UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Understanding Detective Stories

Permalink

Journal

Authors

Pratt, Ian
Xu, Luopinh
Leudar, Ivan

Publication Date

1992

Peer reviewed

Understanding Detective Stories

Ian Pratt, Luoping Xu and Ivan Leudar
Departments of Computer Science and Psychology
University of Manchester,
Manchester M13 9PL, UK
email: luoping@cs.man.ac.uk

Abstract

In this paper, we illustrate a general approach to psychological inference by considering its application to a simple detective story. Detective stories provide a fertile ground for the investigation of psychological inference, because their plots so often hinge on the mental states of the characters involved. Although our analysis contains several, logically independent suggestions for how to tackle some of the different problems that arise in understanding the story, one guiding principle underlies our approach: the *re-use thesis*. According to the re-use thesis, certain inferential mechanisms whose primary function has nothing in particular to do with psychological inference can be re-used for psychological inference tasks. In the course of the paper, we present several examples of the re-use thesis in action. Finally, we sketch how these applications of the re-use thesis can contribute to an understanding of our detective story.

1 Introduction

In this paper, we illustrate a general approach to psychological inference by considering its application to a simple detective story. The development of computer programs able to understand simple stories has long been seen as a major challenge for artificial intelligence, and the human being's remarkable ability to infer details left implicit in almost any narrative text, and thereby to forge the story into a coherent explanatory whole must be the envy of any present-day computer reasoner.[1] We focus on detective stories because of their relation to the present authors' wider interest in psychological inference—inference about the beliefs, desires, intentions, and other mental states of agents. The study of psychological inference occupies a central place in cognitive science, not only because of its obvious practical applications, but also because the concepts of *belief, desire, intention* etc. hold centre stage in the philosophy of mind. An account of how these concepts function in thought would thus address a central issue in the philosophy of mind.

Detective stories provide a fertile ground for the investigation of psychological inference, because their plots so often hinge on the mental states of the characters involved. Which of the suspects knew the contents of the heiress's new will? Did the mysterious Australian couple want to kill her for some reason? And why did the art dealer lie about the value of her grandfather's portrait? Answering and—more importantly—raising questions of this kind are essential in understanding a detective story; indeed, it is harder to think of a richer mine of problems connected with psychological inference.

2 A simple example

Consider the following simple detective story.[2]

> Gunner wanted to kill Ridley. He went to see him at his flat shortly before Ridley was due to play a match at the tennis club, and shot him dead. Then he dragged the body over to the door, took the telephone off the hook, and drove to the tennis club. When Ridley failed to turn up for the match, Gunner offered to telephone to see what happened. He went into the telephone booth and dialed. After a few minutes, he dashed out and said that, while he was talking to Ridley, there was a knock at the door, a scuffle of voices, and then shooting, followed by the door closing. The police were called and found Ridley's body. Later, the inspector talked to Gunner again about the telephone call. Then he quietly informed Gun-

[1] There is a large literature on story-understanding in AI. For a representative example, see (Dyer, 1983).

[2] This plot is taken from one of the stories in a beginners' textbook in German.

ner that, according to the telephone company, Ridley's telephone had been disconnected last week. He then charged Gunner with the murder of Ridley, and Gunner was taken away by two waiting policemen.

The above story illustrates some of the inferences readers of detective stories are called upon to make. The central problem in understanding this story—and no one to whom we showed the story found it difficult to solve—is to work out how the detective inferred that Gunner was the murderer, not just an innocent (telephonic) witness. The pivotal move, of course, is to hit upon the question (from the detective's point of view) of how Gunner knew that Ridley was dead, given that the telephone was disconnected. For known he must have done. And unless he was implicated in the murder, Gunner's behaviour simply does not make sense.[3]

We can get some inkling of just how tricky this understanding process is by noticing that the sentence "Gunner knew that Ridley was dead" appears nowhere in the story; still less is there any reference to *the detective's beliefs as to whether* Gunner knew that Ridley was dead. Readers must therefore infer without prompting that the detective can reason that Gunner must have known that Ridley was dead, and they must then raise the question as to how (from the detective's point of view) Gunner could have come by this knowledge. Then, and only then is it a (relatively) short step to seeing how the detective can infer that Gunner may have shot Ridley himself and faked the telephone call to give himself an alibi.

