

Old Sins & New Confessions

Patrick Hayes & Kenneth Ford

Many years ago, when AI was young, one of us (Pat) wrote a short article for the *AISB Newsletter* listing some sins to which AI work was at that time prone. Since those long-gone days, of course, our field has matured and grown wealthy beyond its wildest dreams, and these early faults are no longer committed by the kind of people who read this *Magazine*. To emphasize how far we have advanced, we reproduce the old list here. Our readers will immediately see that these early mistakes are now completely eliminated.

NINE DEADLY SINS

(courtesy of *AISB Newsletter* 20, 1975)

(1) Look ma no hands (first noted by J. McCarthy)

The sinner writes a program which does something which no program has done before, and writes a paper saying so; but does nothing else. (A sin of omission.)

(2) Plato

The sinner points out that Plato (Aristotle, Kant, Hegel, Wittgenstein, Hume, Sartre,...) has also tackled problems like those AI workers are tackling, and that he (they) might have something useful to say about it. He gives no actual quotes, however. People think the sinner is very erudite.

This sin has an obverse (which we call *Otalp*), where the sinner points out that he is working on the same problems as Plato (etc.), but claims to have got further than Plato (etc.) did because of the Power of Computer Science.

(3) Little boxes

The sinner writes a program which interfaces, or is intended ultimately to interface, to another program written by X. Performance is mediocre. The sinner points out that if only X's program worked several orders of magnitude faster, or was a bit smarter, all would be well, and expresses the devout hope that X will soon get it together.

Often, X was a theorem-prover: more recently, now that theorem-provers are dying out or migrating, X is working on some new programming language.

(4) Gee, what shiny bricks

The sinner is trying to get a program to see/hear/feel/smell some real world (e.g. a pile of bricks on a table). It doesn't do so well. So the sinner re-makes the world, so as to fit it to his program's abilities, e.g. by buying really bright lights, polishing the edges of his bricks, or speaking *very slowly*. Then it works better: the sinner claims a success.

(5) One day my son

The sinner is working in some fairly well defined

area of AI, maybe with some slight actual success. He points joyfully that *really substantial* progress in this direction will enable extraordinary feats of technology to be performed (such as robot cars, kettles that obey spoken commands, airline reservation systems that check political affiliations, etc.). Often, the sinner is writing an intermediate report to a funding agency.

(6) Got no time for you right now

The sinner is working hard in some AI area. He knows that there is nothing in the literature on his unique problem, so he doesn't read anything. He is far too busy to waste time writing anything. Anyway, all the people who are worth a damn are in the same laboratory anyway, so what's the point of talking to anyone else?

Sometimes, if the laboratory is well funded, whole academic careers can be built in vacuo.

(7) Tablets of stone

The sinner wants to represent knowledge about something in a program. So he invents a new notation for writing the knowledge down. This formalism is similar in expressive power to many formalisms, but it's the sinner's own baby, so he publishes some papers about it. Eventually, he gets his own students working on it, and a whole literature builds up around this particular formalism.

Nobody tries to compare or contrast the various formalisms, since each proprietor has his own appearances to keep up.

(8) Rally round the flag boys

The sinner is persuaded of the essential truth of a slogan, often invented by one of his colleagues. He attends conferences and meetings and vehemently criticises what other people are saying, because it does not conform, in surface structure, to the slogan. He does not attempt to reconcile apparently opposing views, or engage in discussion to discover whether surface disagreements are merely verbal or reflect a real disagreement in content.

(9) Its all really only

The sinner claims that everything (or almost everything) in AI is "really" only a special case of some particular mechanism or technique *M*. In the past, *M* was often some kind of heuristic search process. More recently, *M* is some kind of non-deductive reasoning engine (keywords: analogy, matching, actors, controlled inference, frames, Frames, ...).

This is perhaps a venial rather than a mortal sin, since it *can* lead to genuine insight. The mortal sin is in actually *believing* it. As Minsky says, we will proceed by pretending that such a theory exists.

These sins are particularly bad in combination. For example, a system consisting of a central figure committing *Tablets of stone* and *Its all really only*,

surrounded by students/disciples committing *Rally round the flag boys*, is almost immune to all outside influences, and can be treated in only two ways: you can let it take you over, or you can refuse to communicate with it at all.

By the way, I hereby confess to having committed *Plato*, *Little boxes* and *Its all really only* at various times in the past, present and (probably) future. Confess your sins, O my sisters and brothers, that they shall be lifted from your shoulders.

This final exhortation did produce results at that time, and several of the confessors indeed reported scapular levitation and a new spring in their step. Searching (so to speak) for a way to make this rite of renewal available to a new generation, we have implemented a confessional agent. ELIJA is not very smart, but it is sincere. It believes that it can forgive anything, it desires to forgive everyone, and its intention is to forgive whoever it is talking to. It will forgive any sins you tell it about, if you can manage to persuade it that you have actually described the sin. If ELIJA asks you to "tell me more about" your sins, don't be reticent: let it all hang out. You will know you are forgiven when ELIJA tells you "go forth and sin no more."¹

Please do not try to persuade ELIJA that you are lying to it. While lying is undoubtedly a sin, creating a paradox in a knowledge-base is a form of damage, and damaging a piece of software across a state line is a federal crime according to the Computer Fraud and Abuse Act, 18 U.S.C. § 1030 (amended Oct 11, 1996) sec (a) para. 5(A). The Secret Service (see 18 U.S.C. § 1030 (d)) are, unfortunately, less forgiving than ELIJA. Obviously, things are more complicated than they were in 1975.

(Oh, and one other thing: rereading the old piece now, we are struck by its careless use of the male pronoun. Of course when Pat wrote "he," he meant "he or she." Sorry about that.

Things are more complicated than they were in 1975. Sigh.)

Note

1. This claim has not been warranted by a government agency or registered Church. ELIJA is not known by the State of California to contain anything that might cause anything. Whatever happens to you after talking to ELIJA, it's not our fault, OK? "ELIJA," "sin" and "forgive" are registered trademarks, copyright ScapLev, Inc. Do not use without permission.)