RAISING THE STANDARDS OF AI PRODUCTS

Alan Bundy
Department of Artificial Intelligence
University of Edinburgh

Richard Clutterbuck
School of Social Sciences
University of Sussex

Keywords
Code of practice, association of companies, Artificial Intelligence, social implications, legal implications.

Abstract
We propose a mechanism for the promotion of high-standards in commercial Artificial Intelligence products, namely an association of companies which would regulate their own membership using a code of practice and the precedents set by previous cases. Membership would provide some assurance of quality. We argue the benefits of such a mechanism, and discuss some of the details including the proposal of a code of practice. This paper is intended as a vehicle for discussion rather than as the presentation of a definitive solution.

Acknowledgements
We are grateful to members of the Edinburgh Computing and Social Responsibility Group for feedback on an earlier draft of the code, and to Maggie Boden for comments on the paper itself.

1. The Need for High Standards in AI Products

Credibility has always been a precious asset for AI, but never more so than now. The current commercial interest in AI is giving us the chance to prove ourselves. If the range of AI products now coming onto the market are shown to provide genuine solutions to hard problems then we have a rosy future. A few such useful products *have* been produced, but our future could still be jeopardised by a few, well publicised, failures.

Genuine failures - where there was determined, but ultimately unsuccessful, effort to solve a problem - are regrettable, but not fatal. Every technology has its limitations. What we have to worry about are charlatans and incompetents taking advantage of the current fashion and selling products which are overrated or useless. AI might then be stigmatised as a giant con-trick, and the current tide of enthusiasm would ebb as fast as it flowed. (Remember Machine Translation - it could still happen.) Both companies selling AI products and academic AI research groups would suffer in the resulting crash.

AI companies are very dependent on the good-will of their customers. The current life-span of typical AI products is about 1-5 years. The customers of AI products are likely to stay in the market for several times this period; typically they are themselves companies or academic groups engaged in AI research or interested in the long term application of AI techniques. To stay in business the AI company must sell successive upgrades of its products to the same group of customers and, therefore, must build up and maintain a good reputation. If AI business is to expand then new customers must be brought into this existing group. This will only happen if the overall range of AI products is of high-quality and the reputation of this particular company is good. Thus it is in the interests of each company to raise both the general and its particular standard. It must also convince customers that its products are of high standard. When the market was small a company with high-quality products could win new customers by word of mouth. Now the market is growing they must use advertisements, and it becomes harder for a company to convince potential customers that its products *are* of high quality.

Apart from improving the public image of AI and increasing the market for AI products, producing more high-quality products would raise morale and standards in AI itself, leading to a virtuous circle of standards being raised, better work being done, good people being attracted to the field, and even more high quality products emerging. Poor-quality products will produce a vicious circle going in the opposite direction.

But these internal reasons for wanting high standards, while important to insiders, are perhaps less important than external reasons. AI products look destined to play a major role in society. That society deserves, and has the right to expect, protection from exploitation by AI companies and from being harmed by AT products.

An extreme, potential example of such harm is described in [Thompson et al 84], which argues that it is not possible to build an automatic or semi-automatic launch-on-warning system for nuclear weapons with anything like an acceptable failure rate. Anybody who claimed to have done so, or who claimed to be able to do so, would be guilty of misleading the public in a way that could have disastrous consequences. If such a claimant were an AI company then the whistle might be blown on it by the mechanism described below.

However, the main purpose of this proposal is to catch less apocalyptic, but more common-place, mis-

leading claims, whether or not they might give rise to a legal remedy. Examples might be an expert systems shell whose advertised range far exceeds the problems it is really suitable for, or a natural language front end which is presented as being able to deal with a much wider input than it, in fact, can.

2. A Professional Association

The academic field guards itself against charlatans and incompetents by the peer review of research papers, grants, PhDs, etc. There is no equivalent safeguard in the commercial AI field. Faced with this problem other fields set up professional associations and codes of practice. AI needs a similar set-up. We propose that the responsible AI companies should get together now to found such an association. Continued membership should depend on a constant high-standard of AI products and in-house expertise. Members would be able to advertise their membership, and customers would have some assurance of quality. Charlatans and incompetents would be excluded or ejected, so that the failure of their products would not be seen to reflect on the field as a whole nor on the companies in the Association.

Since the trade in AI products is international, with multi-national companies involved both as vendors and customers, the Association would also need to be international. Otherwise, there would be difficulty over membership of multi-national companies and about dealing with complaints resulting from international sales. It is particularly important to make membership attractive to multi-nationals because, with their existing reputation, they have less to gain from the cachet of membership and are more able to avoid the full impact of national registration.

