Socially Sensitive Computing:
A Necessary Paradigm Shift for Computer Science

Thomas R. Addis[1,a], Bart-Floris Visscher[a], Dave Billinge[b], David C. Gooding[c]

[a] *Department of Computer Science and Software Engineering, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth PO1 3AE*
[b] *Department of Creative Technology, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth PO1 3AE*
[c] *Science Studies Centre, Department of Psychology, University of Bath, Bath BA2 7AY, U.K*

"Rules are for the obedience of fools and the guidance of wise men"
Douglas Bader (1910 – 1982)

Abstract

In response to the grand challenge for computer science during a workshop held in Edinburgh in November 2002 we identified an essential problem in computing that has yet to be addressed directly, see Appendix A for UKCRC Grand Challenge criteria (Addis et al., 2004). We show that this problem, originally identified by Wittgenstein circa 1945 (pub. posthumously, revised translation 1966), explains a barrier that prevents people from communicating seamlessly with computer systems. It explains many of the apparently insoluble problems that beset human computer usage such as; context dependency in natural language understanding, the generalisation problem in machine learning/neural networks and effective data retrieval. Although many solutions have been attempted, nobody has yet directly addressed the underlying cause. This cause can be characterised by the notion of the ubiquitous existence of irrational sets that emerge from the fundamental nature of human language and a continually changing knowledge of the world. We suggest two possible routes to solving this problem.

Keywords: Paradigm shift; HCI; Socially sensitive computing; Irrational sets; Inferential Semantics; Natural language; Pragmatics; Dual semantics; Complexity; Informal; Purpose; Common sense; Indefinite world;

1. The Essence of the Challenge

We argue from the Church-Turing Thesis (Kleene, 1967) that a program can be considered equivalent to a formal language similar to predicate calculus where predicates can be considered as functions. We can relate such a calculus to Wittgenstein's first major work, the *Tractatus* (Wittgenstein, 1921), and use the *Tractatus* as a model of the formal classical definition of a computer program. However, Wittgenstein found flaws in

[1] Corresponding author. Tel +44-23-92355721
E-mail address: tom.addis@port.ac.uk

his initial great work and he explored these flaws in a new thesis described in his second great work; the *Philosophical Investigations* (Wittgenstein, 1966). The challenge we make is "*can computer science make the same leap?*"

We are proposing that because of the flaws identified by Wittgenstein, computers will never have the possibility of natural communication with people unless they become active participants of human society. We will come to this conclusion by examining the way people communicate and by considering the two major works on the philosophy of language by Wittgenstein. We suggest that such a lack of natural communication is evident by the common complaint of 'computer rage' and that the same paradigm shift that Wittgenstein took is a viable way forward for Computer Science.

2. Inferring Internal Experience

Classical linguistic philosophy suggests that programming language understanding arrives from denotational (referential) semantics (Watt, 1991) and the same was suggested for natural language through syntactic (e.g. transformational) structures mapping onto semantic structures (Chomsky, 1964). These semantic structures are constructed from predicates, and other primitives, chosen such that each construct could only have a single meaning (Katz and Fodor, 1963); there could be no ambiguity. Computer science has used this approach since its suggestion in 1963 in an attempt to create natural language understanding. These attempts have been unsuccessful except within restricted domains and with particular linguistic paradigms (Addis, 2003). It became a challenge in the 70's & 80's to propose counter paradigms for each attempt to show its limitations and dismiss any possible claim for a general solution. In most cases the mechanism by which the semantic primitives derived meaning was not questioned: it was enough that they had a single interpretation, although how this singularity of meaning might be determined formally was similarly rarely addressed. The only "formal" mechanism, we can assume, was referential.

We have shown in an experiment that people can hold conversations that are descriptions of their internal life (Billinge and Addis, 2003, Billinge and Addis, 2004). People seem to achieve this transfer of meaning through the use of metaphor and an appeal to common human experience (Lakoff, 1986, Lakoff and Johnson, 1980). If, for example, the non-technical music literature is examined[2] it becomes evident that the common experience does not even have to be the music itself in order for one person to describe an experience to another. The rich and extensive use of metaphor suggests that emotional resonance and association to a commonly understood situation can be employed to trigger what, to the author of the description, is his 'accurate' emotional response to a piece of music. Communication, in this case, will depend mostly upon our shared humanity, sometimes upon our personal experiences but unlike computers, little upon any referential semantics. Since nobody can have direct access to another's internal experiences, then the only way in which such experiences can be understood is indirectly. We suggest that each of us can infer another's experience because we share the state of being a person, in a culture, using a language and sharing external experiences. It is hence possible through conversation to build an internal model of another person's view

---

[2] Examples are, record reviews, concert reports, descriptive, as opposed to analytical, music histories and biographies.

of the world. The only requirements for this model is to be able to make predictions from conversations about:

- one's own possible future experiences
- the way one should respond to another person
- an interpretation of what is said by that person (or even by others)
- new ideas and ways of looking at the world

3. A Philosophical Paradigm and Computing

The implications of such observations on the communication of internal experience are radical. They have led us to take Wittgenstein's *Tractatus* [Wittgenstein 1921] as a paradigmatic description of the current state of computer science. We can take this step because the Church-Turing Thesis shows that the Turing Machine (the classical computer) is equivalent to Lambda calculus and recursive functions. Lambda calculus and recursive functions together form the description of a functional programming language (e.g. ML). Such a functional language embodies all seven of Wittgenstein's theses.

