

Acquiring Information from Books

Dr Antoni Diller
School of Computer Science
University of Birmingham
Birmingham
B15 2TT
England

Abstract

People need a vast amount of knowledge in order to live in an advanced technological society. Most of this has to be obtained from others by believing what they say and what they have written. Androids and sophisticated AI systems would also have to be able to learn in this way. This obvious fact tends to be overlooked by AI researchers (such as Pollock and Brooks) involved in the design of androids. They concentrate almost exclusively on belief-formation by means of perception. However, before we can program the ability to learn from others into an android we need to have a better understanding of human belief-acquisition. Elsewhere I have proposed a two-stage model of belief-acquisition. In the first stage we do acquire beliefs by means of our senses, but also from other people. In this latter case we make use of a defeasible rule, ‘Believe what you hear or read’. The second stage consists in the use of a sophisticated critical methodology in order to carefully assess a small number of our beliefs. 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 statements found in a book. This turns out to be far more complicated than may at first sight appear.

1 Introduction

Some people working in AI see its ultimate goal as being that of constructing an artificial person [3, 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’ [2]. 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 which has largely been overlooked. As well as having a large amount of in-built 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 be able to interact with human beings in any meaningful way. People acquire a large number of beliefs every day in this way and to mimic human behaviour

androids would also have to be able to do this. Understanding what is involved in learning from others, so as to be able to implement it in a machine, is much more complicated than may at first sight appear and in this paper I investigate part of this human ability.

It is well known that human beings have the ability to learn from others. They have the ability to acquire knowledge and beliefs by listening to what other people say and by reading what other people have written. This ability is not optional. In order to live in a human society, especially an advanced, technological one, a person needs a vast amount of knowledge and information. The knowledge required 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 the institution of money and how it works, how goods are bought and sold, where various goods can be bought and so on. He also needs to know some basic facts about electricity, so that he can turn lights on and off and work various electrical appliances safely. He would also need to have some knowledge about different kinds of electrical appliances such as the television, radio, dishwasher, washing-machine, microwave, telephone, computer 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 a technologically advanced human society. He would, furthermore, need to have a certain amount of specialised knowledge in order to carry out his various social roles and to do his job (assuming that he has one). Thus, a medical doctor requires a huge amount of information in order to practice medicine and even a bricklayer needs a substantial amount of knowledge in order to do his job properly.

It is impossible for a person to acquire all the knowledge that he needs in order to live in a human society by himself. Most of this knowledge has to be obtained from other people. It is acquired by believing what they say and by trusting what they have written. Furthermore, 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 in order to live in a human society 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 re-

spected 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. The way in which I accommodate these observations is by proposing a two-stage or two-phase model of belief-acquisition. This model was introduced and explained elsewhere [4] and I say more about it in the next section of this paper, but here I just want to mention that anti-authoritarianism appears in the second stage of my model. People differ in how extreme their anti-authoritarianism is. The most radical version involves the acceptance of the principle that no belief or theory is beyond criticism and may, in fact, be criticised and revised if necessary. Such a variety of anti-authoritarianism is represented by pancritical rationalism [1]. Less extreme versions may ring-fence a number of beliefs and principles as being beyond criticism. The larger the number of such things that are beyond criticism the less radical is the version of anti-authoritarianism involved. If the mass manufacture of androids ever becomes a reality, then the diversity of approaches to the extent of criticisability that exists amongst human beings would, no doubt, be mirrored in their mechanical progeny.

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 many features 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. 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. It does not, however, involve thoroughly testing an assertion before it is accepted. 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.

So far in this introduction I have been writing mainly about human abilities. An android or AI system that was sufficiently advanced to be capable of interacting with human beings, talking to them, learning from them and maybe also teaching them would clearly have to have similar abilities to those described briefly above. In order to produce an android with these abilities,

we first need to have a good understanding of them 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 the first stage of belief-acquisition. I look in detail at the features of assertions found in books that make us wary of accepting them outright and I isolate many of the factors that cause us to override our default principle to accept what we read in books. My long-term goal is to formulate these things in sufficient detail so that they can be programmed into an android or AI 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. Looking at how we assess information given during a personal communication [5] should convince the reader of this.

2 The Two-phase Model

In this section I summarise the two-phase model of belief-acquisition that I introduced elsewhere [4] in a simpler form. 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.

