# CHALLENGES TO MACHINE LEARNING:

## Relations between reality and appearance

John McCarthy, Stanford University

• Apology: My knowledge of of machine learning is no more recent than Tom Mitchell's book. Its chapters describe, except for inductive logic programming, programs aimed at classifying appearances.

• We live in a complicated world that existed for billions of years before there were humans, and our sense organs give us limited opportunities to observe it directly. Four centuries of science tell us that we and the objects we perceive are built in a complicated way from atoms and, below atoms, quarks.

• Science, since 1700, is far better established than any kind of philosophy. Bad philosophy has stunted AI, just as behaviorism stunted psychology for many decades.

• Besides the fundamental realities behind appearance studied by science, there are hidden every day realities—the three dimensional reality behind two dimensional images, hidden surfaces, objects in boxes, people's names, what people really think of us.

• Appearance is quite different from reality. Most machine learning research has concerned the classification of appearances and has not involved inferring relations between reality and appearance. Robots and other AI systems will have to infer such relations.

• Human common sense also reasons in terms of the realities that give rise to the appearances our senses provide us. Indeed young babies have some initial knowledge of the permanence of physical objects.

• Perhaps if your philosophy rejects the notion of reality as a fundamental concept, you'll accept a notion of *relative reality* appropriate for the design and debugging of robots. Thus the robot needs to be designed to determine this relative reality from the appearance given by its inputs.

• We'll discuss:

• Dalton's atomic theory as a discovery of the reality behind appearance.

- The use of touch in finding the shape of an object. Results of an experiment in drawing an object which one is only allowed to touch - not see.

- A simple problem involving changeable two dimensional appearances and a three dimensional reality.

- Some formulas relating appearance and reality in particular cases.

- What can one know about a three dimensional object and how to represent this knowledge.

- How scientific study and the use of instruments extends what can be learned from the senses. Thus a doctor's training involving dissection of cadavers enables him to determine something about the liver by palpation.

## ELEMENTS, ATOMS, AND MOLECULES

• Some scientific discoveries like Galileo's $s = \frac{1}{2}gt^2$ involve discovering the relations between known entities. Patrick Langley's Bacon program did that.

• John Dalton's postulation of atoms and molecules made up of fixed numbers of atoms of two or more kinds was much more creative and will be harder to make computers do.

• The ancient ideas of Democritus and Lucretius that matter was made up from atoms had no important or even testable consequences. Dalton's did.

• Giving each kind of atom its own atomic mass explained the complicated ratios of masses in a compound as representing small numbers of atoms in a molecule. Thus a sodium chloride (NaCl) molecule would have one atom of each of its elements. Water came out as $H_2O$.

• The simplest forms of the atomic theory were inaccurate. [Early 19th century chemists didn't soon realize that the hydrogen and oxygen molecules are $H_2$ and $O_2$ and not just H and O.] Computers also need to be able to propose theories adventurously and fix their inaccuracies later later.

• Only the relative masses of atoms could be proposed in Dalton's time. The first actual way of estimating these masses was made by Maxwell and Boltzmann about 60 years after Dalton's proposal. They realized that the coefficients of viscosity, heat conductivity, and diffusion of gases as explained by the kinetic theory of gases depended on the actual sizes of molecules.

• The last important scientific holdout against the reality of atoms, the chemist Wilhelm Ostwald, was convinced by Einstein's 1905 explanation of Brownian motion. The philosopher Ernst Mach was unconvinced.

• The first actual pictures of atoms in the 1990s were a big surprise. An actual picture of a proton showing the quarks would be even more surprising and seems quite unlikely.

• Philosophical point: Atoms cannot be regarded as just an explanation of the observations that led Dalton to propose them.

Maxwell and Boltzmann used the notion to explain entirely different observations, and modern explanations of atoms are not at all based on the law of combining proportions. In short, atoms were discovered, not invented.

## ELEMENTS, ATOMS, MOLECULES—FORMULAS

• Most likely, it is still too hard to make programs that will invent elements, atoms, and molecules. Let's therefore try to write logical sentences that will introduce these concepts to a knowledge base that has no ideas of them.

• We assume that the notions of a body being composed of parts and of mass have already been formalized, but the idea of atom has not. The ideas of bodies being disjoint is also assumed to be formalized.

• The following formulas approximate a fragment of high school chemistry and should be somewhat *elaboration tolerant*, e.g. should admit additional information about the structure of molecules.

