Inferring Ignorance from the Locality of Visual Perception

Ernest Davis
New York University
251 Mercer St.
New York, NY 10012

Abstract

This paper presents a logical theory that supports high-level reasoning about knowledge and perception. We construct a formal language in which perception can be described. Using this language, we state some fundamental axioms, and we show that these are sufficient to justify some elementary but interesting inferences about perception. In particular, our axioms make it possible in some cases to infer that an agent does not know about a particular event, because he has had no way to find out about it.

1. Introduction

Intelligent creatures learn much of what they know through direct perception. Therefore, reasoning about the acquisition of knowledge over time often requires a high-level understanding of the power and limits of perception. This paper presents a logical theory that supports high-level reasoning about knowledge and perception.

One type of reasoning that may be supported by a theory of perception is the inference that an agent must be ignorant of a particular fact because he has had no way to find out whether it is true. Such means of inferring ignorance may be important, either to infer that a secret can be kept from another party, or to help an agent plan to find out a given fact.

We focus on the following problem:

I. Steve is in a closed room with no windows and walks across the room. Claire is outside the room. Infer that Claire does not know now that Steve has crossed the room.

We further require that our theory support the following inferences:

II. Andrew is in his office and does not see any cows there. Infer that he knows that there are no cows in his office.

III. Joanne does not know whether there are flowers on the dining room table. However, she knows that the top of the dining room table is visible from any point in the dining room. Infer that Joanne knows that she can find out whether there are flowers on the dining room table by going and looking.

IV. Fred has seen that Max has been with him all night. Infer that Fred knows that Max was not five miles away any time that night.

V. Judy sees Sharon standing facing a bus. Infer that Judy knows that Sharon knows that there is a bus in front of her.

These problems illustrate various aspects of perception. In (I) we infer that an agent is ignorant of a fact from our knowledge of the physical limits of vision. Since Claire cannot see Steve inside the room, she cannot know what is happening in the room. (II) shows the gaining of knowledge from inferences based on both prior knowledge and perception. Given what Andrew knows about the size of cows, the presence of a cow is physically incompatible with what he sees. (III) involves reasoning about possible future states of perception and knowledge. (IV) involves perception extended over an interval of time. (V) shows that one agent can infer the perceptions and knowledge of another agent by perceiving their physical situations and knowing their perceptual and inferential powers. The formal model developed in this paper is rich enough to support the statement and proof of close analogues of all the above inferences.

Our purposes are quite different from those of computer vision research; we therefore need a theory of perception at a different level of description. The facts about vision needed to solve problems (I) through (V) above are little used in computer vision programs. For example, consider the fact that it is impossible to see what is happening outside a closed room from inside it. Such a fact has little importance in computer vision, since it rarely constrains the interpretation of a particular image. However, it is of importance in reasoning qualitatively about how vision will augment knowledge in a given situation; one can use it to deduce that an agent will have to leave a closed room to know what is happening outside it.

Some limitations of our theory should be noted. First, we ignore learning through perceptions of conventional signs such as writing. Second, we model only visual perception. Finally, our models of vision and of physical causality are greatly simplified. We are interested here in the connection between perception and knowledge, and a more complex physical theory would have added irrelevant complexity. We believe that the essential structure of our theory will carry over to more realistic physical models.

This work was supported by NSF grant #DCR-8603758.

2. The Toy World

We formulate our theory within a "toy" world. Many of the assumptions below are obviously invalid in general; however, the basic structure is largely independent of these assumptions. In particular, the relations between knowledge and perception are independent of the particular physics of motion and vision assumed.

Our toy world is constructed as follows: There is a fixed set of physical objects moving about through space. Objects are rigid, maintaining a constant shape; however, their position may change continuously. At each time instant, each object occupies a connected regular set of points. The places occupied by two objects at a single time may not overlap. Besides their shapes, objects have time-invariant visible properties, such as their coloring. Objects may be also be characterized in terms of non-physical properties and relations, such as "being a Republican".

There are no other physical restrictions on the world. A course of events is physically possible iff each object maintains a constant shape and properties and moves continuously, and no two objects ever overlap.

Some objects are *agents*. At each instant of time, an agent has a body of knowledge with the following properties:

A.1 Knowledge of axioms: All general axioms — axioms of predicate calculus, geometry, time, physics, knowledge and perception — are known.

A.2 Consequential closure: Any logical implication of the agent's knowledge is known.

A.3 Veridicality: All knowledge is true.

A.4 Positive introspection: If an agent knows a fact, he knows that he knows it.