The *Tractatus* was modelled on Hertz' *Principles of Mechanics* [1956 (first published in 1900)] (T4.04, T6.361). Hertz believed that his book would be a full and final statement of the principles of mechanics; Wittgenstein thought that Frege, Russell and Whithead had done the same for mathematics and that he would do the same for language" (Gooding, 2004a). This early work encapsulated a formal and logical representational schema into a descriptive form that was based upon denotational (or referential) semantics. In this case, the referents (the objects) have some logically necessary properties (Addis, 1980). In our discussion we will refer to paragraphs in the *Tractatus* by the prefix T and those in his other great work the *Philosophical Investigations* by the prefix P. In order for denotational semantics to work the objects that are referenced must be:

- *independent* in that they can freely combine to form "states of affairs" that can be described (T2.01, T2.0272, T2.0122, T2.0124). The language should not be artificially limited by what it can depict by its syntax.
- *atomic* in that there are no smaller constituents (T2.02, T2.021). Otherwise it would be structured and such structures are statements that have sense drawn from the components that either have sense through construction or meaning through reference.
- *in all possible worlds* (T2.022, T2.023) in that there can be no conceivable world that does not have these objects in them for reference. If there is a world that does not contain at least one of these objects then it cannot be described. Since it cannot be described then it cannot even be thought about. This is expressed simply by T7 "What we cannot speak about we must pass over in silence".
- *immaterial* (T2.0231, T2.0233) in that the object is not a physical notion in the same sense as a chair is a physical object. He suggested that they might be physical in the same way that Force or Mass are physical in that we see their effect on the world but not their substance. In contrast, a chair has structure and cannot be an atomic object; it is however a named proposition.

- *indescribable* except by their behaviour (form) (T2.021, T2.0121, T3.0271) since this is all that is left. For example, Force and Mass are only known by the way they interact with each other. They cannot be conceived of in isolation. Wittgenstein saw these as candidates for being referents.
- *self governed* in that they have their own internal rules of behaviour (T2.0141, T2.033, T2.012, T2.0121, T2.0123, T2.01231, T2.03) metaphorically similar to atoms bonding according to valencies.

These referents (objects) are intended to be more than just elements of description; they *form* the real world (Wittgenstein, 1921; T2.04, T2.06). From these referents, the full force of Logic, Predicate and Propositional calculus retains stability of meaning and sense. Such a stance results in the position that *everything* is potentially unambiguously describable (Wittgenstein, 1921; T2.225, T2.224, T7).

We introduce a distinction between 'rational' and 'irrational' sets. The distinction is analogous to Cantor's (Aczel 1996 p23) irrational numbers (Boyer 1949 pp287-298). In the case of rational numbers the rule was that a member number could be expressed as a ratio of integers. However in contrast the irrational numbers have no such clear definition. Examples of irrational numbers are $\sqrt{2}$ and π. It can be shown that there are infinitely more irrational numbers than rational numbers. However, as for irrational numbers an irrational set can always be approximately represented by a rational set in the same way as the value of π can be $^{22}/_{7}$ for certain purposes. Different purposes or contexts will require different approximations.

A ***'rational' set*** *is a set where there is a finite set of rules that can include unambiguously any member of that set and unambiguously excludes any non-member of that set* (see also the definition of irrational sets – section Paradigm Leap).

It should be noted that all the sets referenced by the *Tractatus* are rational where set membership is always specifiable and context independent or has an explicit context that is also rational. This is because there can be no ambiguity about object assignment (denotation) or the semantic relationships that bind objects into descriptions.

The *Tractatus* provides an extensive model of computer languages. The argument is that names (in practice signs; the visible part of an expression or name) in propositions do not always refer to primitive objects but are themselves referencing propositions (Wittgenstein, 1921; T3.14, T3.31, T4.03, T4.22, T4221, T5.135, Wittgenstein, 1966; P43-P60). These, in turn, are complexes that finally end up as compound statements whose ultimate referent is the bit. Here the bit is the mechanical equivalent of Wittgenstein's referent objects. The bit, if taken as a detectable distinction, has all the strange properties of Wittgenstein's object. For example, a world cannot exist (or at least be detectable) unless it contains at least one distinction. A 'bit' is a concept that can only be embodied in a distinction but it has itself no definitive physical form. A particular 'bit' is referred to as an argument place in the *Tractatus* (T2.0131). So the concept 'bit' can be used for many purposes and in many descriptions. Bits can be represented physically by differences in voltage, current, height, length or just the position of beads on a wire.

Further, it is at the bit that the program links to the world and has meaning. It is this meaning that allows the program to have “sense” with respect to the computer. This formal semantics and the ability for programmers to create procedures and sub-routines (sub-propositions or expressions) is the primary characteristic of all high level and assembler programming languages.

The consequence of such a formal model is that any set of names can be used in a program to represent a proposition. All that is necessary is that there is a formal definition that gives the name meaning within the program in terms of the proposition it represents. Since a proposition can take on an infinite number of forms through the use of tautologies and other formal equivalences then there is an infinite but bounded set of possible organisations that can be adopted for a program. Such a set is bounded by the meaning of the essential program (the base or minimum program). However, the additional adopted structure is also represented, in the end, by bits on a computer. This will appear as a program overhead that is used to support a chosen program organisation or structure and in this sense only the program interpretation has changed.

There are also social consequences of the view adopted by the *Tractatus* in that it is assumed that rules can be created for all situations and as such these rule can bypass human judgement and are superior to it. It also assumes that there is only one correct way of seeing the world and so human existence can be governed by some finite set of laws.

4. Dual Semantics

One of the problems we address here is that computer languages have a *dual semantics* in that the program signs (e.g. the names/labels given to data items, procedures and sub-routines) at the highest level also have referents in the world (figure 1 – the Problem Domain). This is generated by the analysis of the problem domain in terms of records (as in database and program structures), relations (as in normalised data structures) and objects (as in object-orientation). This analysis identifies constructs in the world that are meant to be stable and unchanging (as per *Tractatus* referents) to which names can be given and meaning assigned. This is how names/symbols denote meaning in the world.