Although most of our knowledge comes from other people, this is a fact that is either ignored by epistemologists or relegated to the periphery of the subject. Centre stage is occupied with issues relating to perception. Pollock, for example, writes [10, p. 52], ‘The starting point for belief formation is perception. Perception is a causal process that produces beliefs about an agent’s surroundings.’ 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 this ability is much more complicated than Pollock suggests and it is not the only way in which people acquire beliefs. There are, for example, an unlimited number of judgments that an agent can make about his immediate surroundings. Using myself as an example, I can make the following judgments about my current surroundings: ‘It’s not raining’, ‘The radio is switched off’, ‘This room is a mess’, ‘There is a bookcase near to the door’, ‘It’s peaceful in here’, ‘The door is open’, ‘Birds are singing outside’ and ‘There are several piles of books on the table’. The judgments an agent actually makes depend on a variety of factors in addition to the perceptual properties of his immediate surroundings. These might include his current belief-system, his goals and his values. In my case, if I was not writing this article, I would not have made any of these judgments. My purpose in making them was to make an epistemological point. Some of

the factors that influenced which judgments I made originated with me, but others came from other people. We acquire some, but not all, of our beliefs through perception and those that we do are acquired against the background of a large amount of knowledge that we have not obtained through perception. Acquiring beliefs by means of perception involves not only the perceiver and his surroundings but it also involves the perceiver's beliefs, values, goals and so on and a large part of these additional things he has obtained by listening to other people, reading what they have written and by going through an extended period of enculturation.

If I was restricted to acquiring beliefs by making judgments about my immediate surroundings, I would be extremely limited in the beliefs that I could acquire. My knowledge would be very restricted and would not be sufficient for me to be able to live and function in an advanced technological 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 take account of the sort of issues that I am interested in, because we do not simply accept everything that we read or hear. I will illustrate the sorts of factor that we take into account in assessing what we read in books by considering a few examples.

When listening to other people or reading books, our belief-acquisition is governed by means of the defeasible rule, 'Believe what you hear or read'. Most of the time when we read a book there is no reason for us to override this rule. For example, in *The Oxford Companion to Philosophy* [7, p. 378] I read that David Hume lived for a while in Paris. Having read that I now have the belief that Hume lived for a while in Paris. There is no reason for me to doubt the veracity of this fact. For example, Oxford University Press is a well-known publishing house with a reputation for producing reliable and authoritative reference books. Furthermore, I know very little about Hume's life and this piece of information does not clash with any of my pre-existing knowledge.

There are times, however, when I do not accept what I read. For example, in *Chariots of the Gods?* [12] von Däniken writes about the Nasca lines on the plains of Peru and says that they are giant runways for space-craft. Although I have read this, I have not added the belief that the Nasca lines are the markings of giant runways to my belief-system. This is because I have overridden the rule, 'Believe what you read or hear'. I have learned from other sources that von Däniken is unreliable. Furthermore, I know that the Nasca lines are drawn on the pebbly surface of the desert and would be destroyed if an aircraft tried to land on them.

In the two examples just discussed I was dealing with a situation in which we assess the information that we receive from a single source, but it sometimes occurs that we come across several sources relating to the same event and these sources are mutually inconsistent. The sources involved do not all have to be of the same kind. Thus, what we read in a book may conflict with what we hear on the radio or what a friend tells us may conflict with something that we have read in a newspaper. In these circumstances we sometimes use rules that compare the relative merits of these different sources. Although Quine and

Ullian think that we have meta-beliefs about the relative merits of different sources of information, rather than rules, as I am suggesting, they make a similar point [11, p. 14]:

We all hold . . . that those [beliefs] gained from respected encyclopedias and almanacs are more to be relied on than those gained from television commercials.

There are additional factors involved in this example which Quine and Ullian overlook. In acquiring information when our sources conflict we may have to take other things into account in addition to the relative merits of the sources involved. There are, for example, differences between people in how they evaluate reference works. I doubt that Quine and Ullian would regard *Harper's Encyclopedia of Mystical and Paranormal Experience* [6] or *The Encyclopaedia of Occult, Paranormal and Magick Practices* [8] as respected encyclopedias, but there are people who would rate them very highly as authorities. Thus, in addition to comparing sources a person's pre-existing knowledge and general outlook on life may affect what he does with a piece of information that he comes across in a book. When Quine and Ullian write 'We all hold', they seem unaware that they are members of a particular social group with many shared attitudes and beliefs, but which may not be shared with other social groups in the larger pluralistic society that they are members of. This also shows the value of the sort of investigation that I am carrying out in looking at the various factors involved in our assessment of what we read in books.