The Association would be self-regulating. If its decisions were seen to be too arbitrary then the value of Association membership would be devalued in the eyes of the public, the customers and the vendors, and the importance of its decisions would decrease in proportion; customers would take no account of Association membership when deciding to buy, so vendors would not bother to join. For such self-regulation to work it is necessary for the Association to be publically visible. Both vendors and potential customers must be aware of the Association and must see its decisions as fair and its sanctions as effective. Customers will then use Association membership as a major determinant when deciding whether and what product to buy. Vendors will regard Association membership as a valuable asset to their company, and will aim for high quality in their products in order to retain membership. They will want to use Association membership in their advertising, and this will, in turn, improve the visibility of the Association. It will be necessary for the Association to maintain a high profile of both its existence and its actions, to be open about its decisions, and to employ effective sanctions.

The Association would need a panel to consider applications for membership and to hear complaints against members. Its main sanction would be refusing membership or expelling existing members, backed up by lesser sanctions like a public admonition, payment of compensation, etc. The rules of the Association might include a contractually binding committment by members to fulfil any compensation order made by the panel. The Association might also insist that contracts issued by members contained various standard clauses, e.g. giving customers the right of reimbursement if returning products within a certain period, guaranteeing compensation under certain circumstances, insisting that precise, testable statements be made about the product, etc."[2] The panel need only take a passive role in determing disputes; it would publicise its address and its willingness to hear complaints. The burden of making a case complaint would fall on the complainant. The panel might then need to employ a small team of experts to investigate discrepancies between the evidence brought by the complainant and that by the company complained of. This investigative team could be recruited on an ad hoc basis, e.g. academic researchers as consultants. The complainant would usually be the dissatisfied purchaser of one of the companies products, but could be any member of the public with a legitimate interest in the product. There might be a multi-stage process, so that cases were only heard by the full panel when a prima-facie case had been established. This is to filter out malicious complaints and those that are outwith the remit of the panel. It would not be necessary actively to investigate AI products or companies before any complaint had been received, and would probably be prohibitively expensive to do so.

The panel needs a code of practice to which members would agree to adhere and which would serve as a basis for applying sanctions. What form should such a code take, i.e. what counts as malpractice in AI? We suspect malpractice may be a lot harder to define in AI than in insurance, architecture or travel agency.

- Due to the state of the art, AI products cannot be perfect. No-one expects 1002 accurate diagnosis of all known diseases. On the other hand a program which only works for slight variations of the standard demo is clearly a con. Where is the threshold to be drawn and how can it be defined?

- It is unlikely that any current AI product could fulfil a claim to: understand any natural language input, or to make programming redundant, or to allow the user to volunteer any information what-so-ever. However, the claimant could defend the

[1] E.g. the Vehicle Builders and Repairers Association and the National Association of Estate Agents. Note that, unless otherwise stated, all examples are of UK institutions or laws.

[2] C.f. the Vehicle Builders and Repairers Association which has standard forms for estimates and recommends clauses in repair contracts.

claim by debating the meaning of 'understand'. 'programming', or 'information', and this would muddy the water. What constitutes an exaggerated claim?

- Given the ambiguity of such terms it would be difficult to decide whether an exaggerated claim was intended to deceive or was due to a difference in terms or was just a genuine oversight on the part of the vendor. How is a vendor's claim to be assessed? Should one try to assess the vendor's intention, or should one ignore this and only assess how a reasonable customer might interprete the claim?

- Because of the ambiguity of such terms the full description of an AI product must describe its limitations, e.g. what sentences it cannot understand, as well as its abilities. It is not enough to refrain from false claims.

- A vendor may claim that it cannot accurately describe the limitations of its product without revealing confidential information about the technique and/or software it is based on. The problem is particularly acute for software products because of the lack of protection afforded by patent and copyright law. Where do we draw the line between a complete and accurate description of the capabilities of the product and the protection of trade secrets?

The difficulty is to give a precise definition of what constitutes reasonable behaviour on the part of a vendor. It seems impossible to cover all the possible situations, in advance, with a list of precise standards to be attained, but it is often possible retrospectively to detect unreasonable practice in particular cases. The usual legal solution to such problems is to use a high-level code in combination with judgements about individual cases in order to build up gradually a picture of *the reasonable* vendor, i.e. to establish *case law',* and to use this to evaluate complaints rather than to pre-vet products. Note, however, that it takes a long time to build up an extensive range of cases - most of the early judgements must be made solely on the basis of the high-level code.