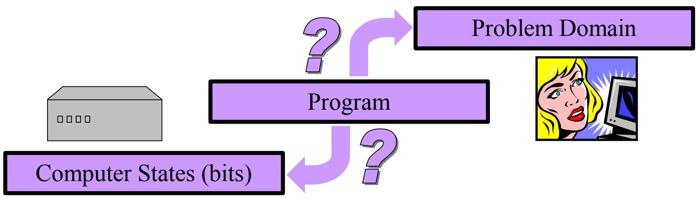

Figure 1 The problem of dual semantics

It is acceptable that propositions can represent material properties (Wittgenstein, 1921; T2.0231), relationships (Wittgenstein, 1921; T2.031), and any complex model of the world (Wittgenstein, 1921; T3.1, T3.11, T3.32, T4.01, T4.021) *but a proposition can have one and only one complete analysis* (Wittgenstein, 1921; T3.25). These propositions form semantic structures of the kind identified by Katz and Fodor (1963). Such an analysis is dependent upon only the essential features of the proposition (program) that link it to the referent objects (the bit in the case of programs).

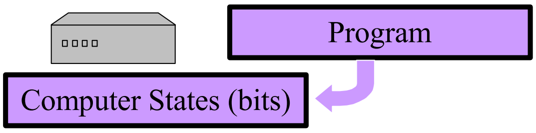

Figure 2 The only rational interpretation of a computer program

A computer program, as we have already seen, has such an analysis with respect to the computational engine (figure 2), so the 'alternative' interpretation of a program can only depend upon its accidental features (Wittgenstein, 1921; T3.34). This duality of interpretation develops a peculiar tension in program design that is hard to keep stable, particularly with respect to the informal, and often undefined, mechanism which links the program names and organisation with the user's domain. Further, the 'objects' that are usually chosen to be referenced in the informal analysis of the problem domain do not normally have all the features required of *Tractatus* type objects and thereby they keep on changing and this change can be seen as an underlying force to the phenomena of software evolution (Lehman et al., 1997).

5. The Paradigm Leap

The *Tractatus* is an established piece of work in the field of Philosophy and is an effective paradigm of how programming languages should be linked to a computer. There is no problem with the engineering necessity of this approach to sense and meaning. On a broader scale it sidesteps many of the paradoxes of the linguistic philosophy of the day, paradoxes which mostly stem from a simplistic application of referential semantics (Quine, 1964). However, it has a fatal flaw when applied to the human use of language and Wittgenstein eventually exposed this flaw. He noted that it is not possible to unambiguously describe everything within the Propositional paradigm (Wittgenstein, 1966). Wittgenstein found that the normal use of language is riddled with example concepts that cannot be bounded by logical statements that depend upon a pure notion of referential objects. This means that it is almost impossible to make any general statement about the world that is both true and has empirical content, e.g. "Birds can fly".

One of his illustrations is an attempt to define a game (Wittgenstein, 1966; P69 – P71). Such a definition cannot be achieved that will either exclude all examples that are not games or include all examples that are. This process has similarities to the field of Computational Linguistics where language paradigms are countered by exceptions. It is through such considerations that Wittgenstein proposed a new linguistic philosophy that was based upon what we are calling inferential semantics. The view epitomised by Wittgenstein's *Philosophical Investigations* is that meaning, grammar and even syntactic rules emerge from the collective practices (the situated, changing, meaningful use of language) of communities of users (Gooding, 2004b). Several decades later, critics of AI used this same difficulty to argue the impossibility of simulating human intelligence by the methods of traditional AI (Dreyfus1972, 1992, Collins 1990)

This observation by Wittgenstein has since been substantiated by ethnographic studies (Suchman 1987, Hutchins 1995) and by post Kuhnian studies of science (Kuhn 1962, 1985). To address it we make the distinction between rational and irrational sets.

*An **irrational set**[3] is where no finite set of rules can be constructed that can include unambiguously any member of that set and, at the same time, unambiguously exclude any non-member of that set.*