So far I have been explaining the first phase of belief-acquisition, but I now want to say something about the second phase. We do not have the time to thoroughly check every statement that we read or hear, but we do examine critically a small number of statements that are particularly important to us. For example, earlier in this section I quoted Pollock's assertion that the starting-point for belief-formation is perception [10, p. 52]. Many people reading his statement would simply accept it and add the corresponding belief to their belief-system. However, I think that Pollock is incorrect on this point and above I presented various reasons why I think that Pollock's statement is false and why I do not believe it. In its place I have various other 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 often get our beliefs by accepting what others say and what they have written.

The second stage involves making use of some sort of critical methodology in order to thoroughly investigate the correctness of some of our beliefs. It involves argumentation and reasoning. 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 [4, pp. 24–26].

There is a lot of interaction between a person's belief-system, the factors that he takes into account in order to decide whether or not to override the defeasible rule 'Believe what you hear or read' and his second-stage critical

methodology. The following are the three main ways in which these elements interact:

- (1) Beliefs are added to a person's belief-system if they do not trigger any of the factors that cause him to override the rule 'Believe what you hear or read'.
- (2) The factors that cause overriding involve various beliefs and so can change as the person's belief-system changes. For example, someone may rate a particular reference work very highly until he discovers a number of errors in it. From that time on information obtained from that source will be treated differently from the way in which it was treated before his opinion was revised. Furthermore, such a change in evaluation may entail a revision of beliefs obtained from that source in the past.
- (3) Beliefs are also added and removed from a person's belief-system as a result of the operation of that person's critical faculties.

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 interest, but I have not said much about it here, because my concerns in this paper are different.

3 Assessing Information Received from Books

Some of the information that a person has he has obtained by reading books. 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. If someone rigidly applied this rule and believed everything that he read in a book, then he would very quickly end up with a massively inconsistent belief-system. To see that this is so imagine someone reading Marx's *Das Kapital* followed by reading Popper's *The Open Society and its Enemies*. If such a person believed everything that he read, he would believe everything in Marx's book and everything in Popper's and, thus, he would have very many mutually inconsistent beliefs. Clearly, we do not accept everything that we read.

Given that the rule 'Believe what you read' is defeasible, we need to enquire into the circumstances when in fact it is defeated. What factors relating to a statement that we read in a book 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 book that is being read and the others relate to the *author* or *authors* of the book, the *content* of the statement being assessed and how that content may affect the *reader* of the book. I will next look in more detail at the various factors that occur in these four categories. 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 a book. In other words, the

reader should not assume that the factors that follow are all independent of each other.

3.1 External factors

- We take the kind of book involved into account. There are very many different kinds of book and we do not treat them all in the same way. The following is a selection of the various sorts of book that there are: novels belonging to various genres, textbooks, encyclopaedias, manuals, guide books to foreign countries, travel books, cookery books, biographies and autobiographies, various kinds of reference book, self-help books, religious books, true crime books, history books and so on. Although novels, for example, are written in a similar way to that in which factual accounts of real events are written, we do not accept what we read in novels as being factual statements to be added to our belief-system. We may, however, remember them as accounts of a fictional world.
- We may take the publisher of the book into account. There are, no doubt, differences between people concerning the status of various publishers. Many people in England, however, think highly of publishers like Oxford University Press and Cambridge University Press. There is a presumption that books, especially reference books, produced by such publishers are accurate and authoritative.
- We may take the place where the book was published into account. If we come across a book published in a country whose traditions of scholarship and publishing we are unfamiliar with, then we may be slightly more wary of accepting what we read outright. If we find that the information from such a source is generally reliable, then our confidence in similar books published in the same country may increase. Recently, for example, I came across a book published in Budapest by the Central European University Press [9]. I was slightly wary because Hungary used to be a communist state, but upon learning more about the publisher and the author I became more confident in the book's content as a source of information about Popper's thought. (The Central European University was founded by George Soros, the well-known supporter of the open society, and Notturmo is a researcher in the Karl Popper Archives.)
- We may take the year of publication into account. For example, if we want to learn about recent findings in a discipline with which we are not very familiar, then we would go for a book published in the last year or so. Standards of scholarship change over the years and, thus, knowing when a history book, say, was published may help us in assessing the quality of the information that it contains.
- We may take the edition of the book into account if it has gone through several editions. For example, if a book is in its seventh edition, then that edition is likely to be more accurate than the third edition, say.

