Evaluating Information Found in Journal Articles

Antoni Diller
School of Computer Science
University of Birmingham
Birmingham
B15 2TT
England
A.R.Diller@cs.bham.ac.uk

Abstract

People need a vast amount of knowledge in order to live in a human society. Most of this has to be obtained from others by believing what they say and write. Androids and sophisticated cognitive systems would also have to be able to learn in this way. Before we can program this ability into them, however, we need to understand human belief-acquisition. Elsewhere I have proposed a two-stage model of belief-acquisition. The first stage consists in the employment of a defeasible rule to believe others. The second stage consists in the use of a sophisticated critical methodology. In this paper I develop one part of this model in more detail. I look at the factors that cause us to override the defeasible rule to believe others in the situation when we are reading an article.

Introduction

Some people working in AI and cognitive science see their ultimate goal as being that of constructing an artificial person (Charniak and McDermott, 1985, p. 7). Those working on the MIT Cog Project are more explicit: 'Building an android, an autonomous robot with humanoid form and human-like abilities, has been both a recurring theme in science fiction and a "Holy Grail" for the Artificial Intelligence community' (Brooks et al., nd). Many of the theoretical and practical problems associated with the task of manufacturing an android are currently being tackled by a large number of researchers, both in academia and in industry, but there is one very important human ability that such an android would have to possess that has largely been overlooked. As well as having a large amount of built-in knowledge and the ability to acquire beliefs by means of observation, the android would also have to have the ability to extend its knowledge by listening to other people and by reading what they have written. Such an ability is essential if the android is going to live in a human society and interact with human beings in any meaningful way. This is because, in order to function in a human society, an agent needs a vast amount of knowledge. This includes what is known as *world knowledge* in AI and *stock* or *commonsense knowledge* in sociology. Although people need to know many practical things, such as how to answer a telephone or how to behave in a restaurant, in this paper I am interested in propositional knowledge. Someone living in an advanced society would

need to know, for example, many things about money and how it works, how goods are bought and sold, where various goods can be bought and so on. A person would also need to know something about the legal and political system of the society in which he lives and something about its social institutions as well. In addition, he would need to know something about the transportation system of the country he lives in. There are also many further things that he would need to know, but I hope that the above gives a flavour of the knowledge required to live in an advanced human society.

The ability to learn from other people may seem obvious, but its importance is usually overlooked by cognitive scientists and epistemologists. For example, Pollock (1987, p. 486) writes, 'Perception represents the basic source of human knowledge. Nonintellectual mechanisms put us into various perceptual states, and being in those perceptual states constitutes a prima facie reason for conclusions about the world around us.' I do not deny that agents do have the ability to make judgments about their surroundings and to acquire beliefs as a result of this, but if they were restricted to acquiring beliefs by making judgments about their immediate surroundings, they would be extremely limited in the beliefs that they could acquire. Their knowledge would be very restricted and would not be sufficient for them to be able to live and function in an advanced society. We need to make use of beliefs that we acquire by listening to other people and by reading what they have written. Anyone who accepts this is forced to consider the human ability to learn from others and this turns out to be far more complicated than may at first sight appear.

An android or artificial cognitive system that was sufficiently advanced to be capable of interacting with human beings and living in a human society would have to have the ability to acquire information from what people say and from written sources. In order to produce an android with this ability, we first need to understand it and that is what I am after. In this paper I make a start on the task of looking in more detail at one aspect of belief-acquisition. I look in detail at the features of assertions found in journal articles that make us wary of accepting them and I isolate many of the factors that cause us to override our default principle to accept what we read. My long-term goal is to formulate these things in sufficient detail so that they can be programmed into an android or artificial cognitive system.

People acquire knowledge from a variety of sources. For example, they get beliefs by listening to other people, by reading books and articles, from the media, from the Internet and so on. This paper is one in a series in which I look in detail at each of these sources. There are enough significant differences between how we evaluate the information coming from these sources for each to be handled separately, though it is inevitable that there should be some overlap. Looking at how we assess information given during a personal communication (Diller, 2000b) or found in a book (Diller, 2000a) should convince the reader of this.

The Two-phase Model

In this section I summarise the two-phase model of belief-acquisition that I introduced elsewhere in a simpler form (Diller, 1999). Since then the model has been considerably refined, extended and improved.

In the first phase we acquire beliefs by reading what other people have written, by

listening to what they say and by making judgments about our surroundings. These processes, however, do not always result in us acquiring true beliefs and so 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.

