

CROSS DOMAIN INFERENCE AND PROBLEM EMBEDDING

Robert Cummins

I. INTRODUCTION

I.1. Two reasons for studying inference.

Inference is studied for two distinct reasons: for its bearing on *justification* and for its bearing on *learning*. By and large, philosophy has focused on the role of inference in justification, leaving its role in learning to psychology and artificial intelligence. This difference of role leads to a difference of conception. An inference based theory of learning does not require a conception of inference according to which a good inference is one that justifies its conclusion, whereas, obviously, an inference based theory of justification does require such a conception.¹ Because of its focus on normative issues of justification, philosophy has taken a retrospective approach to inference, whereas a focus on learning naturally leads to a prospective approach. A focus on learning leads us to ask, "Given what is known, what should be inferred? How can what is known lead, via inference, to new knowledge?" A focus on justification has led philosophers to concentrate instead on a retrospective question: "Given a belief, can it be validly inferred from what is known? How can what is known justify, via inference, a new belief?" Thus, for philosophy, inference can be regarded as permissive: one needn't worry about what to infer, only about whether what has been arrived at somehow or other is or can be inferentially justified. A theory of learning, on the other hand, requires a conception of inference that is directive, for the problem of inference based learning is precisely the problem of what to infer.

I.2. Inference and learning.

Inference is relevant to learning because it is a method of expanding or revising existing knowledge. There are two main limitations of inference based accounts of learning.

(1) Inference has to be inference *from* something; it requires premises. Inference based learners must be "seeded". Thus, inference based theories of

¹ Acquiring a concept or a motor or intellectual skill are standard cases of learning, yet neither concepts nor skills need be justified to count as learned. Indeed, it isn't even obvious what it would mean to say that a concept or skill is justified. Even knowledge acquisition, as this is conceived in psychology and AI need not, and typically does not, involve justification: Knowledge, as this concept is used in the theory of learning, needn't be true, let alone justified.

learning must either be nativist, or they must acknowledge some non-inferential learning. This consequence is well-known, and I will say no more about it.

(2) Less often noticed is a difficulty that arises in connection with the acquisition of knowledge expressed in novel concepts/terms. Inference, on the face of it, cannot lead from existing knowledge K to new knowledge K' in cases in which K and K' are representationally disjoint. Premises expressed on one vocabulary do not relate inferentially to conclusions expressed in another. Thus "cross-domain" learning, as I shall call it, poses a problem for inference based accounts of learning. The problem has a Kantian ring: How is cross-domain inference possible?

We deny the possibility of cross-domain inference at some cost. If there is no cross-domain inference then each domain must be seeded by its own proprietary innate knowledge, or by some kind of non-inferential learning. Pan-nativism is preposterous, however, and non-inferential learning mechanisms are thin on the ground, and not plausible for the introduction of "higher knowledge" in any case. We do well, therefore, to inquire what resources there are for cross-domain inference.

1.3. Cross-domain inference.

Only two kinds of cross-domain inference have received any extended attention: explanation by and confirmation of theory, and analogy.

Theoretical explanation and confirmation. Many scientific theories are expressed in a vocabulary that is disjoint from the vocabulary appropriate to the expression of the theory's intended explananda and confirming data. Given that explanation and confirmation are typically explicated in terms of (or as forms of) inference, the problems of theoretical explanation and confirmation have long been recognized as instances of cross-domain inference in the philosophy of science. But, although there is a large literature on these issues, there is nothing in it to help us with the problem of learning by cross-domain inference. There are two reasons for this. First, as has been typical in philosophy, a normative, retrospective approach to the problem has dominated. (but see Hanson, 19~~). The distinction between the context of discovery and the context of justification, made popular by positivist philosophy of science, was designed to canonize such an approach by relegating problems of discovery -- i.e., of learning -- to psychology on the grounds that they were not amenable to rational assessment. Thus, the problem of how knowledge in one domain might lead to knowledge in another was ruled out of bounds in this context in favor of the standard normative and retrospective problem: Given a theory and a body of data, or an alleged explanandum, what inferential relations must hold

if the one is to be confirmed by or explain the other? (The classic text is Hempel, 19~~).