The panel would evaluate a complaint against this code. They would be able to take account of the state of the art and compare the product with the claims made for it. A high-level code and the injunction to judge 'reasonableness' would enable the panel to assess whether the spirit of the code had been broken, rather than the letter of some spuriously precise, low-level code. The accumulation of case 'law' would ensure some uniformity of treatment, and prevent favouritism or victimisation. Openness about the grounds for decisions would also help ensure uniformity. There is a current trend in other areas towards the giving of reasons - a procedural safeguard which promotes the the quality of decision making.[5] Difficult decisions, like those outlined above, would be decided in particular cases. rather than in general. The general answers would emerge over time with the accumulation of judgements. Previous ajudications would guide future ones without pre-empting them. The case 'law' would provide a guide to vendors as to how to practice reasonably.

3. A Proposed Code of Practice

As a basis for discussion, we propose below such a code of practice for AI vendors. Before we give this we must define our terms. In what follows below, the term:

- AT *product,* means any piece of software or hardware or any service or any combination of these which is based on AI techniques and which is offered for sale;

- *vendor,* means a company selling an AI product, either to a customer direct or to a middleman, whether that company made the product or not;

- *customer,* means a person or group who buys or attempts to buy an AI product from a vendor;

- user, means the person who uses the AI product, in particular the person who interacts with the product if it is interactive.

The proposed code \$:

> The vendor of an AI product should describe to the customer, and where appropriate the general public, the abilities and limitations of the product as accurately as possible, taking account of the likely expectations of the intended customer. In particular, the vendor should accurately describe, in so far as the state of the art and its own knowledge enables this to be done:
>
> 1. what the product does, including an account of its scope, limitations and reliability;
> 2. known bugs in the product;
> 3. the consequences of failure of the product;

Law is in scare quotas bacause wa ara dafining an axtra-lagal mechanism.

Compare, for instances, the Unfair Contract Terms Act 1977. which uaaa caaa law to dafina raaaonablanaaa in ralation to exclusion clauaaa.

Saa a.g. tha Criminal Justice act 19B2.

4. the amount and type of user interaction required and how and at what cost it is to be obtained;

5. the skill and knowledge required of the user;

6. the computational requirements of the product, e.g. hardware and software environment. space and time requirements in different environments;

7. the amount and type of maintenance required and the cost of this;

8. any social, economic or legal implications of the use of the product, where these can be assessed.

It is the responsibility of the vendor to see that this code is observed by any agent acting on its behalf, e.g. a salesman. The vendor is also responsible for ensuring that any middlemen which sell its products are fully acquainted with the necessary information to enable them to comply with the code.

An informed customer will, in any case, ask about 1-7 above, and a reasonable vendor should supply the information unasked. Thus this part of the code merely makes good practice explicit, as it was intended to do. In conjunction with the case 'law', it should help protect the uninformed customer and define the standards to be met to become a reasonable vendor.

Point 8 is rather different from the others. It was inserted to try to protect the wider interests of society as well as those of the customers. It might be criticised as being impractical to realise *or* as not appropriate in this context. However, we feel that something like it is required somewhere, and we would welcome suggestions as to how best to meet this requirement. Maybe it needs to be dealt with by separate machinery.

It is not our intention that vendors be required to state political or ethical opinions, nor that the Association be asked to judge such opinions; it would be beyond their competence to do so. In point 8 we wanted only to encourage vendors to make statements which were within their technical, legal, etc. competence so as to enable others to form accurate political and ethical opinions about the impact of the product. For instance, all the photocopiers at Sussex University have a prominent notice above them detailing the law relating to copyright. Vendors of AI products should, similarly, draw the attention of users to illegal uses of their product.

It is not intended that complainants actually intend to be customers of the vendors they complain of, but they should have a legitimate interest in the product over and above commercial competition. For instance, suppose an expert in the field believes that a vendor is misleading customers about a product or that the vendor is producing a product that will be harmful to society; we would like that expert to be able to bring a case to the Association. On the other hand the Association would have to be alive to attempts by vendors to undermine their rivals by bringing malicious complaints - and should filter out such complaints at an early stage.

4. Related Codes and Laws

The above proposal is complementary to the existing system of codes and laws applying to AI.