A.5 Memory: If an agent knows a fact (with no time indexicals) at one time, he knows it at all later times.

A.6 Internal Clock: An agent always knows what time it is.

An agent also has perceptions. A point is visible to an agent if it is not occluded from him by an object in between. An agent A can see the visible properties of an object at a visible point. Moreover, A can see the properties, not only of individual points but also of connected sets of points. Specifically, let X be a connected set such that each point of X is visible to A. If all of X lies inside an object O with visible properties $P1$, $P2$, ..., then A sees that X lies in some object with properties $P1$, $P2$, ... If X lies in free space, then A sees that X lies entirely in free space.

Perception provides knowledge about the world, and it is the only source of information as to which of the many physically possible courses of events is actually happening. Specifically, we assume the following:

A.7 Anything that is perceived is known.

A.8 If a physical statement is physically possible, and it does not contradict any past or present perceptions, then it cannot be known to be false.

3. Formal Model

Our model of knowledge derives from Hintikka's (1971) possible worlds semantics for knowledge. Moore (1980) combined this with a temporal logic, by identifying epistemic possible worlds and temporal situations. We modify Moore's model in two ways. First, we adopt a continuous model of time [McDermott, 82]. Second, we use two levels of possible worlds: *layouts* and *situations*. A *layout* is a timeless physical description of the instantaneous state of the world. A *situation* is a placement of a layout within a temporal structure and a system of knowledge relations. Perceptions are associated with layouts; knowledge is associated with situations.

In our toy physics, an object is an atomic individual with a set of visual properties. A layout specifies the objects in the world and the figure occupied by each object. The figure occupied by an object O in layout L is denoted "place(O,L)".

A *behavior* describes the progression of layouts over time. The function "scene(B,T)" maps a behavior B and a time T to the layout of B at time T. Behaviors are constrained by the requirements that all layouts in the behavior have the same objects; that objects have constant shape over time; and that objects move continuously.

We distinguish certain layouts and behaviors as *physically possible*. In our simple physics, a layout is physically possible if no two objects overlap, and a behavior is physically possible if each of its layouts is physically possible. We allow physically impossible layouts and behaviors as valid objects of thought; this simplifies the physical axioms. Even physically impossible behaviors must obey the above constraints.

What the agent can see in a given layout is determined by the laws of vision. The perceptions of an agent A in a layout L fix all aspects of L at points which are not occluded from A, and fix no aspects of L at points which are occluded from A. We say that a layout $L1$ is visually compatible with L with reference to A, written "v_compatible(A,L,$L1$)" if $L1$ is consistent with everything that A can perceive in L. V_compatibility is an equivalence relation over layouts.

Our theory of the scope and limits of vision is expressed in terms of the properties of the v_compatible relation. In our model of vision, $L2$ is v_compatible with $L1$ with respect to agent A if both $L2$ and $L1$ are physically possible and the following condition holds: Let X be a connected set of points, such that every point in X is visible to A in $L1$. Then each point of X is visible to A in $L2$. Moreover, if X lies entirely in A in $L1$, then X lies entirely in A in $L2$; if X lies entirely in some object $O1$ in $L1$, then, in $L2$, X lies entirely inside some object $O2$ with the same visible properties as $O1$; if X lies entirely in free space in $L1$ then X lies entirely in free space in $L2$.

A behavior $B1$ is visually compatible with behavior $B0$ for agent A up to time T if, as far as A can see in $B0$ up to time T, the world could be going through $B1$. We write this relation, "bv_compatible(A,$B0$,$B1$,T)". In our model of vision, we assume that all such information comes through the layouts; that is, two physically possible behaviors are visually compatible if corresponding layouts are visually compatible.

$$\text{bv\_compatible}(A,B1,B2,TS) \equiv \forall_{T \le TS}\ \text{v\_compatible}(A,\text{scene}(B1,T),\text{scene}(B2,T))$$

Two states of the world may be identical in their physical layout and yet differ in other respects. To accommodate this, we define a *situation* as a state of the world, including the physical layout, the non-physical properties and relations of objects, and the knowledge states of agents. The knowledge states of agents, however, are not a component of the situation, but are encoded in accessibility relations between situations. The function "layout(S)" maps a situation to its layout.

Non-physical properties of objects are made parts of situations rather than of layouts, in order to allow different agents to have different degrees of knowledge about them. In our system, we can allow Tom to know that all cows are large, but Sid not to know this. This would not be possible if we associated the non-physical property of being a cow with layouts.