In spite of this neglect of the problem of learning, one might hope that the literature on theoretical explanation and theory confirmation might provide some ideas about how distinct domains can be linked inferentially. Unfortunately, the only idea one finds is that there must be "bridge principles", laws linking the two domains explicitly. Little is said about where bridge principles come from. Again, the worry is only about what, if anything, might justify them once they have been formulated. From the point of view of the learning theorist, nothing could be more disappointing: one is presented with the most obvious brute force patch -- to link Σ 's to S's, introduce a conditional 'If Σ then S' -- and no hint as to how the patches might be generated.²

Analogy. Surprisingly, analogy is the only recognized *form* of inference that explicitly addresses the cross-domain problem. For this reason, I am not here to bury analogy, but to praise it. Perhaps the discussion so far will stimulate more research on analogy by indicating its unique importance. Still, it would be nice if there were another tool in the box to help cope with cross-domain inference.

II. Task embedding.

There is another way that cross-domain inference can be facilitated. The basic idea is to embed a task requiring knowledge of domain D' in a task requiring knowledge of D in such a way that success in the D task is contingent on success in the D' task. Reasoning about D thus constrains reasoning about D'. I propose to illustrate this approach in some detail by describing a technique for learning the meanings of arbitrary symbols/icons and simple syntactic constructions for combining them.

II.1. Illustration: learning to understand simple symbols and syntax.

The embedding task. Someone, call her Leader, blazes a trail through a simple branching maze like that shown in figure one. The blazes

² In the context of justification, this isn't a problem: theorists typically supply the intended connections between theory and observation or explanandum. It is not my intent to criticize the literature on theory justification; my point is only that we cannot hope that that literature has already invented the wheel learning theory is looking for.

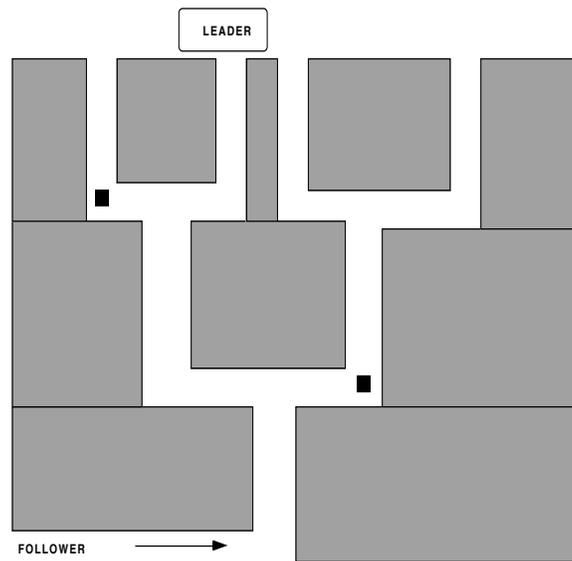


FIGURE 4 ■ MEANS "AVOID THIS PATH"

marks the route to avoid. You retrace your steps to the previous intersection and take the unmarked path. At the next intersection, you discover a square to your left. You go right, in accordance with your current hypothesis, and find Leader. Your hypothesis is confirmed. Presented with other mazes blazed (correctly) with squares, you go directly to Leader without any false steps. You now have a convention with Leader governing the use of the square.

Next, consider figure five. You know what the circle means, but not the bar. If you are like most subjects, you stick with what you know, taking the path with the circle. The next intersection prompts you to retrace you steps, and subsequently to avoid paths with the circle-bar

"Cross domain inference and task embedding," in Cummins and Pollock, eds., *Philosophy and Artificial Intelligence: essays at the interface*, M.I.T Press, a Bradford Book, 1992, pp. 23-38.

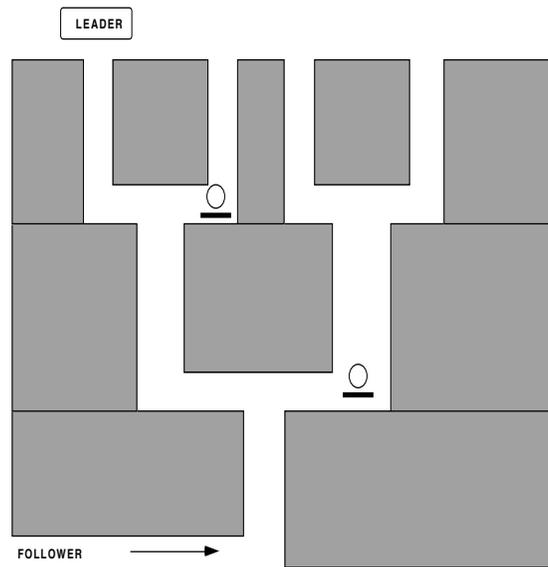
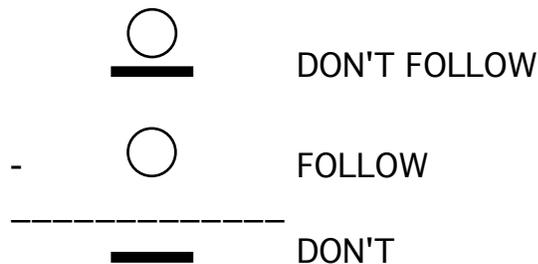