3 Putting yourself in the detective's shoes

The current literature on psychological reasoning is replete with references to the idea that psychological reasoning may proceed by psychological simulation.[4] The idea is simple enough: if one wants to reason

about what some one else will infer or decide to do who believes p, q, r and desires x, y, z, one imagines that p, q, r are true and that x, y, z are desirable, and one sees what one infers or decides to do. One thus uses one's own cognitive state (as it evolves within the imaginary environment) to simulate, and thereby predict, the cognitive state of another person. Well, that is a complicated and certainly controversial story, and we cannot possibly review all of its strengths and weaknesses here. Nor can we discuss the complex problems involved in temporarily suspending certain of one's beliefs. Nevertheless, it is hard to resist the idea that, at least sometimes, one reasons about the thoughts of other people by temporarily entering the make-believe world of those people.

Our detective story is a case in point. Understanding that story really amounts to understanding how the detective worked out that Gunner was the (likely) murderer, a fact of which the reader is informed in the second sentence. And it is plausible that, in trying to understand the detective's inferences, one temporarily suspends that information and imagines knowing what the detective knows: Gunner's rushing out of the telephone booth, the subsequent discovery of Ridley's body, the fact that the telephone was found not to have been working, and so on. Once one is in that cognitive predicament, one's thoughts race on as the detective's must have. The result is the answer to understanding the story.

Certainly, the thesis that psychological reasoning proceeds by psychological simulation can at best constitute a partial account of psychological inference. In particular, simulation (as the term is understood here) is inherently unsuitable for inferring causes from effects ("The *reason* why Gunner faked the telephone call was such-and-such") or for gaining *universal* or *necessary* information ("Gunner *must have* known that Ridley had been shot.") Nevertheless, in favourable situations, simulation is attractive because it holds out the prospect of a certain economy in thought. One needs no special psychological theory to reason about how the detective reasons; one just needs to be able to reason *as* he reasons, and then to observe how that reasoning goes and what it produces. Thus, according to the simulation idea, inferential mechanisms that support one's ability to make inferences and decisions generally can be *re-used* in imaginary situations to reason specifically about what inferences and decisions other people will make.

---

[3]Some people to whom we put this story point out alternative explanations (from the detective's point of view) on which Gunner is not really guilty at all. However, no one we asked failed to understand why the detective might at least strongly suspect Gunner. That is: all those we asked understood the critical issue (for the detective) of how Gunner knew that Ridley had been shot.

[4]Pratt (Pratt, 1989) includes a brief guide to the simulation idea and its manifestations in the literature. Goldman (Goldman, 1989) also analyses the simulation idea (but with a different philosophical orientation); Leudar (Leudar, 1991) discusses the underlying supposition of psychological *similarity* between different persons. Perner (Perner, 1991) provides a comprehensive survey of recent work of the development of psychological concepts in children, in which the simulation idea plays an important role.

4 Understanding Lying

In general, understanding detective stories involves the ability to understand lying and pretence, and why people engage in it. In our story, for example, the detective finally makes sense of Gunner's actions because he sees them as a ploy to fabricate an alibi. How should our story-understanding program work this out?

A useful heuristic in explaining why an agent S performs an action α is to imagine oneself in S's position and see what one can infer about the likely effects of α. If consequences C arise which are, as far as one can judge, positive from S's point of view, then it is sensible to consider the hypothesis that S did α because he intended to bring about C. [5] (Notice that this process is *not* one of simulation.) It goes without saying that that is only part of the story about how one might assign motives: for one thing, S's actions can have unforeseen beneficial consequences; for another, if one is not privy to all of S's beliefs, one cannot unproblematically put oneself in S's position in order to determine the consequences of α by S's lights. Nevertheless, some version of the suggested rule is likely to be a good heuristic for *forming hypotheses* about why S did what he did.