The knowledge of an agent in a situation S is represented by an accessibility relation between S and other situations that are consistent with his knowledge. Let A be an agent and let $S1$ and $S2$ be two situations. $S2$ is accessible from $S1$ relative to A, written "k($A,S1,S2$)", if as far A knows in $S1$, the state of the world might be $S2$. We say that A knows in $S1$ that ϕ is true if ϕ is true in all situations that are accessible to A from $S1$.

Note that in both the visual compatibility relation and the knowledge accessibility relation, more information corresponds to a smaller extension of the relation. The more you know, the more variations in the world you can rule out as false, and therefore the fewer possible states of the world are consistent with your knowledge.

A *chronicle* describes a progression of situations over time. The function "situation(C,T)" maps a chronicle C and a time T onto the situation in that chronicle at that time. Each situation S has a unique time in a unique chronicle, denoted "time(S)" and "chronicle(S)". Each chronicle C has associated a behavior, which is the progression of layouts of the scenes.

$$\text{scene}(\text{behavior}(C),T) = \text{layout}(\text{situation}(C,T))$$

If situations $S1$ and $S2$ are in the same chronicle, and the time of $S1$ is earlier than the time of $S2$, we say that $S1$ precedes $S2$.

$$\text{precedes}(S1,S2) \equiv [\ \text{chronicle}(S1) = \text{chronicle}(S2) \wedge \text{time}(S1) \le \text{time}(S2)\ ]$$

We can achieve properties A.1 - A.8 by imposing the following requirements:

A.1, knowledge of the axioms, and A.2, consequential closure, follow immediately from the definition of knowledge, together with the axiom that all chronicles have a physically possible behavior.

$$\forall_C\ \text{phys\_poss}(\text{behavior}(C))$$

A.3, veridicality, holds if knowledge accessibility is reflexive.

$$\forall_{A,S}\ \text{k}(A,S,S)$$

A.4, positive introspection, holds if knowledge accessibility is transitive.

$$\forall_{A,S1,S2,S3}\ \text{k}(A,S1,S2) \wedge \text{k}(A,S2,S3) \Rightarrow \text{k}(A,S1,S3)$$

A.5, memory, holds given the following: Let situation $S1B$ be accessible from $S0B$, and let $S0A$ precede $S0B$ in the same chronicle. Then, since everything the agent knows in $S0A$, he also knows in $S0B$, and $S1B$ is consistent with everything he knows in $S0B$, there must be a scene $S1A$ in the chronicle of $S1B$ which is accessible from $S0A$. (Figure 1)

$$\forall_{A,S1B,S0B,S0A}\ [\ \text{k}(A,S0B,S1B) \wedge \text{precedes}(S0A,S0B)\ ] \Rightarrow \exists_{S1A}\ [\ \text{k}(A,S0A,S1A) \wedge \text{precedes}(S1A,S1B)\ ]$$

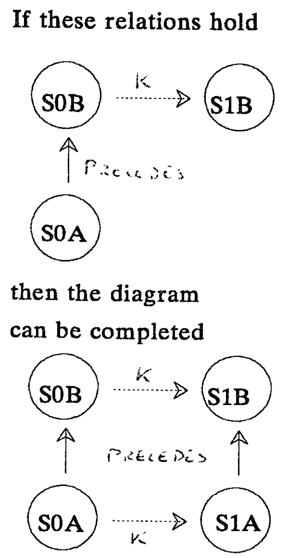

Figure 1

Axiom of Memory

A.6, internal clock, holds if any two accessible situations occur at the same time.

$$\forall_{A,S0,S1}\ \text{k}(A,S0,S1) \Rightarrow \text{time}(S0) = \text{time}(S1)$$

A.7, that perceptions are known, holds given the following: If $S1$ is knowledge accessible from $S0$, then the layout of $S1$ is visually compatible with the layout of $S0$. That is, for a situation $S1$ to be consistent with an agent's knowledge, the layout of $S1$ must be compatible with what the agent sees; conversely, if a layout is not compatible with what the agent sees, the agent knows that that cannot be the real layout.

$\forall_{A,S0,S1}$ k(A,S0,S1) $\Rightarrow$
v_compatible(A,layout(S0),layout(S1))

A.8, that perception is the only source of knowledge of the course of events, holds given the following: Let C0 be the real chronicle. Let B1 be a behavior that is visually compatible with the behavior of C0 up to time T relative to agent A; thus, as far as A could have seen up to time T, B1 could be the real behavior. Then it is consistent with A's knowledge that B1 actually was the real behavior; that is, B1 is the behavior of some knowledge accessible chronicle C1.