The situation argument $s$ is included only to point out that material bodies change in chemical reactions.

$$
\begin{aligned}
& Body(b,s) \rightarrow (\exists u \subset Molecules(b,s))(\forall y \in u)(Molecule(y) \wedge Part(y,b)), \\
& y1 \in Molecules(b) \wedge y2 \in Molecules(b) \wedge y1 \neq y2 \rightarrow Disjoint(y1,y2), \\
& Part(x,b,s) \rightarrow (\exists y \in Molecules(b,s)) \neg Disjoint(y,x), \\
& Body(b,s) \rightarrow Mass(b,s) = \textstyle\sum_{x \in Molecules(b,s)} Mass(x,s).
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
&Water(b) \wedge x \in Molecules(b) \\
&\rightarrow (\exists h1\ h2\ o)(Atoms(x) = \{h1, h2, o\} \wedge h1 \neq h2 \\
&\wedge HydrogenAtom(h1) \wedge HydrogenAtom(h2) \wedge OxygenAtom(o)), \\
\\
&Salt(b) \wedge x \in Molecules(b) \\
&\rightarrow (\exists na\ cl)(Atoms(x) = \{na, cl\} \wedge SodiumAtom(na) \wedge ChlorineAtom(cl)).
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
&Molecule(x) \rightarrow Mass(x) = \textstyle\sum_{y \in Atoms(x)} Mass(y), \\
\\
&HydrogenAtom(y) \rightarrow Mass(y) = 1.0, \\
&OxygenAtom(y) \rightarrow Mass(y) = 16.0, \\
&SodiumAtom(y) \rightarrow Mass(y) = 23.0, \\
&ChlorineAtom(y) \rightarrow Mass(y) = 35.5.
\end{aligned}
\tag{3}
$$

## APPEARANCE AND REALITY

• Getting reality from appearance is an inverse problem. Formulas and programs giving appearance as a function of reality and the circumstances of observation are easier to state and less likely to be ambiguous.

• Reality is more stable than appearance. Formulas giving the effects of events (including actions) are almost always written in terms of reality.

• The formulas that follow will need a situation or time argument once we consider changing appearances.

## FORMULAS—STARTING SIMPLE

• We begin with a little bit about touch rather than with vision. Imagine putting one's hand into one's pocket in order to take out one of the objects.

$$Touching(Side(1), x) \wedge PocketKnife1(x, Jmc) \rightarrow Feels(Texture17),$$

$$Texture(Side(PocketKnife1)) = Texture17 \tag{4}$$

For now we needn't say anything about $Texture17$ except that it is distinguishable from other textures. Textures for touch have similarities to and differences from textures for vision. Both are very scale dependent.

## THREE DIMENSIONAL OBJECTS

• How can we best express what a human can know and a robot should know about a three dimensional object? We start from a standard kind of object with particular types of objects and individual objects defined by successive approximations.

• I propose starting with a rectangular parallelopiped, which we'll abbreviate *rppd*. An object is an rppd modified by dimension information, shape modifications, attached objects, information about its internal structure, location information, folding information, information about surfaces, physical information like mass. Perhaps one should start even more simply with just a size, a ball too large to be included in the object and too small to include it.

- My small Swiss army knife is an rppd, 5cm by 2cm by 1.5cm, rounded in the width dimension at each end. Its largest surface has a smooth plastic surface texture, and its other surfaces are metallic with stripes parallel to the long axis, i.e. the backs of the blades. This description should suffice to find the knife in my pocket and get it out, even though it says nothing about the blades.

- Consider a baby and a doll of the same size. Each may be described as an rppd with attached rppds in appropriate places for the arms, legs, and head. The most obvious and significant differences come in a texture, motion, and family relationships.

## A PUZZLE ABOUT INFERRING REALITY FROM APPEARANCE

• Here's the appearance. The puzzle is: What is the reality behind the appearance? Clicking on the $<$ and $>$ signs is how one experiments.

• The reality is three dimensional, while the appearance is two dimensional.

• Those who implement display know that computing appearance is difficult. Those who do computer vision know that inverting the relation is even more difficult.

## HOW HUMANS SOLVE THE PUZZLE

• The appearance in the puzzle is a genuine appearance. The reality behind the appearance is rather abstract. Thus the bodies have no thickness or mass. This doesn't seem to bother people; we're used to abstractions.