A variation on this heuristic can be used to hypothesize motives for lying. Suppose S tells a lie, P, and one wants to know why. Well, as a special case of the above heuristic, one can consider what S thinks the effects of saying that P are. And one way to solve this problem is to put oneself in the imaginary state of someone hearing the utterance P, but without the special information that S is lying. If, in this imaginary state, one draws a conclusion which S might want one to draw, or reaches a decision which might be favourable to S, then it is sensible to consider the hypothesis that S lied because he wanted one to draw that conclusion or reach that decision. Again, it goes without saying that that is only part of the story about how one might assign motives for lying (or for pretence generally): for one thing, S's utterances have unforeseen consequences; for another, since one is not privy to all of S's beliefs about one's own state of mind, one cannot be sure what S will take one's reaction to his statement to be. Nevertheless, some version of the suggested rule is likely to number among the good heuristics for forming hypotheses about why S lied.

Our detective story is again a case in point. Why, from the detective's point of view, might Gunner have faked the telephone call? Well, if one imagines not knowing that the telephone was disconnected, the natural explanation of Gunner's behaviour is that Gunner heard shooting on the phone because someone shot Ridley while Gunner was on the telephone to Ridley. This would, after all, explain why Ridley's body was found in his flat, and why the telephone was off the hook. Let us call this the *naive explanation*. To be sure, there *are* other explanations of Gunner's actions (there are always other explanations), but the naive explanation seems—without the benefit of knowing that Ridley's telephone was not working—the most plausible.[6] And, of course, on this explanation, Gunner could not have shot Ridley. According to the strategy suggested in this section, then, one can reason as follows about why Gunner lied: if one first imagines not knowing that Gunner was lying about the telephone call, one infers the naive explanation, and concludes that Gunner did not kill Ridley. Now, since Gunner might want one to draw this conclusion, a possible explanation of his actions—we might call it the *alibi-explanation*—is that it was a means to get people to believe he is innocent. Notice how, here again, we are suggesting that a general inference mechanism—the mechanism whereby one determines the consequences of actions—can be *re-used* in psychological inference.

5 Raising questions about beliefs

Let us return to the question of how the detective came to see that Gunner must have known that Ridley had been shot. As mentioned above, this issue is crucial to understanding the story. We proceed via an informal experiment. Imagine learning that someone has telephoned the police to say that there has been a shooting in a house not far from where you live. Imagine in addition that you have no more information about circumstances in which the call was made, or about the identity of the caller. (This is all very improbable, but that does not matter for our purposes.) Suppose now, you are asked to conjecture what might happen next. Presumably, you would answer (roughly) thus: "Armed police will go to the house to see if anything is wrong. If all is quiet, they will knock on the door. If someone answers, they will ask questions about the incident. If, on the other hand, there is no answer, they may try to force their way in. Once inside, they will look for a body or signs of a shooting. If they find no body, they will trace the occupants of the house, ..."

---

[5]See (Pratt, 1990).

[6]The problem of how, in general, one constructs explanations and decides between competing explanations is not addressed by the present paper. For some discussions of this huge topic, see (Harman, 1986, Lipton, 1991). See also (Leake, 1990, Antaki & Leudar, 1991).

Being able to conjecture the effects of events in the presence of only very partial background information is an indispensable inferential accomplishment. We might call the inferential process involved *scenario-branching*, because of the tree of alternative scenarios that is generated. Do the police find the house deserted? Does anyone answer the door when the police knock? When the police search the house, do they find a body? And so on: many unanswered questions, and correspondingly many branch-points in projecting the possible course of events. But it is absolutely vital not to consider *all* conceivable questions. One should not consider whether the police turned right or left as they left the police station, or whether they drove past a lady with a pram, or (if the knock on the door was answered) whether they were offered a cup of tea, and so on. In short, branching is to be avoided if it is unlikely to yield to interestingly different results; otherwise, projection would be stymied by an infinity of irrelevant possibilities. How people control the branching of possibilities when asked to consider very partially specified scenarios is a question we prefer not to address here. All we claim is that they do.

