This is a slightly modified version of a review of: Rosalind Picard's *Affective Computing*, **MIT Press, 1997. written for:** *AI Magazine* **1999**

Review of: Affective Computing

Aaron Sloman
School of Computer Science
The University of Birmingham
Birmingham, B15 2TT, UK
http://www.cs.bham.ac.uk/~axs/

1 Introduction

Writing a multidisciplinary book is a risky business. Some experts are likely to be fiercely critical because of omissions or errors. Others with tunnel vision are likely to miss the point. Rosalind Picard, with considerable courage, addresses a broad collection of themes, including:

- The nature of motivation, emotions and feeling.
- The detection of emotional and other affective states and processes.
- The nature of intelligence and the relationships between intelligence and emotions.
- The physiology of the brain and other aspects of human physiology relevant to affective states.
- Requirements for effective human-computer interfaces in a wide range of situations.
- Wearable devices with a range of sensing and communicative functions.
- Philosophical and ethical issues relating to computers of the future.
- A brief encounter with theology.

This is a book with a bold vision. Some readers will find it inspiring and mind-stretching. Some will find it irritating. Some will have both reactions. It gives many pointers to the vast literature on emotions, including useful recent material, e.g. books by LeDoux, Goleman and Damasio.

The book ranges over themes of varying depth. The main deep theme concerns the nature of intelligence and the role of emotions in intelligence, which I discuss below in section 6, after commenting on some of the simpler themes.

2 Ubiquitous computing and sensing

It will increasingly be feasible to install sensors and computing devices in furniture, in walls, in car seats, in driving controls, in clothing, in jewelry and even in implants. So it will be possible to have a wide range of sensors, processors and transmitters, constantly monitoring, analysing, recording, and transmitting information about one's blood pressure, temperature, blood sugar level, muscular tension, and many other physiological states. Some of these devices, suitably hidden, could also monitor various aspects of the environment, including other people. Thus even your friends and colleagues will easily be able to record your conversation, your facial expressions, and perhaps with remote sensors your muscular tension, temperature, sweating, etc.

Picard believes that such machines can learn to predict our reactions, and use such predictions to inform us of risks and opportunities ("You'd really like that film ..."). She argues that such devices will increasingly be able to measure and categorise emotional and other affective states, so as to help us comprehend what is going on in ourselves and others. (*The New Scientist* No 2135 23rd May 1998 (p. 11) reports on a Japanese device which purports tell you what a pet or infant is trying to communicate!)

Some people may be alarmed by the prospect of being "spied on" by machines. Although Picard warns about ethical issues, she apparently welcomes the use of emotion detectors in a wide range of contexts and relationships (e.g. teacher and pupil). The final chapter, in particular, suggests that computing devices will help us choose our mood music, decide which scenes to record for our photo-albums, and find out about exhibitions, plays and other events likely to suit our tastes. Reactions to this prospect will differ widely. Many will dislike the idea of using remote devices to tell them which mood to expect in their spouse: it will strike them as an improper intrusion. On the other hand, there probably are some couples who, having found bedroom ceiling mirrors tame, will relish mutually informative intimate sensors.

If I were a pilot or a bus driver I might accept the right of my passengers to insist on my being wired up so as to minimise the risk of disasters due to my falling asleep at the controls or having a heart attack while in control. But I would not want a computer linked to such sensors to select music for me, and I have no wish to use such devices, especially hidden devices, to tell me what my wife, my colleagues, or my students are feeling, or to inform them of what I am feeling. I'd prefer us all to become more sensitive. Others may have a different view.

This discussion begs the deeper question: can computational devices really be used to detect emotions, as claimed? Yes: a subset, emotions which produce characteristic patterns in measurable physiological states. How such "sentic modulation" (defined on p.25) might be produced and detected is discussed at length in Chapters 5 and 6. I'll argue below that this will not work for most socially important emotions.

3 Communication and affect

The book makes an important claim which is relevant to AI research on communication and teaching, namely that the ability to detect and influence affective states in others is important in human communication and will be necessary for machines to interact effectively with humans. I agree with this, and made similar claims in (Sloman 1992). However, I am not sure the methods proposed by Picard will be very effective.

Can emotions be recognised on the basis of physiological changes? Chapter 6 describes affect recognition "as a pattern recognition problem and affect expression as pattern synthesis" (p.165). Such patterns can be measured by devices like physiological sensors and cameras recording facial expression or posture. "Despite its immense difficulty, emotion recognition is easier than thought recognition" It is easier "largely because there are not as many emotions as thoughts" (p.167).