$\forall_{A,C0,B1,T}$ bv_compatible(A,behavior(C0),B1,T) $\Rightarrow$
$\exists_{C1}$ [k(A,situation(C0,T),situation(C1,T)) $\wedge$
B1 = behavior(C1)]

4. Proof

We can now sketch how an analogue to inference (I) can be formulated in our model, and proven from the above axioms of perception and knowledge, together with suitable axioms of geometry and physics. The full version of this paper [Davis, 88] shows the formulation and sketches the proofs of all the inferences (I) through (V), and it gives a complete formal proof of (I). These are omitted here due to length limitations.

Claire is on one side of a wall, for an interval of time i0. On the other side of the wall, occluded from Claire, is an object omystery. The object lies strictly within some larger region, which is entirely occluded from Claire. (Figure 2). During i0, Claire stays motionless, the object stays within its envelope, and no other object ever intersects the envelope. We wish to prove that there is no way for Claire to know whether the object is motionless or whether it is moving around within its envelope, since either are equally compatible with the motions of the objects that Claire does see. This conclusion can be formalized as follows: At the end of i0, there is one knowledge accessible situation that follows on a chronicle in which the object is motionless; there is another knowledge accessible situation that follows on a chronicle in which the object is in motion.

[$\exists_{C1}$ k(aclaire,s0z, situation(C1,end(i0))) $\wedge$
motionless(omystery,behavior(C1),i0)] $\wedge$
[$\exists_{C2}$ k(aclaire,s0z,situation(C2,end(i0))) $\wedge$
$\neg$motionless(omystery,behavior(C2),i0)]

To prove this, we construct two particular behaviors. In the first, every object moves just as it does in the real world except that omystery stays motionless throughout i0. In the second, every object moves just as it does in the real world except that omystery moves continuously within xenvelope throughout i0. We show that both of these are physically possible, since no other object comes within the envelope, by hypothesis, and so no other object interacts

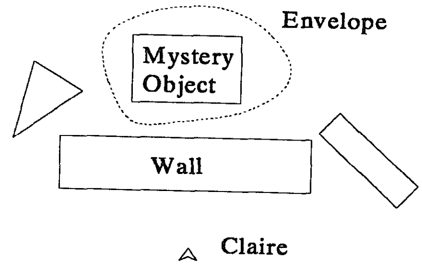

Figure 2
Claire cannot know
whether the object is moving

with omystery. Both compatible with Claire's perceptions, since the identical objects are visible to Claire in the identical places. Hence, by axiom A.8, Claire cannot know of either of them that it did not occur.

5. The Problem with Inferring Ignorance

The above proof rests on axiom A.8, which states that, if a course of events is possible and it is compatible with the agent's perceptions over time, then the agent cannot know that this is not the course of events that took place. Clearly, some such axiom is needed if we are to deduce the ignorance of agents from the limits of their perceptions.

Unfortunately, axiom A.8 is so strong that it rules out many plausible states of knowledge. This problem is particularly exacerbated by the weakness of our physics, which allows all kinds of courses of events as possible. For example, suppose that Fred and a table are the only objects in a closed room. In our theory, there is no way that Fred can ever be sure that the table is the only other object in the room; nor is there any way that he can be sure that he will ever see the table in the future. For it is physically possible that there is a swarm of small bees which has hitherto always hidden itself on the far side of the table, but which is just about to come and completely surround him, so as to occlude his viewing anything in the future. Since this is physically possible, by axiom A.8 he cannot know that it is false.

The ultimate problem here is that our theory of knowledge, like most such theories, approximates rationality in terms of the axiom of consequential closure, the assumption that an agent can make all logical deductions. It has often been pointed out that this axiom is too strong [Konolige, 86] [Levesque, 84]; it is less often noted that it is also too weak, and that plausible reasoning must be allowed as a source of knowledge

6. Previous Work

Little work has been done in AI on reasoning about perception. ATTEND, the focussing of a sensory organ, was a primitive act in conceptual dependency [Schank, 75] and was causally connected to MBUILD, the performance of a mental act; but the logic of these acts was not developed in detail. A number of theories, such as [Appelt, 82], [Allen and Perrault, 80] and [Morgenstern, 87], have studied the acquisition of knowledge through communication, but these have not looked at direct perception. The definition of a perception as a set of physical layouts was put forward in [Davis, 84]. Hintikka (1969) used a modal theory of perceptions to eliminate the need for sense-data as ontological primitives. The situation semantics of [Barwise and Perry, 82] studied the logical structure of sensory verbs in detail, but did not relate it to knowledge acquisition or to physical constraints. Reiter and Mackworth (1987) give a formal account of the relation between an image and a physical situation.