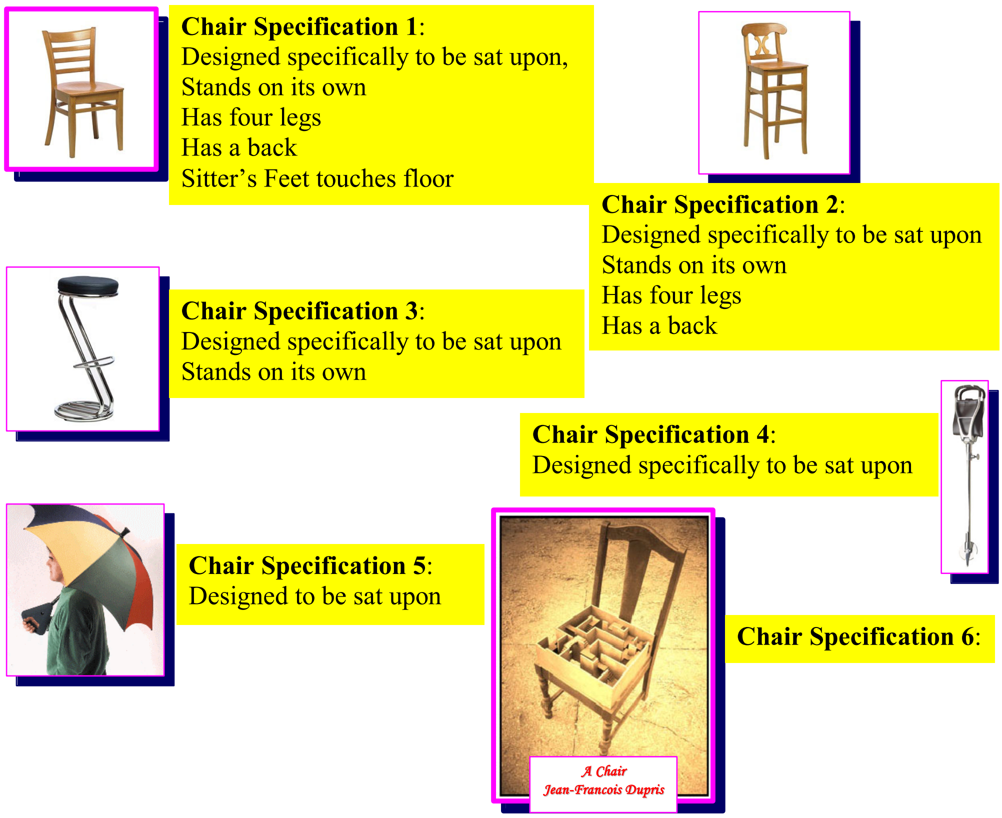

Figure 3 An attempt at identifying a chair using rules

By way of illustration consider the set of chairs and a possible specification (figure 3, Chair Specification 1). It is always possible to find some exception to a finite set of rules that attempts to identify a member of the set 'chair'. The counter to this is that instead of a rule one could just simple provide a list of exceptions. If every exception was then added to a membership list this technique would break down by simply discovering a context in which at least one of the set would cease to be identified as a member through the use of the rules or context. The more extreme cases added to the set, the more opportunities there will be for situations that exclude the accepted members of the set. We are thus in a position where some things are not potentially unambiguously describable. (See Appendix B for examples drawn from the press.)

The problems created by irrational sets are well known, particularly where human computer interaction occurs. The difficulties arise when computer scientists attempt to provide an inappropriate, 'rational' description of what can only be described by irrational sets. One attempted solution to these problems has stimulated extensions to the 'crisp' set by assigning a 'value' to a membership. Examples are fuzzy and probabilistic membership assignments. However, fuzzy sets are also rational in that members are assigned an explicit, ordinal membership number. Such assignments can be expressed by

[3] The idea of rational and irrational sets was originally proposed by Jan Townsend Addis (private communication February 2004).

a finite set of rules. Similarly, a probabilistic assignment of membership is also rational because it is specified by a rule in the form of a ratio of integers.

Even though some sets can be considered irrational sets we still have rational sets and so denotation remains one mechanism for relating meaning to a name. For irrational sets there is an additional mechanism for meaning assignment based upon human usage and context. Meaning derived through this latter route would imply a mechanism that adapts from experience. It is this mechanism that provides the link between the program and the world with which it is designed to work and it is the other half of the dual semantics (see later the 'abductive loop').

6. Some Predictions from this Thesis

We have computer programs with a semantics based upon computer bits but we have no formal way of assigning the meaning of the program to the problem domain: yet this is the very reason we wrote the program. Programs must remain in the domain of rational sets if they are to be implemented on a machine. We have the freedom to use the program's accidental properties without affecting the program's meaning with respect to the computer. We can choose the names we use (variables etc.) and dictate the computer organisation within the possibilities of the programming language.

Alternatively, consider the mapping of the program to the computer as a proposition, and since propositions can adopt many different but equivalent (tautological) forms so can a program. It is the job of a compiler to make a transformation of a program in order for it to run. For any particular computer there are an infinite but bounded number of possible structural forms for a given program that will behave similarly. The possibilities are bounded by the intended final form of the program (the essential program). Apart from these limitations the choice of form chosen is in the hands of the programmer. This produces a tension between the two semantics and based on this we can make the following statements:

- Reverse engineering is impossible unless domain information is used.
- Machine mismatches can be detected through tautology
- Formal 'objects' (e.g. Windows in OO or implementations of mathematical procedures) will be stable. However, informal 'objects' (e.g. persons, chairs or games) will never be fully captured or be stable because they are irrational sets open to the dynamics of human affairs. Thus programs that represent particular domains in the world can never be finished products; they will always require changing.
- Backward compatibility for a domain in the world can rarely be maintained for any length of time; it will always breakdown eventually (e.g. a database).
- It will not be possible to completely represent certain human functionality such as natural language understanding on any machine that is not adaptable.
- Increasing a training set for machine-learning algorithms will eventually cause degradation in its recognition performance if the set includes irrational distinctions. This is because the new rules (say) required to include the new extreme cases will tend to exclude some older examples.

7. Inferential Semantics

From an engineering point of view the only information that can be experienced by an individual is the result of the interaction of the individual's sense organs with the world. This is not a passive view since these organs are also controlled by an inference mechanism; namely the human mind. It is through inference derived from sensory data that we experience the world and relate to other people. So like the computer we might be able to trace the sense of our understanding of the world by the tracing of internal constructs derived from our senses. However, this would not be of any great help to other people since it is unlikely that we are identical in the same way that two computers, constructed according to a defined engineering design, are identical. We could speculate that if we were to be different by as little as one bit we could not ever be sure that a 'program' would mean the same if 'run' in different heads or that it would even 'run' at all. So tracing and knowing the 'program' (or our internal constructions) is not very useful.

What could work, from a purely pragmatic point of view, is if individuals could construct models of the world, and other people, that were sufficient to meet the needs of surviving in the world and with others. This model does not have to be exact, just sufficient. However, to do this we have to extend our semantic model to have another definition of meaning; a definition that does not depend upon the direct referencing of objects. As we noted earlier, for Wittgenstein, the *meaning* of a word was also defined as its *use in language* (Wittgenstein, 1966; P43).

We can interpret this extended definition of meaning to imply a *process* of inference. During conversation, both observed and participating actively, a process is going on where a model of the meaning of words is being constructed through inference from word usage. This is a *group* activity and one designed to construct something common in the way language and the world may be perceived; a way that allows communication to occur. However, these models are only understood by their effectiveness, their ability to make predictions and their coherence within a group-dynamic situation. They can never have been 'seen' directly since they only exist within an individual. It is the hidden dimensions of the model that express concepts. Since these dimensions are likely to be different for different people we have the effect of distinctions (like 'chair') having no proper boundaries that can be logically defined.

This lack of boundaries for concepts is the *family resemblance* effect detected by Wittgenstein and illustrated by his example (Wittgenstein, 1966; P67). It is an effect that fuzzy sets, in some cases probability, and belief networks were intended to overcome (Wittgenstein, 1966; P71) without losing the power of referential assignment. An alternative to Wittgenstein's family resemblance is Lakoff's (Lakoff 1986, Lakoff & Johnson 1980) use of prototypes (paradigms) and metaphor instead of reference. With either route we have a more acceptable approach to human relationships in that there will always be a need for human judgement because what is acceptable behaviour or performance is a time sensitive and socially dependent notion. A research team in Mexico, in conjunction with Salford University, have also explored the use of family resemblance with a learning system in order to approach human performance in categorization (Vadera et al., 2003).