• We use concepts like like *solid body*, *behind*, *part of*, *length*, etc.

• Some of these concepts may be learned by babies from experience, as Locke proposed. However, there is good evidence that many of them, e.g. *solid body* and *behind* were learned by evolution and are built into human and most animal infants.

• The quickest and most articulate human solution was by Donald Michie. Eventually machines will do better.

## FORMULAS FOR APPEARANCE AND ACTIONS

We introduce positions. There is a string of 13 positions. Bodies are also represented by strings of squares of length appropriate to the body. $Content(sq)$ is either a color or a letter depending on the version of the puzzle.

$$
\begin{array}{l}
Body(b) \wedge sq \in b \wedge Location(sq,s) = pos \\
\wedge (\forall b' \neq b)((\exists sq' \in b')(Location(sq',s) = pos \\
\rightarrow Higher(b,b'))) \\
\rightarrow Appearance(pos,s) = Content(sq).
\end{array}
\tag{5}
$$

$$
\begin{aligned}
&Body(b) \wedge sq \in b \wedge Location(sq, s) = pos \\
&\wedge (\forall b' \neq b)((\exists sq' \in b')(Location(sq', s) = pos \\
&\rightarrow Higher(b, b'))) \\
&\rightarrow (\forall sq' \in b)(Location(sq', Result(ClickCW(pos), s)) \\
&= CWloc(Location(sq', s))) \\
&\wedge (\forall b' \notin b)(Location(sq', Result(ClickCW(pos), s)) \\
&= Location(sq', s)).
\end{aligned} \tag{6}
$$

Here's the formula for the effect of counter-clockwise motion.

$$
\begin{aligned}
&Body(b) \wedge sq \in b \wedge Location(sq, s) = pos \\
&\wedge (\forall b' \neq b)((\exists sq' \in b')(Location(sq', s) = pos \\
&\rightarrow Higher(b, b'))) \\
&\rightarrow (\forall sq' \in b)(Location(sq', Result(ClickCCW(pos), s)) \\
&= CCWloc(Location(sq', s))) \\
&\wedge (\forall b' \notin b)(Location(sq', Result(ClickCCW(pos), s)) \\
&= Location(sq', s)).
\end{aligned} \tag{7}
$$

The last parts of the last two formulas tell what doesn't change.

## HOW SHOULD A COMPUTER DISCOVER THE REALITY?

• A point of view common (and maybe dominant) in the machine learning community is that the computer should solve the problem from scratch, e.g. inventing *body* and *behind* as needed. It is not dominant in the computer vision community.

• Our opinion, and that of the knowledge representation community, is that it is better to provide computer programs with common sense concepts, suitably formalized. There is some success, but the formalisms tend to be limited in the contexts in which they apply. I think, but won't argue here, that formalizing *context* itself is a necessary step.

• Here are two sample formulas relevant to the present problem but perhaps not general enough to be put in a *knowledge base* of common sense.

$$Color\text{-}Appearance(scene, x, s) = Color(Highest(scene, x, s)) \tag{8}$$

$$Behind(b2, b1, s) \wedge Opaque(b1) \rightarrow \neg Visible(b2, s). \tag{9}$$

• Solving the puzzle involves inferring formulas like

$$\begin{aligned}
&Body(b) \wedge Present(b, Scene) \equiv b \in \{B1, B2, B3, B4\},\\
&Color(B1) = Blue \wedge Color(B2) = Orange \wedge Color(B3) = Green\\
&\qquad\qquad \wedge Color(B4) = Red,\\
&Length(B1) = 6 \wedge Length(B2) = 8, \text{etc.},\\
&Higher(B1, B2) \wedge Higher(B2, B3) \wedge Higher(B3, B4),\\
&Higher(B4, Background) \wedge Length(Background) = 13.
\end{aligned} \tag{10}$$

• We haven't put in effects of actions and some relations among the predicates.

• The lengths and colors of the bodies are assumed not dependent of the situation. Human language tolerates *elaborations* such as actions that affect color better than do present AI formalisms.

• The ideas of the last two slides about what knowledge should be given to the program have benefitted from discussions with Stephen Muggleton and Ramon Otero.

## ENTITIES EXTENDED IN TIME

- The most obvious example is a tune. Maybe jokes, especially practical jokes, are another example.