FIGURE 5 ○ MEANS "DON'T TAKE THIS PATH"

blaze, a policy that takes you to Leader with no further problems. You now know that the circle-bar combination means, and what the circle means by itself. If you are like most subjects, asked what the bar means, you will say something like this: "It means: do the opposite of what the other symbol means." That is, you will do a kind of subtraction problem:



Moreover, presented with the maze of figure six, you will find Leader with no false steps.

"Cross domain inference and task embedding," in Cummins and Pollock, eds., *Philosophy and Artificial Intelligence: essays at the interface*, M.I.T Press, a Bradford Book, 1992, pp. 23-38.

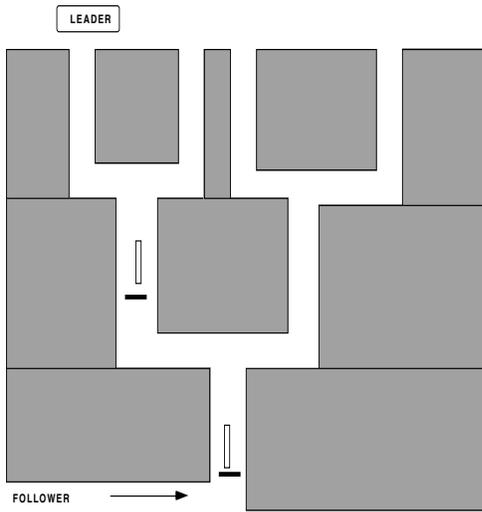


FIGURE 9
MEANS "DON'T GO RIGHT AT THE NEXT BRANCH"

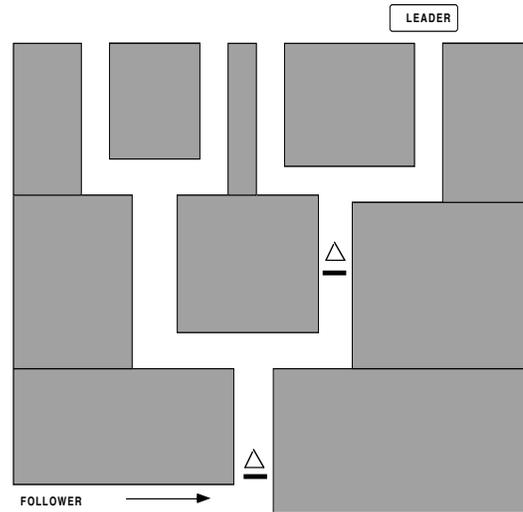


FIGURE 10
MEANS "DON'T GO LEFT AT THE NEXT BRANCH"

Figure eleven introduces a new twist: symbols that refer to particular objects in the maze. Having no experience with the "squiggle", subjects typically ignore it and proceed on the basis of the vertical bar which they know to mean "go right". Eventually, they discover that going right meets with success when it is done at the church. After a few mazes blazed in this way, subjects will say the squiggle means "church". Analogous experience with \square , combined with previous experience, allows for rather complex situations such as those in figures twelve and thirteen.

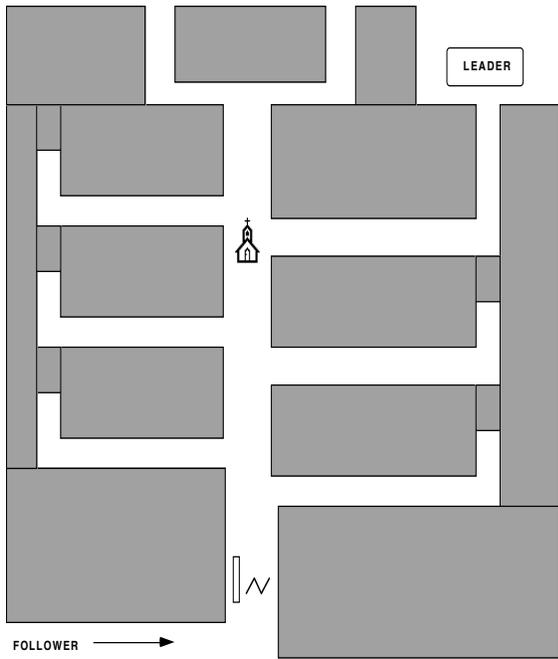


FIGURE 11  MEANS "GO RIGHT AT THE CHURCH"