Now put yourself again in the position of the detective in the story. You have heard Gunner's account of the telephone call, and you know that it cannot be true. You might then wonder what Gunner could possibly have thought he would gain by such a pretence, and to do this, you might simply consider, from Gunner's point of view, what would happen when the police are called in such a situation. Since you are largely in the dark about Gunner's beliefs, there are a number of possibilities. We suggest that, in trying to determine the most interesting of these, you can simply engage in scenario-branching as just described. That is, you can simply consider what happens when the police are called with a report of a shooting, and ask yourself what might transpire. As we have seen, a tree of alternative scenarios will branch out, as the significant unknown factors occur to you. This time, of course, the object of the exercise is not to determine what may happen, but what *Gunner might have thought may happen*. And this changes the way you use the branch-points as the scenario-projection mechanism throws them up. Thus, you ask not whether there will be anyone at home to answer the door when the police call, but whether *Gunner thinks there will*; not whether the police will find a body, but whether *Gunner thinks they will*, and so on.

As we have already observed, the crucial move in understanding the detective story above is to make the unprompted inference (reasoning within the detective's predicament) that Gunner must have known Ridley was dead. The problem here is not so much *verifying* this fact as *thinking of it in the first place.* And, in this section, we have put forward a mechanism for how one might be led to consider this possibility. (The problem of verifying it is not considered in this paper.) Our suggestion is that one can deploy the mechanism for scenario-branching—the mechanism, that is, whereby one conjectures the effects of actions in very partially specified situations. The attractiveness of our suggestion resides in the fact that, according to it, there is no need for a special psychological theory from which we can raise important questions about the beliefs of agents; the ability to conjecture the effects of their actions in unspecified circumstances works, we suggest, well enough.

In each of the previous three sections, we have described how inferential mechanisms whose primary function has nothing in particular to do with psychological inference can be re-used for psychological inference: (i) inferential mechanisms generally can, in favourable cases, be re-used in 'pretend-mode' to reason forwards about the psychological state of another person; (ii) inferential mechanisms for projecting the consequences of actions can be re-used to hypothesize motives of the agents who perform them (a variation on this process involves divining the reasons for lies); and (iii) inferential mechanisms controlling the branching of possibilities when projecting the consequences of actions in incompletely specified situations can be re-used to raise important questions about the beliefs of agents. The idea of the re-use of inferential mechanisms in the keystone in our approach to psychological inference. So much so, it deserves a name:

The re-use thesis: Inferential mechanisms whose primary function has nothing in particular to do with psychological reasoning can be re-used for psychological reasoning tasks.

We stress that, for the purposes of this paper, the re-use thesis is to be considered as a suggestion in AI— that is, as a suggestion for how to design a computer program capable of effective psychological inference. The question of whether *human beings* re-use non-psychological inference mechanisms for psychological reasoning is one we shall not discuss.