For instance, in the UK, the British Computer Society and, in the USA, the Association for Computing Machinery both provide codes of conduct for their members, [BCS 81, ACM 8?]. We imagine that most other national computing societies have similar codes. Neither of these codes apply to AI products or vendors, as such. There is a small area of overlap in that a salesman who is a member of the BCS or ACM is required to behave honestly and competently in promoting a product. However, any sanctions for breaking the BCS or ACM codes would fall on the individual rather than the company. Our code is intended to apply to vendors, which would usually be companies rather than individuals.

In the UK, the Trade Descriptions Act uses criminal sanctions to protect customers from false claims, and the Sale *g±* Goods Acts and the Unfair Contract Terms Act 1977 permit civil remedies for defective goods. Other countries have similar laws. However, as illustrated above, there is a large grey area between illegal behaviour and the behaviour of a reasonable vendor. It is the purpose of the above proposal to deal with this grey area, where the vendor has clearly behaved unreasonably, but not in such a way as to constitute a criminal offence or grounds for a civil action. For this reason it would not be appropriate for the proposed Association to impose the kind sanctions that would be imposed by a court of law. An unreasonable vendor should still be allowed to trade, but should not be allowed to use the cachet of membership of an association of reasonable vendors, with whatever assurance of high-quality that that was generally felt to imply.

The British Standards Institute defines standards for many products. Vendors whose products meet these standards are able to advertise that fact with a 'kite mark'. In some cases the standards set have been adopted by the law.[6] There is currently an attempt to define a BSI standard for Prolog. Unfortunately, few AI products lend themselves to such definitions of standards. It is not worthwhile to try to define one unless very similar products are being produced by a number of vendors and there is a wide agreement on a de facto standard. Major and stabilised programming languages seem possible candidates, but customised expert systems, and even expert systems shells, do not.

[6] E.g. motorcycle crash helmets must be worn and must comply with a minimum BSI standard.

5. A Discussion of Problems

There is a danger of a few companies annexing the Association to themselves and excluding worthy competition. But this is not a major danger. Firstly, in the current state of the AI market, AI companies have a lot to gain by encouraging high-quality in other AI companies. Every success increases the market for everyone, whereas failure decreases it. Until the size of the market has been established and the capacity of the companies has risen to meet it, AI companies have more to gain than to lose by mutual support. Secondly, excluded companies can always set up a rival association. There is room for more than one association, and they could compete by trying to set the highest standard for membership. However, too many associations would be confusing to consumers and would diminish their influence and effectiveness.

There is also a danger of the Association developing into a trade protection society, i.e. of maintaining low standards by protecting its members from disgruntled customer by offering weak excuses for faulty products and by not employing effective sanctions. If this happened then customers would lose confidence in the Association and membership would cease to carry any assurance of high-quality. Membership would still be attractive to vendors who wanted to use the Association's excuses to fob off disgruntled customers. However, provided the Association did not have a monopoly, there would be nothing to stop a more principled group of vendors forming a rival association as above.

The Association would always be at risk of legal action against it from disgruntled vendors who might sue for libel. It would need to take care that its pronouncements were fair comment' and to take legal advice on them. It should also try to create the conditions under which vendors would have more to lose from the bad publicity accruing from the court case than they would gam from any damages awarded. But to guard against such actions, the Association would have to maintain a legal defence fund contributed by the members.

6. Passive vs Active Role

We have proposed that the Association take a passive role, reacting to complaints, rather than an active role, pre-vetting products and/or instigating investigations of products with its own team of investigators. Our reasons are practical rather than principled. An active role would require money and people. Most AI companies are small and newly set-up; they might be loath to provide the large sums of money required to set up the investigative machinery, but might be prepared to fund a, much cheaper, passive association. AI experts are currently rare and expensive. Most such experts want either to conduct their own research or set-up their own companies. It would be hard to recruit good full-time investigators. One might find academics prepared to work part-time, but most candidates would have some existing consultancy arrangements that might disqualify them. We have also argued that, while the field is still immature, it is more difficult to set standards that a pre-vetted product must meet than to evaluate the complaints of a customer against the code of practice and previous cases.

For the same reasons we have not proposed the direct registration of AI products. To issue a certificate of good quality to an individual product would require a prior investigation of that product; it would no longer be possible to employ the default assumption that the product of a member of the Association was assumed good unless proved otherwise. That is, direct registration of products would require an active Association with all the associated expense, employment of rare expertise and setting of prior standards. The passive mechanism proposed above indirectly ensures good quality products by encouraging the vendors to produce products that will not attract complaints. The burden of criticising the product and proving that criticism falls on the complainant. The panel need only investigate differences in the evidence presented to them by the complainant and vendor about the product complained of; this investigation would be relatively cheap compared with that required to register every product.