The tension caused by the dual semantics that pivots on the essential and accidental meaning of the signs used in programs has been recognised, as can be seen by the continued search for new languages, program structuring and systems design methods (e.g. Java, conceptual modelling and object orientation). The central problem of the human context has also been addressed through the pursuit of natural language understanding, naïve physics, case-based reasoning and adaptive interfaces (Addis, 2003, Hayes, 1979, Thompson et al., 2004). There is a belief that given sufficient power or moving beyond the Turing machine would somehow solve the problem (Wegner and Eberbach, 2004). This has not been demonstrated with such efforts as many-fold increases in computer power or parallel mechanisms such as neural nets. None of the approaches tried so far have really succeeded. Many of the pursuits have been constrained by the formal bounds represented by the *Tractatus* and those approaches that have broken away have not bridged the gap identified here.

8. The Task

An alternative to Wittgenstein's family resemblance is Lakoff's use (Lakoff 1986, Lakoff and Johnson 1980) of prototypes (paradigms) and metaphor instead of reference. With either route we have a more acceptable approach to human relationships in that there will always be a need for human judgement because what is acceptable behaviour or performance is a time sensitive and socially dependent notion. The requirement to encapsulate the wide range, and ever changing perceptions, of a problem domain is a continuous link with human activity. Such perceptions cannot be predicted and hence planned for in advance. So many of the current principles of design that assume stability of concepts will have to be shelved and two distinct design paths will need to be forged that involve the two independent elements of a program; the formal rational and the informal irrational (figure 4).

The questions to be answered are: Can we construct a computing mechanism based upon family resemblance rather than sets, paradigms rather than concepts, and metaphor rather than deduction? Can we devise systems that have judgement rather than decisions? One possibility is that we might be able to write dynamic, socially sensitive interfacing-compilers that can match any program to any user (see figure 4).

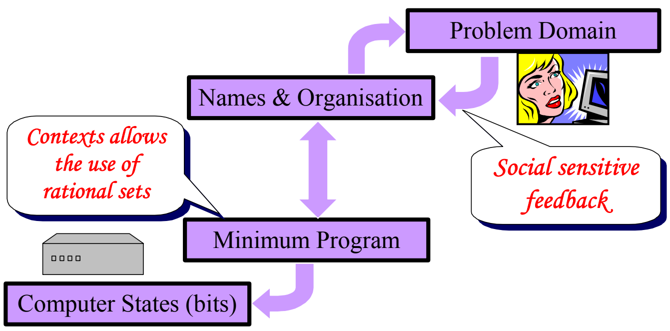

Figure 4 Showing where change can occur to solve the dual semantic problem

Such a compiler would be in 'conversation' with its user, other users and machines via (say) the Internet absorbing the human cultures and language so that its generated semantic and semiotic mappings make a program usable by a person. This might stop 'computer rage' through a more natural communication between people and machines; it may identify what is really meant by common sense.

9. A Mechanism to Handle Socially Sensitive Feedback

In the case of a formal system based upon rational sets as defined by the Tractatus deductive inference depends upon the preservation of truth. This mechanism is purely syntactic provided the coherence of the rational set is kept at all times. The introduction of irrational sets destroys the coherence of the sets and this means that the normal syntactic inference mechanisms can no longer be relied upon. So with what can we replace deduction?

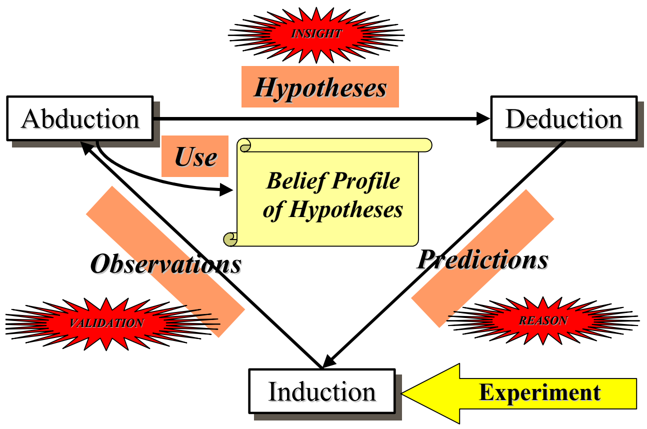

Figure 5 The abductive inference loop

A possible solution is to use the abductive inference loop (figure 5) originally proposed by C.S. Peirce (Weiner, 1966, Tursman, 1987) and modelled by Addis and Gooding (1999). In this model truth is confined to the deductive element and everywhere else truth is replaced by belief. Here belief is defined in terms of the probability of behaviour of an agent given a purpose. The need for purpose becomes central to any intelligent system as argued by Addis (2000) since without purpose there is no criteria for success and thus no mechanism for validation. Validation (or Peircean induction) serves the same role that syntax does in deduction; it provides a mechanism through which conclusions can be justified. This happens automatically with deduction since the syntax of an argument is sufficient criteria to justify the preservation of truth. This is further supported by the associated truth tables (Wittgenstein, 1921). In the case of validation the change in belief that governs action must relate to the likely success of that action and actions can only be judged to have success if there is a purpose.

For the process of validation to work and for the whole system to be effective the consequences of actions need to be known and assessed. In most cases this means that any system that is to keep track of the new set boundaries needs to be interactively engaged in its environment *by achieving goals*, therefore there must exist *a purpose* (Addis, 2000).