It is impossible for any person to check every piece of information that he receives and that he has to make use of in his life. This is because it is usually very time-consuming to investigate the truth of an assertion that we hear or read and so a person just does not have the time available to thoroughly test every statement he needs to make use of. In addition, the critical evaluation of a statement itself involves taking for granted very many other statements which themselves have not been thoroughly checked out. It simply is a fact, that some people may regard as unfortunate, that we have to accept most things on trust.

Although most of the information that we need we obtain by trusting others, this does not mean that we have to accept an authoritarian theory of knowledge. It is a truism that even the most respected authorities can make mistakes. Although a person cannot check *every* piece of information that he accepts, he can test *some* of the assertions that he hears and reads.

Most of a person's belief-system has been obtained by believing other people, but this does not mean that that person accepts absolutely everything that he hears or reads. Believing every assertion and theory that a person encounters would very quickly lead him to have a massively inconsistent belief-system. By this I mean that he would have a large number of obviously inconsistent beliefs. My proposal is that the first phase of belief-acquisition involves making use of a defeasible rule to the effect that we accept other people's assertions at face value. There are features, however, concerning the making of an assertion and its content that make us wary of accepting it outright. For example, a person may be very wary of accepting the assertions of a government spin-doctor when these are presented in the context of a press briefing. Any feature that we take into account in the first stage of belief-acquisition has to be, of necessity, easy to recognise. It does not involve thoroughly testing an assertion before it is accepted. Such a factor has to be easy to recognise because we hear and read so many statements every day that we have to decide very quickly whether or not we are going to accept them. To use some computing metaphors, our decisions have to be made in real time and on-line. Because these features of assertions have to be straightforward to recognise, they cannot be very sophisticated. This means that people do end up having quite a few false beliefs and several incorrect pieces of information. This is another reason for holding a two-phase model of belief-acquisition. In the second stage we look more carefully at a small number of our beliefs and thoroughly check them out. In this way we can try to minimise the number of false beliefs that we have about issues that are particularly important to us.

The main research contribution of this paper is the list of factors that may cause us to override the defeasible rule to believe what we hear and read. Here I want to give some examples of how those factors operate. (These factors are contained in the section "Assessing Information Received from Articles" below. The annotations used in the examples, like 'E5' and 'A1', come from that section.) Faced with a statement in a journal article three main options are available to us, namely to accept it and

to add the information contained in the statement to our belief-system, to reject the statement (which does not necessarily involve accepting its negation) or to flag it as something that is worth thinking about further at a later stage. For example, while reading John Pollock's article "Defeasible Reasoning", I came across the assertion, on p. 488, 'Perception and memory provide the starting point for reasoning.' Although this article appeared recently (E5) in a highly respectable journal (E1, E3) by a well-regarded philosopher and cognitive scientist (A1, A2, A3, A4), I did not accept it straightaway when I read it. This is because of my pre-existing knowledge (R2). I may be wrong in believing that most of a person's knowledge comes from other people and only a small amount comes from personal observation, but it would take much argumentation to convince me that I was wrong. So, I rejected Pollock's statement.

As another example, consider the assertion, 'We attempt to impose upon a seemingly chaotic universe an order we find satisfactory and, therefore, reassuring.' This occurs in an article in the journal *Ultimate Reality and Meaning* by Vivelo (1984, p. 133). I assume that readers of my paper are in the same position that I am in, namely they know nothing about this journal (E1) or the author of the article (A1, A2, A3, A4). The title of the journal (E1) may, however, make many people sceptical of its value. In order to decide what to do with this assertion we have to go mainly on its content (C1) and on our pre-existing knowledge (R2).

Having given examples of how we use the defeasible rule to believe what we read, I turn my attention to the second stage of belief-acquisition and give an example of what happens there. We do not have the time to thoroughly check every statement we read or hear, but we do examine critically a small number of statements that are particularly important to us. For example, earlier I quoted Pollock's assertion that the starting-point for belief-formation is perception (Pollock, 1987, p. 486). Many people reading that would simply accept it and add it to their belief-system. However, I think that Pollock is incorrect on this point and above I presented various reasons why I think his statement is false and why I do not believe it. In its place I have various beliefs. For example, I believe that one of the starting-points of belief-formation is perception, but there are others as well. In particular, I believe that we usually get our beliefs by accepting what others say and what they have written.