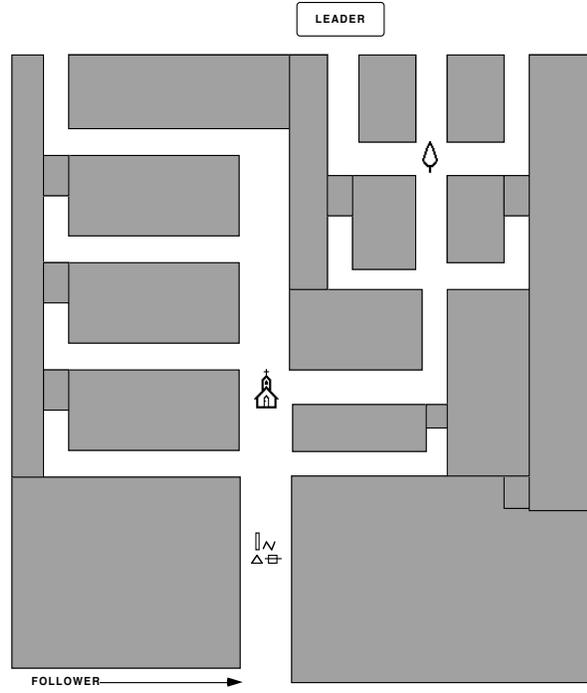


FIGURE 12  MEANS "GO RIGHT AT THE CHURCH, GO LEFT AT THE TREE"

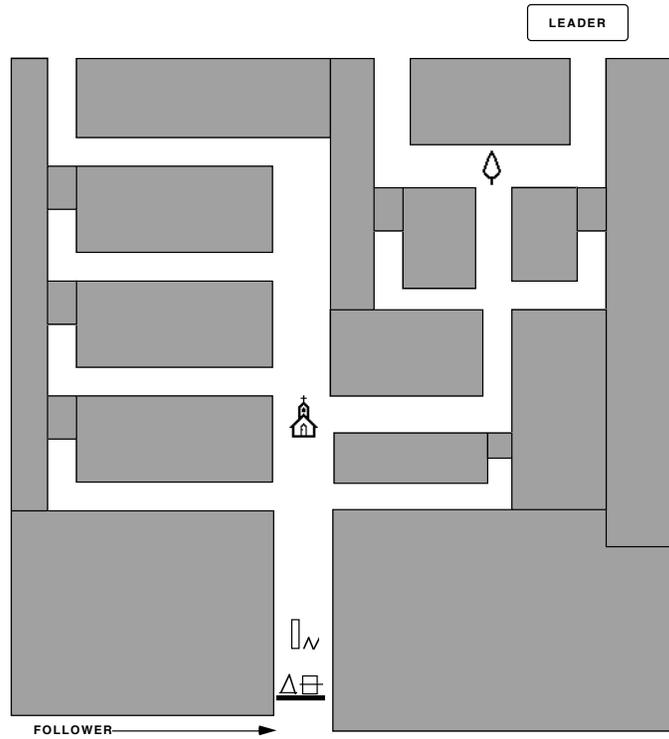


FIGURE 13  MEANS "GO RIGHT AT THE CHURCH, DON'T GO LEFT AT THE TREE"

II.2. Remarks about the illustration.

In the illustration just rehearsed, both the embedded task and the embedding task are coordination problems (Lewis, 1969), i.e., problems in which two or more agents share a goal and must coordinate their actions in order to achieve it. While it is obvious that the embedding task is a coordination problem, it isn't so obvious that the embedded task is a coordination problem. To see that it is, notice that it is a communication problem: the shared goal is successful communication between Leader and Follower. What successful communication requires, roughly, is that Understander adopt correct beliefs about Meander's intentions in producing the communicative signals -- blazes, in this case; Understander's beliefs must coordinate with Meander's intentions.⁴

Communication is an especially difficult coordination problem for two reasons. (i) Intentions and beliefs are hidden, hence it is difficult to determine whether coordination has been achieved. (ii) Since any symbol or construction could mean anything, the range of hypotheses Understander might have to canvass is essentially unbounded. Embedding the communication task in the maze task allows a solution to both of these problems by allowing knowledge of the embedding task to constrain inferences required to perform the embedded task. Feedback about success or failure to properly recognize Leader's communicative intentions is provided in a readily accessible and easily appreciated manner by the fact that success in the maze task is contingent in an obvious way on success in guessing Leader's communicative intentions. Embedding the communication task in the maze task has the effect of turning the communication task into a case of supervised learning, even though there is no teacher. The environment, together with knowledge about the embedding task, play the role of a teacher in an entirely natural way, allowing Follower easy access to information about Leader's communicative intentions.