3.2 Factors relating to the author or authors

Books can be written by one or more people or they can be edited with contributions from various authors. To make the following discussion easier to follow I shall assume that we are dealing with a book written by a single author. With suitable changes, the discussion can be made to apply to other sorts of authorship as well.

- If we know something about the author, then we may take this into account when considering the quality of the information involved. For example, if he belongs to a different social, cultural or religious group, 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. We also 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. If the book we are reading is a technical monograph or a textbook or something similar, then the affiliation of the author may influence our assessment of the information it contains. For example, if we regard the University of Oxford as being one of the best in the world, then a publication by a professor there would weigh heavily in our assessment of its content. If we know something of the author's goals or his agenda, this may make us wary of accepting his assertions uncritically. For example, a political or religious tract is written in order to convince the reader of the truth of some political ideology or religious doctrine. Knowing this about the tract influences the way in which we treat the information that it contains.

3.3 Factors relating to the content of the message

- The content of an assertion that we read may have characteristics that make us wary of believing it without further ado. 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 book. People rarely write straightforwardly inconsistent assertions, like 'It is raining and it is not raining', but may write two or more assertions that others may, being more logical and rational, see as being inconsistent.
- The content of someone's assertions may create an intense emotional reaction in the hearer and this may influence the way in which that person assesses further assertions from the same person.
- 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 a book by someone who claims to have been abducted by aliens. There are, of course, individual differences between people and not everyone would react to an account of alien-abduction in the same way.

3.4 Factors relating to the reader

- 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. For example, if somebody has little or no interest in Egyptian history and he reads that Ramesses II reigned from 1279 BC until 1213 BC, then he is likely to simply accept this. However, if he has taken a keen interest in David Rohl's new chronology, then he is going to be wary of accepting this and will probably investigate the matter quite thoroughly using some sort of critical methodology.
- 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. This consideration includes the case when our knowledge is that of what this author wrote elsewhere.
- 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.
- The character of the recipient may influence his assessment of assertions that he hears. 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.
- A person's maturity may influence his assessment of the assertions that he reads. Thus, an adult is likely to be far less credulous than a child. Experience would have taught him that people are not always as truthful as they should be.

4 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 any human society. It is impossible for them to generate all this knowledge by 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 AI research. In this paper I have focused on how we acquire beliefs from books 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 AI 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 book and elsewhere [5] I have identified the factors that may be invoked when we are listening to another person talk. In future papers I plan to look at those factors that are at work when we read a journal article or a newspaper, listen to the radio or 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 AI research.

Acknowledgement

I am grateful to Debra Barton for reading a number of drafts of this paper and for making many helpful suggestions.

References

- [1] W. W. Bartley, III, *The Retreat to Commitment*, second edn, Open Court Publishing Company, La Salle and London, 1984.
- [2] R. Brooks, C. Breazeal, M. Marjanović, B. Scassellati and M. Williamson, The Cog project: Building a humanoid robot. This paper will appear in a volume of the Springer-Verlag Lecture Notes in Computer Science series of books. It is also currently available on the Internet at <http://www.ai.mit.edu/projects/cog/publications.html>.

- [3] E. Charniak and D. McDermott, *An Introduction to Artificial Intelligence*, Addison-Wesley, Reading (Massachusetts), 1985.
- [4] A. Diller, The belief-filter component, *Cognitive Science Research Papers CSRP-99-9*, School of Computer Science, University of Birmingham, 1999.
- [5] A. Diller, Everyday belief-acquisition, *Cognitive Science Research Papers CSRP-00-7*, School of Computer Science, University of Birmingham, 2000. This paper has been accepted for presentation at the Argentine Symposium on Artificial Intelligence (ASAI 2000).
- [6] R.E. Guiley, *Harper's Encyclopedia of Mystical and Paranormal Experience*, HarperCollins, New York, 1991.
- [7] T. Honderich, (ed.), *The Oxford Companion to Philosophy*, Oxford University Press, Oxford, 1995.
- [8] B. Lane, *The Encyclopaedia of Occult, Paranormal and Magick Practices*, Warner Books, London, 1996.
- [9] M. A. Notturmo, *Science and the Open Society: The Future of Karl Popper's Philosophy*, Central European University Press, Budapest (Hungary), 2000.
- [10] J. L. Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, Cambridge (MA) and London (England), 1995.
- [11] W. V. Quine and J. S. Ullian, *The Web of Belief*, second edn, McGraw-Hill, New York, 1978.
- [12] E. von Däniken, *Chariots of the Gods? Unsolved Mysteries of the Past*, Souvenir Press, London, 1969.