he was correct in rejecting it and not alter his assessment of it. When a person does change his mind about the status of an assertion, he will probably need to make further revisions to his belief-system in order to remove any obvious inconsistencies arising out of the change of mind. I do not deal with belief-revision in this paper.

Many different critical methodologies are used in the second stage of belief-acquisition to thoroughly test the accuracy of our beliefs. This is because there is not a single methodology that can be used to check the correctness of every kind of factual assertion. Different methods are used for different kinds of assertion. For example, the way in which someone checks a physical assertion, such as ‘The speed of light in a vacuum is 299,792,458 metres per second’, is very different from the historical methodology that has to be used to decide whether, for example, General Sikorski died in an accident or was murdered and, if he was assassinated, who was responsible. Furthermore, not everybody has the ability to check every kind of belief. A historian would know how to investigate the circumstances of General Sikorski’s death and a physicist would know how to determine the value of the speed of light, but the historian is unlikely to have the knowledge to work out what the speed of light is and the physicist is unlikely to be able to conduct historical research into issues relating to what happened in the past. The topic of how and why people go about checking their beliefs is a vast one, but, unfortunately, I cannot pursue it further here.

2 Assessing Information Heard on the Radio

2.1 Introduction

A person listening to the radio does not accept every assertion he hears. There are many reasons for this. I group these into five categories: (1) We sometimes take *external* factors into account. These relate to such things as the country in which the radio station is based and the time at which we happen to be listening to the radio. (2) Sometimes information about the *assertor* may cause us to be wary of accepting an assertion he makes. He may, for example, be well known to be unreliable. (3) There may be something about the *manner of delivery* of an assertion that makes us hesitate to accept it. (4) The actual *content* of the assertion may make us think that it is not correct and incline us to reject it. (5) There may be something about the *listener* that makes him wary of accepting a particular assertion.

2.2 External Factors

There are features of the context in which an assertion is made, rather than of the assertion itself, that may incline an agent to reject that assertion.

These include the time when the assertion is broadcast, the owner of the radio station which broadcast it, the location of that radio station, the type of programme during which it was uttered and the political situation in the country at the time the assertion was made. Examples of how these factors work can be found in the full version of this paper [8, pp. 9–11].

2.3 Factors Relating to the Assertor

Sometimes there are things that we know about the assertor that influence our assessment of the correctness of what he says. These relate to the role that the assertor plays in the programme concerned, the assertor's ideology, the character of the assertor, any possible agenda that the assertor may have in appearing on the radio, any relevant expertise that the assertor may have (if talking about some specialist topic), the age of the assertor and the assertor's gender. Examples of how these factors work can be found in the full version of this paper [8, pp. 12–13].

2.4 Factors Relating to the Manner of Delivery

All of us learn our first language in a social context and as we learn how to speak we also acquire beliefs through accepting what we are told. In due course, we learn not to trust everything that we are told. We are taught what to look for in an assertion that may indicate that it is not worth accepting. For example, a person's body language may make us think that he is lying to us. Clearly, such features are not available to us when we are listening to the radio. There are, however, features of the way in which someone speaks that may make us think that he is not being completely truthful. Examples of why an agent may choose not to believe what he hears on the radio, because of the way in which the assertion is uttered, can be found in the full version of this paper [8, p. 14].

2.5 Factors Relating to the Content of the Message

There may be features of the content of an assertion that alert us to the possibility that it may be inaccurate. Such features include the consistency of the message, whether it induces any strong emotion in the listener, whether it is about something that is out of the ordinary and whether it is about the sorts of thing that people often lie about or make mistakes about. Examples of how these factors work can be found in the full version of this paper [8, pp. 14–15].

2.6 Factors Relating to the Listener

There are a number of properties of the person listening to the radio that may make him wary of accepting what he hears. These relate to the listener's pre-existing knowledge, to any possible consequences that result from accepting

the assertion being evaluated, to the relevance of the assertion to the listener, to the listener's character, to the maturity of the listener, to the listener's gender and to the situation in which the listener is in when he hears the assertion. Examples of how these factors work can be found in the full version of this paper [8, pp. 16–18].

3 Conclusion

Many problems have to be overcome before scientists succeed in building a humanoid robot with intellectual abilities analogous to those possessed by human beings. Before we can even begin to design such an android we have first to understand the abilities that humans have. In this paper I have been concerned with developing a model of the human ability to evaluate testimony. I have looked at how people acquire information from the radio and I have mentioned many factors that may cause them to override the defeasible rule to believe others. A great deal of work still needs to be done before we can implement a realistic model of belief-acquisition in a computer program. I have, however, done some of the ground-breaking work on this topic. I hope that some people reading this paper will be stimulated to join me in this exciting, but sadly neglected, field of AI research.

References

1. Adams, D. (1988) *Dirk Gently's Holistic Detective Agency*. Pan, London.
2. Charniak, E., McDermott, D. (1985) *An Introduction to Artificial Intelligence*. Addison-Wesley, Reading.
3. Diller, A. (1999) The belief-filter component. *Cognitive Science Research Papers CSRP-99-9*, School of Computer Science, University of Birmingham.
4. Diller, A. (2000) Evaluating information found in journal articles. In Nepomuceno, Á., Quesada, J.F., Salguero, F.J., editors, *Logic, Language and Information: Proceedings of the First Workshop on Logic and Language*, 71–78. Kronos, Sevilla.
5. Diller, A. (2000) Everyday belief-acquisition. In Henning, G.P., editor, *Argentine Symposium on Artificial Intelligence (ASAI2000) Proceedings*, 221–232. SADIO, Buenos Aires.
6. Diller, A. (2001) Acquiring information from books. In Bramer, M., Preece, A., Coenen, F., editors, *Research and Development in Intelligent Systems XVII: Proceedings of ES2000*, 337–348. Springer, London.
7. Diller, A. (2002) A model of assertion evaluation. *Cognitive Science Research Papers CSRP-02-11*, School of Computer Science, University of Birmingham.
8. Diller, A. (2002) *Assessing Information Heard on the Radio*. *Cognitive Science Research Papers CSRP-02-12*, School of Computer Science, University of Birmingham.
9. Diller, A. (2003) Designing androids. *Philosophy Now*, no. 42, 28–31.
10. Diller, A. (2003) Modelling assertion evaluation. *AISB Quarterly*, no. 114, 4.