10. Discussion and Conclusion

In the late 60s Herbert Simon (1969) suggested a hypothesis drawn from the seminal work of Grey Walter (1961) that explains the complexity of human behaviour within a pragmatic framework. This hypothesis was that "*a man, viewed as a behaving system, is quite simple. The apparent complexity of his* ***behaviour is largely a reflection of the complexity of the environment in which he finds himself*.**" He finally notes "*the evidence is overwhelming that the system* (man's behaviour) *is basically serial in its operation: that it can process only a few symbols at a time and the symbols being processed must be held in special, limited memory structures whose content can change rapidly. The most striking limits on the subjects' capacity to employ efficient strategies arise from the very small capacity of the short-term memory structures (four chunks) and from the relatively long time (five seconds) required to transfer a chunk of information from short-term memory to long-term memory*," He concludes "*we should not expect it* (man's behaviour system) *to become essentially more complex*". Yet despite his initial strong focus on the environment as the source for complexity he led the way to trying to describe the complexity of human behaviour in terms of formal systems (Newall and Simon, 1958); systems that are isolated from the intricacy of the real world. We are suggesting that we need now to take more seriously these earlier perceptions and start describing system behaviour in terms of its response to not just a complex world but also, more importantly, an indefinite world.

So what we have identified here is in a real sense not new. There has been for some years now a very strong sense in the Computer Science and the AI communities that there is a problem. This is illustrated by the plethora of work harnessing ever more powerful computing to solving 'simple' problems such as natural language understanding, information retrieval, data mining and face recognition. The instinct in us by training, as it was for Simon, is to attempt to get to grips with the difficulties through formal analysis. However, it is this very formal analysis that is the problem. Any alternative approach seems to be counter to good science and therein lies madness. We are not proposing that we abandon the scientific method nor are we suggesting that formal analysis should be put to one side. What we are suggesting is an acknowledgement that there is a formal analysis that can take account of irrational sets. We are also suggesting that we acknowledge that computer programs can only deal with rational sets at any instant.

If such ideas are accepted then this should give good guidance to what can be achieved with computer systems and what designs are likely to be successful. In particular, rational sets can be used to represent a snapshot of an irrational set for the purposes of any deductive inference within a restricted and specific context. In this case the normal rules of truth preservation hold, the boundaries of sets remains fixed and hypotheses can be constructed within this framework. In the terms used by Lakotos to describe scientific evolution (Lakatos, 1970) we will have a non-progressive situation where the view of the world becomes more complicated as it is embellished with new constraints and distinctions to make it fit. However, to make such a view progressive and simpler, a new ontology has to be established where revised boundaries between sets and concepts are drawn. It is then that we see the analogue of Kuhn's paradigm shift (Kuhn, 1962) where new groups of people adopt a new perception of the world.

In normal human affairs such snapshots have to be continually updated to take into account shifts in the boundaries of the sets (and concepts) and this requires the use of belief instead of truth and abduction (with validation) instead of deduction. It is abduction that allows new boundaries between sets and concepts to be adopted on the bases of changes in the pragmatic criteria used by the cycle of validation. The criteria may shift and change the ontology of the system to some valuable effect. This shift should result in some positive benefit such as making a scientific hypothesis simpler or, in the case of some dispute, it may mean that a political agreement can be reached, or in business that a contract may be adjusted to gain some benefit for one or all of the participants. This shift is not changing the reality of the world but it is changing, fundamentally, the way the indefinite world is to be described.

There has not yet been a computer program that you can argue with, or persuade to adopt a new point of view or come to some agreement through conversation alone. The adoption of irrational sets requires a mechanism whose function depends upon the activity of searching for a shared ontology between people with different purposes or different points of view. The final assessment of an accepted shared ontology can only be judged on the fine balance between peoples' different requirements and their perceived payoff.

Considering the traditional approach towards the computer understanding of natural language within the framework of rational sets we can identify two distinct phases. There is the first phase where the flexibility of expression is explained by some syntactic theory of language (Chomsky, 1964) and the second phase where stability of meaning is provided by a semantic theory (Katz and Fodor, 1963). However, there has always been the suggestion that there is also a pragmatic phase where the act of assigning meaning may be found. This latter phase is currently little understood (Liu, 1997) and is not well defined.

We are suggesting that the recognition of irrational sets restructures this traditional approach and refocuses it onto a specific version of pragmatics because we can now explain the flexibility of language by changes in ontology, that is the particular choice of boundaries made by a person or group of people that define the concepts or sets that describe the world. The assignment of meaning here is the adoption of a temporary set of rules that define concepts or set membership in a practical sense; it is a pragmatic structure taking into account peoples' objectives.

We have proposed that the adoption of an ontology is dependent upon purpose. If that is the case then one of the major roles of language is to search for and agree a common ontology that suits either a common purpose or can contain more than one purpose. It is this activity that is central to negotiation and the evolution of ideas. It is a study of this activity that will lead us to the practical constructs required for a more comfortable human computer interaction.

11. References

ACZEL, A. D. (1996) *Fermat's Last Theorem: Unlocking the secret of an Ancient Mathematical Problem*, Penguin Books.