In the second stage we use some sort of critical methodology in order to thoroughly test the correctness of some of our beliefs. If the beliefs that we are examining belong to a specialised discipline, like physics, mathematics or archaeology, then methodologies specific to those disciplines may have to be employed in order to check the truth or falsity of our beliefs belonging to those disciplines. I say more about different sorts of criticism elsewhere (Diller, 1999, pp. 24–26).

When we come to accept a new belief or reject an old one, that may have a knock-on effect on our pre-existing knowledge. We may have to engage in some form of belief-revision. This is an issue that is the subject of much research, but my concerns in this paper are different.

Assessing Information Received from Articles

Some of the information that a person has he has obtained by reading journal articles. Human beings can be thought of as having a rule to the effect that they should believe everything that they read. This rule is, however, defeasible. We, thus, need to

enquire into the circumstances when, in fact, it is defeated. What factors relating to a statement that we read in a journal article make us wary of accepting it outright? The factors that we take into account can be grouped into four categories. The first of these comprises of a number of *external* features of the journal that is being read and the others relate to the *author* or *authors* of the journal article, the *content* of the statement being assessed and how that content may affect the *reader* of the article. It should be noted that sometimes a number of factors, maybe belonging to different categories, combine together to make us wary of accepting outright an assertion that we read in an article. In other words, the reader should not assume that the factors that follow are all independent of each other.

Before listing the factors that may cause us to override the rule to believe what we read, I just want to say that I am aware that research has been done and may continue to be done on how we react to information found in different journals, but here I am mainly concerned with how a person unaware of such research reacts to such information. Most of us are in this position and only a few of us will go to the trouble of digging out such research, but even if we do this we will still have to evaluate the fruits of the research in some way similar to that which I have described in this paper in outlining the two-phase model of belief-acquisition.

The first group of factors are the *external* ones. (In addition to listing these factors I would have liked to have given further examples of how each of them operates, but the space limitations of this conference prevent me from doing that.)

(E1) We take the kind of journal involved into account. There are very many different kinds of journal and we do not treat them all in the same way.

(E2) We take the kind of article into question, since there are several different kinds of article.

(E3) We may take the publisher of the journal into account. There are, no doubt, differences between people concerning the status of various publishers. Many people in England, for example, think highly of publishers like Oxford University Press and Cambridge University Press.

(E4) We may take the place where the journal was published into account.

(E5) We may take the year of publication into account.

The second group of factors that we take into account are those relating to the *author* or *authors* of the article. Articles can be written by one or more people. To make the following discussion easier to follow I shall assume that we are dealing with an article written by a single author. With suitable changes, the discussion can be made to apply to other sorts of authorship as well.

(A1) The affiliation of the author may influence our assessment of the information contained in the article he has written. For example, if we regard Harvard as being one of the best universities in the world, then a publication by a professor there would weigh heavily in our assessment of its content.

(A2) If we know something about the associations to which the author belongs, then we may take this into account when considering the quality of the information he provides. For example, if he belongs to a different social, cultural or religious group from us, then our initial reaction may be to be more critical of what he writes, though we may, for various reasons, seek to counteract this tendency.

(A3) We take into account the author's intelligence, experience and expertise. For example, we are wary of the writings of a person about a specialist topic that he is not an expert on. Of course, on some occasions this tendency of ours might prevent us from learning something useful.

(A4) If we know something of the author's goals or his agenda, this may make us wary of accepting his assertions uncritically.

The third group of factors relates to the *content* of the message:

(C1) The content of an assertion that we read may have properties that make us wary of believing it. One consideration concerns the coherence of the message and its internal consistency. Thus, we would not accept an inconsistent message or an assertion that was inconsistent with something else the author wrote in the same article.

(C2) The content of an assertion contained in a journal article may create such an intense emotional reaction in the reader of the article that this influences the way in which he assesses that assertion.

(C3) The content of an assertion may be so out of the ordinary that we are very reluctant to accept it without further ado. For example, we may read an article by someone who claims to have had an out-of-the-body experience. There are, of course, individual differences between people and not everyone would react to such an account in the same way.

The fourth group of factors relates to the *reader*:

(R1) We consider the importance of the message and its relevance to us. We tend to be less critical of assertions that are not particularly important to us than those that are really important to us.

(R2) Whether or not we accept an author's assertion may be influenced by our pre-existing knowledge. For example, if the assertion is straightforwardly inconsistent with what we already know and we are confident of the truth of the statements that it is inconsistent with, then we are unlikely to accept the assertion outright. We may, though, flag it as something we should investigate more fully later.