Picard's discussion may give some readers the impression that all emotions essentially involve measurable physiological responses i.e. sentic modulation. However this is true of only a subset of emotions, as explained below. Moreover, many emotions involve thought contents and they are no less diverse than thoughts. You can, in principle, be afraid of, hope for, pleased about almost anything that you can believe or expect. Most of the emotions people are interested in are constituted largely by their *contents* and the contents of emotions are also the contents of thoughts. Even if there are a few physiologically detectable universal categories covering all emotions (happy, sad, angry, etc.), which I doubt, it won't be much use knowing that your spouse is happy without knowing whether it is happiness about your promotion or about a message from a secret lover. Measuring sentic modulation will not in general answer such crucial questions about human emotions.

4 Ambiguity in the concept of "emotion"

It has been clear for a long time (Arnold 1968) that the word "emotion" has no unique and clear meaning. A proliferation of definitions can be found in philosophical and psychological literature. It might therefore be wise to avoid discussing whether computers can have or detect emotions. Alternatively, we can introduce new technically defined terms and use those. This is best done using *architecture based* concepts: starting from an architecture we can derive the types of states and processes the architecture can support (Simon 1967; Sloman 1992; Wright *et al.* 1996; Sloman 1998). Animals or machines with different architectures will not be describable in those terms. On this basis we can distinguish primary and secondary emotions.

4.1 Primary emotions

Human brains have many components which are evolutionarily old. Some are responsible for "animal" emotions, e.g. being startled, frozen with terror, sexually aroused, or nauseated. Information from perceptual systems fed to a fast pattern recognition mechanism can rapidly trigger massive global changes. Such mechanisms apparently include the brain stem and the limbic system (Goleman 1996; LeDoux 1996). Engineers will appreciate the need for fast-acting pattern-based global "alarm" mechanisms to ensure that an agent reacts appropriately to important risks and opportunities (Sloman 1998).

Damasio (1994) calls these "primary emotions" (pp 131-4), as does Picard. These products of our evolutionary history are still often useful. Because they involve physiological reactions relevant to attacking, fleeing, freezing, and so on, sensors measuring physiological changes (including posture and facial expression) can detect such primary emotions.

4.2 Secondary emotions

Primary emotions may be less important for civilised social animals than certain semantically rich affective states generated by cognitive processes involving appraisal of perceived, or imagined situations. These are referred to by Damasio as "secondary emotions", and described by Picard on pp 35–6,63–4. They can arise only in an architecture with mechanisms for processes such as envisaging, recalling, planning and reasoning. Patterns in such processes can trigger learnt or innate associations in the "alarm" system which cause rapid automatic evaluations to be performed. Possible effects include:

(1) Reactions in the primary emotion system including physiological changes e.g. muscular tension, weeping, flushing, smiling. These can produce a characteristic "feel", e.g. "a flush of embarrassment", "growing tension", etc. (Try imagining a surgical operation on your eyeball.)

(2) Rapid involuntary redirection of thought processes (compare (Simon 1967; Sloman and Croucher 1981)).

It is not always appreciated that effects of type (2) can occur without effects of type (1).

4.3 Two types of secondary emotions

Damasio conjectures that triggering by thought contents depends on "somatic-markers" which link patterns of thought contents with previously experienced pleasures or pains or other strong feelings. Such triggering enables secondary emotions to play an important role by directing and redirecting attention in dealing with complex decisions (Damasio pp. 174–5). Picard also believes that secondary emotions always trigger primary mechanisms, producing sentic modulation. However, I think we need a distinction between two sub-classes: "central" and "peripheral" secondary emotions.

(a) *Central secondary emotions* involve involuntary redirection of ongoing cognitive processes like planning, reasoning, reminiscing, self-monitoring, etc. Such shifts of attention can occur entirely at the cognitive level without involving sentic modulation. An example might be guilt, which involves

negative assessment of one's own motives, decisions or thoughts, and can produce thoughts about whether detection will occur, whether to confess, likely punishment, how to atone, how to avoid detection, etc. Other emotions (infatuation, anxiety, etc.) will have different effects on attention.
(b) *Peripheral secondary emotions* occur when cognitive processes trigger states like primary emotions, without any disposition to redirect thought processes (e.g. the shudder produced by imagining scraping one's fingernails on a blackboard).

A hybrid secondary emotion could involve a mixture of both types, e.g. guilt or embarrassment accompanied by sensed bodily changes.

Emotions of type (a) are often important to novelists, playwrights, poets and garden fence gossips. There need not be any overt expression, but when there is it will typically be some sort of verbal utterance or intentional action. By that I don't mean that people *label* their emotions: like other animals and young children even human adults may lack the sophistication to recognise and classify their own mental states. Rather, a central secondary emotion can be expressed involuntarily in choice of words, or an extended thought or behaviour pattern such as frequently returning to a theme, or always expressing disapproval of a certain person.