ADDIS, T. R. (1980) *A Theory of Knowledge for Man-Computer Communication*. PhD Thesis, Brunel University.
ADDIS, T. R. (2000) Stone Soup: Identifying Intelligence though Construction. *Kybernetics,* 29**,** 849-870.
ADDIS, T. R. (2003) *Designing Knowledge Based Systems*, 1st Edition 1985, Ab-Libris Ltd.
ADDIS, T. R., BILLINGE, D. & VISSCHER, B.-F. (2004) *Socially Sensitive Computing: A Necessary Paradigm Shift for Computer Science,* Newcastle, Grand Challenges in Computing, UK CPHC, UKCRC.
ADDIS, T. R. & GOODING, D. (1999) Learning as Collective Belief-Revision: Simulating reasoning about Disparate Phenomena. *AISB Symposium.* Edinburgh.
BILLINGE, D. & ADDIS, T. R. (2003) The Functioning of Tropic Communication: A Mechanism for Consistent Figurative Descriptions of Artistic Effect. *AISB Symposium on AI and Creativity in Arts and Science.* University of Wales, Aberystwyth, Wales, UK.
BILLINGE, D. & ADDIS, T. R. (2004) Modelling Inference. *Model-Based Reasoning in Science and Engineering, Abduction, Visualization and Simulation.* Pavia, Italy.
CHOMSKY, N. (1964) Current Issues in Linguistic Theory. IN FODOR, J. A. & KATZ, J. J. (Eds.) *The structure of language.* New Jersey, USA, Prentice-Hall, Inc.
BOWKER, G. C and STAR, S. L. , (1999), *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press
BOYER, C. B. (1949) *The concept of the calculus*, Dover pub, NY
COLLINS, H. (1990) *Artificial Experts: Social Knowledge and Intelligent Machines*, MIT Press
DREYFUS, H. L. (1972) *What Computers Can't Do: A Critique of Artificial Reason*, Harper & Row,
DREYFUS, H. L. (1992) *What Computers **Still** Can't Do: A Critique of Artificial Reason*, MIT Press,
GOODING, D. C., (1986), *How do scientists agree about novel observations?,* Stud. Hist. and Phil. Sci**.**, 17: 205-230.
GOODING, D. (2004a) Note on Wittgenstein's Tractatus. Bath.
GOODING, D. (2004b) Note on Wittgenstein's Philosophical Investigations. Bath.
HAYES, P. J. (1979) The Naive Physics Manifesto. IN MICHIE, D. (Ed.) *Expert Systems in the Micro Electronic Age.* Edinburgh, Edinburgh University Press.
HERTZ, H. (1956), *The Principles of Mechanics*, Dover Publications, New York (first published 1900).
HOARE, T. and MILNER R. (2004) *Grand Challenges in Computing Research*, The British Computing Society.
HUTCHINS, E. (1995) *Cognition in Wild*, MIT press
KATZ, J. J. & FODOR, J. A. (1963) The Structure of a Semantic Theory. *Language,* 39**,** 170-210.
KLEENE, S. C. (1967) *Mathematical Logic,* New York, John Wiley & Sons Ltd.
KUHN, T. S. (1962) *The Structure of Scientific Revolutions,* Chicago, USA, Chicago University Press.
KUHN, T. S. (1985) *The Essential Tension: Selected Studies in Scientific Tradition and Change*, University of Chicago Press
LAKATOS, I. (1970) Falsification and the methodology of scientific research. IN LAKATOS, I. & MUSGRAVE, A. (Eds.) *Criticism and the Growth of Knowledge.* Cambridge, UK, Cambridge University Press.

LAKOFF, G. (1986) *Women, Fire, and Dangerous Things*, University of Chicago Press.
LAKOFF, G. & JOHNSON, M. (1980) *Metaphors we live by*, University of Chicago Press.
LEHMAN, M. M., RAMIL, J. F., WERNICK, P. D. & PERRY, D. E. (1997) Metrics and Laws of Software Evolution - The Nineties View. *4th International Symposium on Software Metrics (Metrics '97).* Albuquerque, New Mexico, IEEE Press.
LIU, S. (1997) *Negative pragmatic transfer by adult Chinese learners of English as a foreign language: classification, distribution and Communicative effect*. Guangzhou, Guangdong University of Foreign Studies.
NEWALL, A. & SIMON, H. A. (1958) GPS, A Program that Simulated Human Thought. IN FEIGENBAUM, E. A. & FELDMAN, J. (Eds.) *Computers and Thought.* McGraw-Hill.
QUINE, W. V. (1964) The Problem of Meaning in Linguistics. *From a Logical Point of View.* 2nd ed. Cambridge, Harvard University Press.
SIMON, H. A. (1969) *The Sciences of the Artificial*, MIT Press.
SUCHMAN, L. A. (1987) *Plans and Situated Actions: the problem of human-machine communication*, Cambridge UP,
THOMPSON, C. A., GÖKER, M. H. & LANGLEY, P. (2004) A Personolized System for conversational Recommendations. *Journal of Artificial Intelligence Research*, 1-36.
TURSMAN, R. (1987) *Peirce's Theory of Scientific Discovery: A System of Logic Conceived as Semiotic*, Indiana University Press.
VADERA, S., RODRIQUEZ, A. & SUCCAR, E. (2003) Family resemblance, Bayesian networks and exemplars. *AISB Quarterly,* 114**,** 1-11.
WALTER, W. G. (1961) *the Living Brain,* Harmondsworth, Middlesex, Penguin Books.
WATT, D. A. (1991) *Programming Language Syntax and Semantics*, Prentice Hall.
WEGNER, P. & EBERBACH, E. (2004) New Models of Computation. *The Computer Journal,* 47**,** 4-9.
WEINER, P. P. (1966) *Charles S. Peirce: Selected Writings,* New York, Dover Press.
WITTGENSTEIN, L. (1921) *Tractatus logico-philosophicus,* London, Routledge & Kegan Paul.
WITTGENSTEIN, L. (1966) *Philosophical investigations,* London, Macmillan.

Appendix A

Criteria for SSC being a Grand Challenge

Foreword

The UKCRC Grand Challenge laid down some rules of acceptance of potential challenges. We consider below our proposal in the light of these demands.

Meeting the Criteria:

- *It arises from scientific curiosity about the foundation, the nature or the limits of a scientific discipline.*

The proposal arises from the puzzle as to why, after sixty years of effort, millions of man hours and technology that boasts of silicon machines that do 1000+ Gigaflops with a 1000+ Terabytes of storage, have still not even addressed many of the important functions of a human brain; a device that looks like a bowl of porridge and consists of only 15 Gigacells working at about 50 cycles per second.

- *It gives scope for engineering ambition to build something that has never been seen before.*

The proposal suggests new ways of looking at current problems. New types of computation would arise and new engines created along different principles; notions such as 'a structure malleable program' to any single solution.

- *It will be obvious how far and when the challenge has been met (or not).*

The challenge will have been addressed when it is no longer a problem that the world cannot be classified or partitioned.

- *It has enthusiastic support from (almost) the entire research community, even those who do not participate and do not benefit from it.*

This remains to be seen.