Working together with the feedback provided by the embedding task is the equally important fact that the embedding task severely limits the range of hypotheses that Follower needs to consider. Early in the learning task, Follower's hypotheses about Leader's communicative intentions (i.e., about the meanings of the blazes encountered) are effectively limited to direct correlates of Follower's possible actions. Subjects faced with the figure one maze always regard themselves as faced initially with two possibilities: The square means that Leader went down the path with the square, or the square means that Leader went down the path without the square. Given knowledge of the embedding task, there is simply nothing else relevant that the square could mean.

⁴ See Grice (1959), Bennett (1973), Lewis (1969) and Cummins (1979) for more on the analysis of communication in terms of intentions and beliefs.

An important feature of the learning achieved via task embedding is that the acquired knowledge can be applied successfully in the absence of the embedding task that enabled learning in the first place. Subjects who learn the meaning of symbols for *left, right, take, avoid, tree, church*, and the like in the maze context have no trouble applying this knowledge when the symbols are encountered in different contexts.

_____ *Canoe paddling*. The subject is in the rear paddling position (marked "P" in the figure), and can paddle either forward or backward on the left or right. Instructions are issued from the forward position in the form of arbitrary icons for which conventions have been established in the maze context. Although both paddlers can see the destination, the person in the forward position is the only one who can see obstacles (e.g., submerged rocks). Conventions for "left," "right," "forward," and "back" , acquired in the maze context, generalize to this context without difficulty.

_____ *Giving/following directions*. This is very similar to the maze task. Director gives directions to a specified location in a city, and Traveler attempts to follow them. "Left," "Right," "Tree," and so on transfer unproblematically.

III. Caveats and Conclusions

The project described above demonstrates that task embedding can effect cross-domain inference in the broad sense in which this means that knowledge of one domain constrains inferences in another. It is important, however, to be clear about what the illustration does not show.

First, the illustration involves embedding coordination problems, and the learning involved is a species of convention acquisition. While it *seems* clear that the technique will work for other kinds of tasks and learning, I haven't actually tried it.

Second, you can't learn -- at least not inferentially -- what a symbol means if you don't already have the capacity to represent the meaning in question (Fodor, 1975). This point *does* generalize: task embedding constrains hypothesis formation and confirmation, but it does not provide new representational resources.

Third, task embedding, like analogy, presupposes appropriately-structured knowledge of the source (embedding) domain. Just as knowing that A is analogous to B doesn't help you infer things about A if you don't know anything, or the wrong things, about B, so embedding A in B doesn't help you infer things about A unless you know the right things about B.

Taken together, the second and third points must temper optimism about inference-based learning. Cross-domain inference (analogy, task embedding) allow the inference-based learning theorist to avoid one kind of pan-nativism,

"Cross domain inference and task embedding," in Cummins and Pollock, eds., *Philosophy and Artificial Intelligence: essays at the interface*, M.I.T Press, a Bradford Book, 1992, pp. 23-38.

viz., seed knowledge in every domain, but leaves nativism of two other kinds intact: You have to know about the source domain, and you have to have the representational power already in place. We still have to explain the origin of the knowledge of the source domain, and we still have to explain the origin of the concepts used to formulate hypotheses in the new domain. These are the central challenges to inference-based learning. Compared to these problems, the problem of cross-domain inference, while interesting and important, is just detail.

References

Bennett, J. (1973), "The Meaning Nominalist Strategy", *Foundations of Language*, 10, 141-68.

Cummins, R. (1979), "Intention, Meaning and Truth-Conditions", *Philosophical Studies*, 35, 345-60.

Fodor, J. (1975) *The language of thought*. New York: Crowell.

Grice, H. (1957), "Meaning", *Philosophical Review*, 78, 147-77.

Hanson, N. (19~~)

Hempel, C. (19~~) <<the collection>>

Lewis, D. (1969), *Convention*. Cambridge: Cambridge University Press.