(R3) We consider the obvious consequences and repercussions of accepting the message. The consequences of accepting an assertion that we read may be so significant that we insist on getting further information before accepting it. For example, if the message is such that accepting it would have a profound effect on my current plans, my life-style or my belief-system, then I am unlikely to accept it outright, even if it comes from a reliable source. In such a case I would

probably flag the assertion as one that I need to consider thoroughly at some later time.

(R4) The character of the recipient may influence his assessment of assertions that he reads. For example, a creative person may be willing to entertain wacky and unusual ideas which a less creative person would be very wary of accepting or even spend time thinking about.

Conclusion

There are very many problems to overcome if we are ever going to build a humanoid robot with intellectual abilities analogous to those possessed by human beings. Although it may be impossible to design and build an android whose abilities replicate those possessed by human beings, it is sensible to design androids, at least initially, whose abilities are similar to human intellectual ones. In designing an android it makes sense to design one that human beings can interact with. If the android was very different from us, then this would not be possible.

In this paper I have concentrated on some of the problems that arise from the fact that human beings need a great deal of information in order to be fully-functioning members of a human society. It is impossible for them to generate all this knowledge for themselves. Most of this knowledge comes from other people. Any android that we design and build would be in the same position. Before we can even begin to design an android we have first to understand the abilities that humans have. In this paper I have made a start at investigating one aspect of how human belief-acquisition works. I have employed a two-stage model of belief-acquisition. In the first stage, as well as forming beliefs by using our senses, we also acquire beliefs by reading what other people have written and by listening to what they say. The ability to learn from others is, surprisingly, not currently being investigated very much. I am trying to rectify this curious omission from research in cognitive science. In this paper I have focused on how we acquire beliefs from journal articles and I have identified many of the factors that may cause us to override the defeasible rule, 'Believe what you read or hear'. A great deal of work still needs to be done before we understand human belief-acquisition sufficiently well in order to be able to implement it in an android or other artificial cognitive system. Currently, I am working on identifying the factors that make us wary of accepting outright what we read or hear. In this paper I have looked at those factors that may be invoked when we are reading a journal article and elsewhere (Diller, 2000b) I have identified the factors that may be invoked when we are listening to another person talk. In another paper (Diller, 2000a) I have looked at those factors that may cause us to be sceptical of what we read in a book. In future papers I plan to look at those factors that are at work when we read a newspaper, listen to the radio, watch television or find information on the Internet. After that the task still remains to further refine the two-stage model. I hope that some people reading this paper will be stimulated to join me in this exciting, but sadly neglected, field of research in cognitive science.

References

Rodney A. Brooks, Cynthia Breazeal, Matthew Marjanović, Brian Scassellati, and Matthew M. Williamson. nd. The Cog project: Building a humanoid robot. This article is going to appear in a volume of the Springer-Verlag Lecture Notes in Computer Science series of books. It is also currently available on the Internet: `http://www.ai.mit.edu/projects/cog/publications.html`.

Eugene Charniak and Drew McDermott. 1985. *An Introduction to Artificial Intelligence*. Addison-Wesley Series in Computer Science (consulting editor: Michael A. Harrison). Addison-Wesley, Reading (Massachusetts).

Antoni Diller. 1999. The belief-filter component. Cognitive Science Research Papers CSRP–99–9, School of Computer Science, University of Birmingham, April.

Antoni Diller. 2000a. Acquiring information from books. Cognitive Science Research Papers CSRP–00–10, School of Computer Science, University of Birmingham, July. This paper has been accepted for presentation at the 20th Specialist Group on Expert Systems International Conference on Knowledge-based Systems and Applied Artificial Intelligence (ES2000).

Antoni Diller. 2000b. Everyday belief-acquisition. In Gabriela P. Henning, editor, *Argentine Symposium on Artificial Intelligence (ASAI2000) Proceedings: Tandil, September 5–7, 2000*, pages 221–232. Sociedad Argentina de Informática e Investigación Operativa (SADIO), Buenos Aires.

John L. Pollock. 1987. Defeasible reasoning. *Cognitive Science*, 11:481–518.

Frank Robert Vivelo. 1984. The search for meaning and its biobehavioral correlates: An essay in philosophical anthropology. *Ultimate Reality and Meaning*, 7:133–155.