There have been a number of "active" vision systems that have reasoned about where to look to acquire relevant knowledge. For example, the SHAKEY robot looked for landmarks to locate itself. There have also been vision programs, such as [Garvey, 76] and [Selfridge, 82] that have used constraints based on the limitations of visual processes in object recognition. The work in the paper can be viewed as presenting a formal mechanism which justifies such inferences, and which will enable a intelligent system to reason about them directly.

7. Conclusions

We have given a formalism in which a few basic problems relating perception and knowledge can be stated and solved. There are two significant technical innovations in this work. The first is the concept of a physical layout, which specifies just the physical state of the world, and the description of perception in terms of layouts. The second is axiom A.8, which limits an agent's knowledge of contingent physical facts of certain types to that which can be deduced from his perceptions, together with physical laws. This axiom, however, places constraints on an agent's knowledge that are often unacceptably strong. In most practical problems, the positive inference, "Since A sees ϕ, A knows ϕ," is more important that the negative inference, "Since A doesn't see ϕ, A cannot know ϕ;" hence, it may be best to drop axiom A.8 or to restrict its scope.

The theory of the connection between perception and knowledge is largely independent of the physical theory. In particular, axioms A.1 through A.7 may be used together with any set of physical laws, and with any laws delimiting the powers of vision in terms of bv_compatibility. Axiom A.8 can also be used with any definition of physical laws and laws of vision, but it may put strong limits on the range of agents' knowledge of physical properties; that is, it may require that either all agents always know the value of a particular physical property or that no agent ever know it.

Acknowledgements

Thanks to Leora Morgenstern and John Sterling for helpful suggestions and to Ralph Grishman and Larry Manevitz for pointing out the difficulties discussed in section 5.

References

[Appelt, 82] Douglas Appelt. Planning Natural-Language Utterances to Satisfy Multiple Goals. SRI Technical Note 259. 1982.

[Allen and Perrault, 80] James Allen and C. Raymond Perrault. Analyzing intention in utterances. *Artificial Intelligence*, vol. 15, 143-178. 1980

[Barwise and Perry, 82] Jon Barwise and John Perry. *Situations and Attitudes*, MIT Press. 1982.

[Davis, 84] Ernest Davis. An Ontology of Physical Action. NYU Tech. Rep. 123. 1984.

[Davis, 88] Ernest Davis. Inferring Ignorance from the Locality of Visual Perception. In D. Miller and D. Atkinson (eds.) *Planning Systems for Mobile Robots*, Prentice Hall, to appear. 1988.

[Garvey, 86] Tom Garvey. Perceptual Strategies for Purposive Vision, SRI AI Center Tech. Rep. 117. 1986.

[Hintikka, 69] Jaako Hintikka. On the Logic of Perceptions. In *Models for Modalities* D. Reidel Publishing, Dordrechts, Holland. 1969.

[Hintikka, 71] Jaako Hintikka. Semantics for Propositional Attitudes. In L. Linsky (Ed.), *Reference and Modality*, Oxford University Press. 1971.

[Konolige, 86] Kurt Konolige. *A Deduction Model of Belief*, Pitman Publishing, London. 1986.

[Levesque, 84] Hector Levesque A Logic of Explicit and Implicit Belief. In *Proc. AAAI-84*, 198-202. 1984

[McDermott, 82] Drew V. McDermott. A Temporal Logic for Reasoning about Processes and Plans. *Cognitive Science*, vol. 6, 101-155.

[Moore, 80] Robert Moore, *Reasoning about Knowledge and Action*. SRI Technical Note 191. 1980.

[Morgenstern, 88] Leora Morgenstern. Foundations of a Logic of Knowledge, Action, and Communication. NYU Ph.D. Thesis. 1988.

[Reiter and Mackworth, 87] Ray Reiter and Alan Mackworth. The Logic of Depiction. Tech. Rep. 87-24, U. British Columbia Computer Science Dept. 1987.

[Schank, 75] Roger C. Schank. *Conceptual Information Processing*, North Holland, Amsterdam. 1975.

[Selfridge, 82] Peter Selfridge. Reasoning about Success and Failure in Aerial Image Understanding, Tech. Rep. 103, U. Rochester Computer Science. 1982.