- *It has international scope: participation would increase the research profile of a nation.*

It clearly has worldwide implications.

- *It is generally comprehensible, and captures the imagination of the general public, as well as the esteem of scientists in other disciplines.*

What is meant by this criterion is excitement about the 'idea' proposed. The notion would strike at the very heart of how we organise ourselves and accept hypotheses. The accelerated growth of laws and regulations are derived from the misapprehensions that concepts can be captured exactly through definition. The rejection of this idea would releases us all from the inappropriate constraints imposed by those in authority; it would give us a rationale on which to reject nonsense.

- *It was formulated long ago, and still stands.*

If by 'long ago' it is meant in the early part of the 20th century then it does still stand.

- *It promises to go beyond what is initially possible, and requires development of understanding, techniques and tools unknown at the start of the project.*

This promise is certainly the case. A whole new technology and science would stem from this proposal.

- *It calls for planned co-operation among identified research teams and communities.*

It will require a wide range of specialisation ranging from psychologists, philosophers, linguists, sociologists and computer scientists of many sub-specialisations (e.g. networking, systems, architecture and interface design).

- *It encourages and benefits from competition among individuals and teams, with clear criteria on who is winning, or who has won.*

We would not like to see this happen. It could do, but we would discourage it.

- *It decomposes into identified intermediate research goals, whose achievement brings scientific or economic benefit, even if the project as a whole fails.*

This is most likely the case. A simple solution to such outcomes as information retrieval that works for people would be of considerable benefit to the community as a whole. Many of the objectives already pursued and abandoned from lack of practical success (such as natural language understanding and adaptive interfaces) could be re-examined from this new paradigm.

- *It will lead to radical paradigm shift, breaking free from the dead hand of legacy.*

It is a radical paradigm shift. The problem is, would we be able to go against our own training and start thinking within this new framework?

- *It is not likely to be met simply from commercially motivated evolutionary advance.*

SSC is not a simple evolution from where we stand and that is a problem (see last point above).

Appendix B

Examples from the Press of Irrational Sets at work

Introduction

The following examples need some explanation.

First the attempts at laying down a set of rules to cover censorship are shown to fail by David Hewson. The problem of dealing with Spam is similar since the spammer attempts to construct messages that will penetrate any filtering system. Sometimes the rules become such that legitimate messages get binned. Jeremy Clarkson and his team identify a similar problem in that the rules that were laid down by the government Health and Safety Inspectors prevent Jeremy and team from doing their job to a point where many useful activities are stopped.

Secondly, Cosmo Landsman notes the fluidity of irrational sets as a base for a theme in a story. It is here that the movement of ideas becomes the fascination and the whole point of the film. The execution of Paul Hill is a tragic example of the mismatch between the definitions of what is human life. The doctor who was murdered by Paul Hill perceived that he was doing a service to women by removing an unwanted growth like one would a cancer. In this way he was giving back a life to many women. For his work he was honoured as a valuable member of society. Paul Hill on the other hand saw that this growth was also a potential for life just as it is for any baby or child. If babies or children were killed, for whatever reason, then many people would support anyone who tried to stop this infanticide in whatever way they can. On the other hand, the law explicitly condemns the killing of people, where people includes infants, so much so that only the ultimate punishment is sufficient and the ultimate punishment is death. All parties had indisputable logic behind them but they each adopted, at different points in their argument, different criteria. Each group looked upon the other in horror and disgust for no better reason than the indeterminacy of the set called 'human life'.

The Problem of Rules:

From a report by David Hewson, Sunday Times, April 4th 2004.

[We did get permission to publish this quotation]

"Can't an intelligent filter analyse a page beforehand and make a machine judgement on its suitability? There are stacks of those around, too, and pretty worthless they are. Peacefire (www.peacefire.org), a web group opposed to online censorship, carried out an interesting experiment recently. It created dummy pages supposedly run by small sites; each carrying examples of anti-gay hate speech. Posing as individuals, the organisation complained about these pages to the big content-filtering programs, including SurfWatch, NetNanny and CyberPatrol.

Sure enough, the filtering companies responded by blocking the offenders. Then Peacefire revealed the true sources of the quotations – all were taken verbatim from the websites of conservative organisations, including the Family Research Council, Focus on the Family and Concerned Women for America. Would the content-filtering companies now block these big and influential lobby groups? Not yet, which means you can read the selfsame daft words on their sites, along with plenty of other material, but not on the bait pages that Peacefire erected to test the system."

From a report by Jeremy Clarkson, Sunday Times, April 11th 2004.

[We did not get permission to publish this quotation because of his rules of contract. However, we are allowed to report what was said]

The essence of the quotation was that the Health and Safety Executive (HSR) has attempted to provide legal control over all behaviour at work to such an extent that it is impossible to actually do his job. The nickname for the HSE is the Programme Prevention Department.

The Problem of Irrational Sets

Report on the film 'Capturing the Friedmans' by Cosmo Landsman, Sunday Times, April 11th 2004.

[We did not get permission to publish this quotation because of his rules of contract. However, we are allowed to report what was said]

He found this film riveting because every time new evidence was presented he kept on changing his mind about who was guilty until by the end of the film he was still unsure. The fascination about the film was concerned with the nature of memory and the way facts can be so fluid. It was a perplexing and poignant film.

Report on the execution of Paul Hill at Starke, Florida, CBSNEWS.com, September 4th 2003

"The execution of Paul Hill for the murder of a doctor who performed abortions and his bodyguard left U.S. abortion providers anxious – and wary that the former minister may become a martyr to the anti-abortion cause and spur others to act violently."

"Paul Hill's final statement ….. If you believe abortion is a lethal force, you should oppose the force and do what you can to stop it"

"Paul Hill should be honoured today, the abortionists should be executed. said Drew Holman"

"We think that unborn children should be protected and it should be law. Said Sheila Hopkins, a spokeswoman for the Florida Catholic Conference. We definitely reject his statement that it was justifiable homicide."