Unfortunately, this means that the burden of proof falls on the customer. But this is not as bad as it seems; currently, most AI customers are themselves AI practioners, to some degree. For instance, the customers for expert systems shells and knowledge representation systems are often researchers from the AI laboratories of other companies. Hence, they are in a position to investigate the product, and bring a complaint. They only lack a body to bring it to.

All this might change: AI companies might get richer and more able to fund an active association. AI experts might get thicker on the ground, standards might get better defined as the field matures, the average AI expertise of customers might decline as the customer base expands. In this case, there is nothing to stop the Association moving to a more active role.

7. Relationship to the Law

We have proposed a extra-legal regulatory mechanism, i.e. one without legally enforceable sanctions. One reason is that we are aiming to regulate *in* the grey area of legal but unreasonable practice. Another reason is that we would like to see an international association covering countries with different legal systems, so no one legal framework can be assumed. A third reason is that we want to make it as easy as possible for the Association to get started; an extra-legal mechanism involves the minimum of bureaucratic hassle.

This proposal, however, creates problems. The boundary between legal and illegal practice is fuzzy; infringement of some aspects of the code of practice (e.g. points 1-3) may sometimes give rise to legal remedies. Even when legal remedies were possible there may still be situations in which both complainant and vendor would prefer to refer the matter to the Association because it provided a cheaper and quicker mechanism for settling the dis-

pute. However, if cases were referred for legal remedv then the workings of the Association might be undermined; such cases would become sub-judice and the Association would no longer be able to comment on them until they were decided (which might take a long time).

However, as it develops the Association might seek statutory authority in the countries in which it operated. This authority might include, for instance:

- the right to grant licences enabling companies to become vendors of AI products;
- protection from libel action in its judgements;
- that infringement of the code of practice would raise a presumption of fault against an infringing vendor in any legal action.

Note that such legal powers increase the dangers of the Association becoming a trade protection society. In particular, the right to grant licences gives a monopoly that might be used to exclude competition. Therefore, it would need to be offset with rights to the customer and the vendor to prevent abuse, e.g the vendor might have the right of legal action if unfairly excluded from membership, the Association might be required to pre-vet and then underwrite the products of its members, so that dissatisfied customers could sue the Association.

8. Conclusion

In this paper we have proposed a mechanism for peer-policing of standards in AI products. The mechanism consists of an association of vendors of AI products who would use a code of practice to sanction vendors who are guilty of unreasonable practice. It is similar to associations used by other groups of companies and professionals offering services or selling products. It is necessary if AI is not to harm itself and society by the products it produces. Some such mechanism is vital to the existing, responsible AI companies if they are to protect their investment in AI; they must try to prevent a few irresponsible companies from exploiting their customers with overrated or useless products and putting those customers off of AI products in general.

Most of the discussion above is relevant to computing generally. We have limited ourself to AI because of the nature of this conference and in order to help focus our ideas and to identify an area small enough to be tackled. AI is growing fast, but there are not yet many major AI companies, and it seems more likely that *they* would be prepared to get together than that IBM, ICL, Honeywell, etc. would be. However, since many AI products will also contain non-AI techniques, it may eventually be necessary to widen the remit of any AI association.

The above mechanism is only able to regulate commercial companies. We have argued above that academics already have a self-regulatory mechanism of peer review. However, there is major gap in that neither mechanism covers government organisations, e.g. the military. That is beyond the scope of this paper, but not beyond our desires.

This paper is intended as a vehicle for discussion of the problems of maintaining high standards in AI products. It does this by presenting a proposal, but this proposal is not intended as definitive. We would welcome feedback - especially on the code of practice itself. We hope that this paper will inspire those companies that care about high standards to get together, and we hope that they will see that it is in their direct interest to do so.

References

[ACM 82] Association for Computing Machinery. ACM Code of Professional Conduct. *Communications* of *the ACM* 25(3):183-184, March. 198?.

[BCS 81] The British Computer Society. Code of Conduct. 13 Mansfield St. London W1M OBP, 1981. Handbook No. 5.

[Thompson et al 84] Thompson, H. and the Edinburgh CSR group. There will always be another moonrise: computer reliability and nuclear weapons. *The Scotsman* , 1984.

See, e.g. the statutory disciplinary powers of the Law Society over solicitors.

See. e.g. that an infringement of the Highway Code raises a presumption, albeit rebuttable, that the driver of a vehicle involved in an accident is at fault.