6 Putting the bits together

In order to test some of the ideas developed here, a program is currently being written to process some stories requiring psychological inference. In this section, we sketch very briefly how to assemble the foregoing elements into an explanation of the above story.
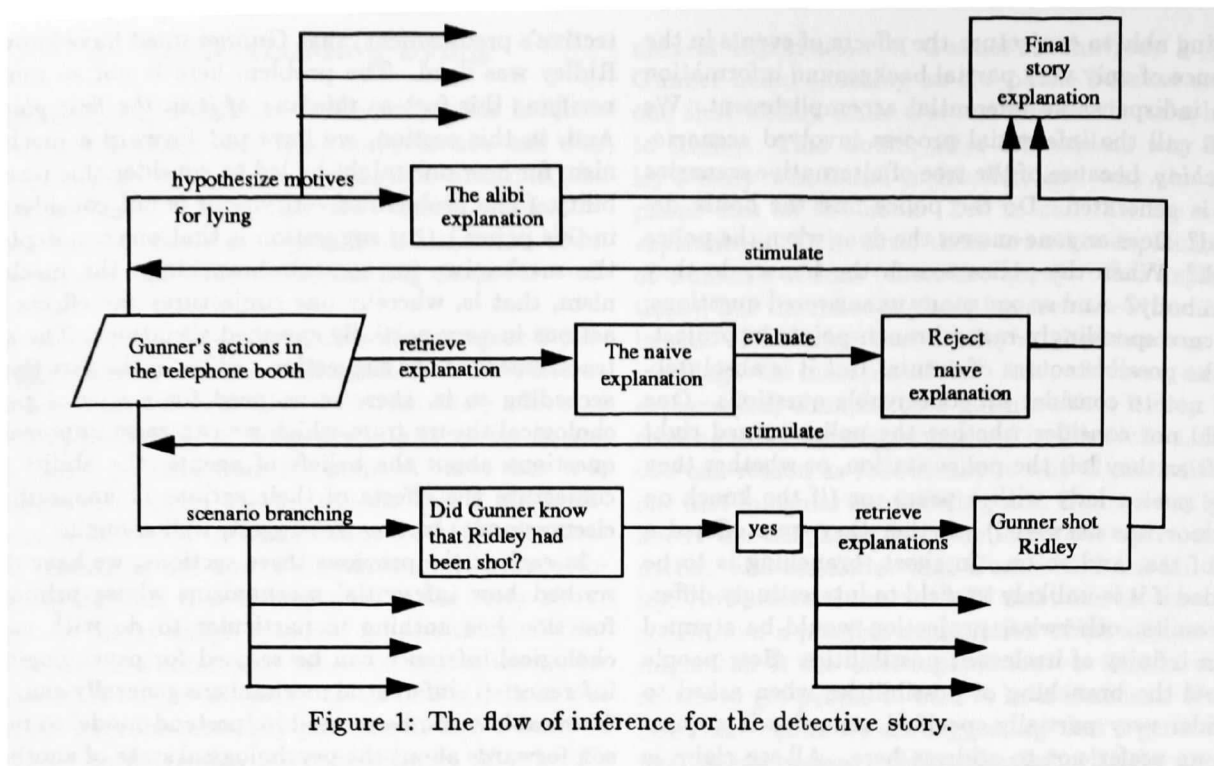
Figure 1: The flow of inference for the detective story.

We certainly do not claim to have a theory of inference general enough to account in an uncontrived way for all of the inferences the story demands of its readers. Our primary concern in this project is the relationship between psychological and non-psychological inference problems, and the present section is included merely to provided an overview of how we envisage our view of that relationship as contributing to the understanding of the above detective story.

The flow of inference is depicted very schematically in fig. 1. The program starts by trying to explain Gunner's assertion that he heard Ridley shot on the telephone. The first explanation the program finds is the *naive explanation*, according to which Gunner really did hear Ridley get shot and wanted to summon help to him. When this explanation is rejected (because it is discovered that the telephone was disconnected) other explanations are considered. At this point, two parallel strands of reasoning are stimulated. The first strand follows the procedure described in section 4 for hypothesizing motives for lying. Here, several alternative explanations are constructed, among them, the so-called *alibi-explanation*, according to which Gunner lied about being on the telephone to Ridley in order to get people to believe he did not kill Ridley. It is important to be clear that, at this stage, the alibi-explanation is only a hypothesis as to why Gunner said what he said—one explanation, that is, among a field of possible contenders. The decisive factor that makes the alibi-explanation

(in contrast to its competitors) ultimately acceptable emerges from the second strand of reasoning.

The second strand of reasoning addresses the problem of reconstructing the detective's inference that Gunner must have known all along that Ridley had been shot. The program adopts the procedure described in section 5 for raising important questions about the beliefs of agents, by performing scenario-branching on the proposition that Gunner said that he heard Ridley being shot on the phone. As a result of this process, the question arises as to whether Gunner thinks that Ridley has been shot. Once the program has hit upon this issue, it can confirm that indeed Gunner must have known. The process of confirming this conjecture is somewhat complicated and lies outside the detailed issues discussed in this paper, but, briefly, the program determines that, unless Gunner had good reason to believe that Ridley had been shot, he could expect nothing but negative consequences (including a conviction for wasting police time) to flow from his actions. The next step in this second strand of reasoning is to explain *how* Gunner could have known of Ridley's shooting. A number of possible standard explanations for coming-to-know are considered, many of which can be rejected for reasons that need not detain us here. The one standard explanation for coming to know something that bears fruit is the one on which, in this case, Gunner knew that Ridley had been shot because he actually did the shooting. Having inferred that Gunner shot Rid-

ley, the program deduces that Gunner will have the goal of getting the detective to believe that he did not shoot Ridley (again, the details of this inference are not considered in this paper).

This is where the two strands of reasoning link up. The first strand has hypothesized a number of possible motives for Gunner's actions, among them, alibi-explanation that Gunner wanted people to believe he did not kill Ridley; the second strand has inferred that Gunner did kill Ridley and will therefore have the goal of getting people to believe that he did not. For this reason, the program chooses the alibi-explanation from among its competitors. The result is the inference that Gunner killed Ridley and devised the scene in the telephone booth to fabricate an alibi.

7 Conclusions

In this paper, we have considered a simple detective story with a view to analysing some of the inferential mechanisms needed to understand it. In particular, we have highlighted the need for the reader to reason within a number of imagined cognitive predicaments. Although our analysis contains several, logically independent suggestions for how to tackle some of the different problems that arise, one guiding principle underlies our approach: the re-use thesis. According to the re-use thesis, certain inferential mechanisms whose primary function has nothing in particular to do with psychological reasoning can be re-used for psychological reasoning tasks. We described three examples: (i) inferential mechanisms generally can, in favourable cases, be re-used in 'pretend-mode' to reason forwards about the psychological state of another person; (ii) inferential mechanisms for projecting the consequences of actions can be re-used to hypothesize motives of the agents who perform them (a variation on this process involves divining the reasons for lies); and (iii) inferential mechanisms controlling the branching of possibilities when projecting the consequences of actions in incompletely specified situations can be re-used to raise important questions about the beliefs of agents whose actions are to be explained. For present purposes, we interpret the re-use thesis as a suggestion for AI—that is, as a suggestion for how to construct effective *computer* reasoning systems. All we can claim to have established in the work reported here is a plausible case for its efficacy. The proof of the pudding will be, as ever, in the eating.

8 References

Antaki, C. and Leudar, I. 1991. Explaining in Conversation,*European Journal of Social Psychology*, 22(2):181-194.

Dyer,M.,1983. *In-depth Understanding: A Computer Model of Integrated Processing for narrative comprehension.* Cambridge, MA: MIT Press (1983).

Goldman, A. 1989. Interpretation Psychologized. *Mind and Language*, 4(3).

Harman, G., 1986. *Change in View.* Cambridge, MA: MIT Press.

Leake, D. 1990. Task-Based Criteria for Judging Explanations, In Proceedings of the Twelfth Annual Conference of the Cognitive Science Society, 325-332 Cambridge, MA.

Leudar, I., 1991. Sociogenesis, Coordination and Mutualism. *Journal for the Theory of Social Behaviour*, 21(2):197-220.

Lipton, P. 1991. *Inference to the Best Explanation.* London: Routledge.

Perner, J. 1991. *Understanding the Representational Mind.* Cambridge, MA: MIT Press.

Pratt, I. 1989. Psychological Inference, Constitutive Rationality and Logical Closure, in Vancouver Studies in Cognitive Science, vol.1:366-389. Information, Language and Cognition, Vancouver, BC: University of British Columbia Press.

Pratt, I. 1990. Psychological Simulation and Beyond, in Proceedings of the Twelfth Annual Conference of the Cognitive Science Society, 654-661. Cambridge, MA.