Subtle patterns expressing anger, jealousy, pride, or infatuation may be clearly visible to others long before the subject notices the emotional state. In (Sloman 1992) I conjectured that some involuntary emotion expressions are necessary in a society of intelligent agents because a society where *all* expression was voluntary would be too unstable: nobody would know whom to trust.

In humans, emotional reactions become less physical, more central, and more controlled, as part of the process of growing up and becoming emotionally mature (Goleman 1996). Damasio's architecture could easily be modified to explain this. For instance:
(i) The global alarm system might be able to control cognitive mechanisms directly, without causing sentic modulation.
(ii) There might be a second global alarm system whose outputs are directed only at the deliberative (cognitive) mechanisms.

On either alternative, some secondary emotions could cause cognitive interrupts, redirection of attention and new evaluations, without necessarily triggering primary emotion symptoms. Of course, the older primary emotion mechanism might *sometimes* be triggered concurrently, producing hybrid emotions.

4.4 Can secondary emotions be detected by physical sensors?

If Damasio's and Picard's theory of secondary emotions were correct then secondary emotions would always generate sentic modulation, which could be measured externally and used to detect emotions using pattern recognition techniques. Two hypotheses need to be considered:

> H1. The *general* nature of a secondary emotion can be detected (e.g. anger, jealousy, joy, apprehension, consternation), but not the semantic content (e.g. who you are angry with and what you are angry about).
>
> H2. Both the general class of emotion (e.g. anger) and the semantic content (anger at Fred for giving away a secret) can be detected from such patterns.

H2 is no more plausible than the hypothesis that our thought contents can be inferred from externally observable physiological patterns. For instance, primary emotion mechanisms which evolved long before political systems cannot be expected to produce a distinctive physiological pattern for consternation over an election defeat or pride at being elected president.

Is H1, the weaker hypothesis, true? Picard acknowledges that there will be considerable variation in the mapping between emotions and sentic modulation, depending on the person and the type of emotion. The culture can also make a difference (e.g. the British "stiff upper lip"). Subject to the need to calibrate patterns for individuals, is H1 true?

I suspect the answer depends on the person. Even if secondary emotions in most people sometimes trigger the primary emotion mechanisms it does not follow that they *always* do, for all people, or that they need to in artificial emotional agents. So, as an unqualified generalisation, I suspect even H1 is false.

It is specially problematic for mixed emotions, like those reported by the captain of a woman's yachting team at the end of a round the world race: She was looking forward intensely to seeing friends and relatives whom she had not seen for many weeks, and to eating again (since food had run out two or three days earlier); she was delighted at completing the race successfully, but disappointed at not winning it, sad at the thought that the adventure was over and the team would have to separate, but looking forward to the next such event. Could any collection of grimaces, shudders, posture changes, tears, sweating, etc. have expressed that mixture? Words are more than apt to the task.

4.5 How are mixed emotions possible?

Picard offers blending and rapid alternation as possible models of such mixed emotions (p.171). Neither is plausible, since coexisting emotions (jealousy and guilt at feeling jealous) endure and preserve their identity. A more accurate model would be a collection of coexisting dispositions, possibly implemented as concurrent (mostly unconscious) cognitive processes striving for attention and control. Of course a combination of infatuation with X, and jealousy of X's lover could be regarded as a blend, but only insofar as one is the cause of the other. The infatuation could also coexist with unrelated jealousy over a colleague's promotion. There's no more need for these to blend or alternate than for different coexisting beliefs or expectations to blend or alternate, or all the many attitudes that we have to different things, people, policies, places, etc. A highly parallel information processing architecture makes blending and alternating unnecessary.

4.6 Linguistic expression

Picard mentions the power of written communication regarding emotions (pp 13 and 97), yet she underestimates its importance:

> "If computers are to utilise the natural channels of emotional communication used by people, then when computers learn to recognise human emotions, they will have to rely primarily on sentic modulation, as opposed to having people explicitly tell them the names of their emotional feelings." (p.26 my underlining).

Contrary to Picard I suspect that for the emotions that matter most to humans the *primary* and most natural mode of expression is linguistic. As indicated previously, I don't mean that people recognise and label their own emotions. Rather, as novelists and playwrights know well, our choice of language can convey rich information about mental states by making our thought processes "readable" externally. This allows richer and more diverse affective states to be expressed than either a set of physiological patterns or a set of verbal labels. Part of the evidence is the long history of intensely emotional letter writing and poetry. There are also the profound outpourings (e.g. of grief or sympathy) in internet news groups concerned largely with mutual consolation and support.

The writer of the following message clearly feels guilty, apologetic and a little apprehensive, without saying so:

"I know I promised to give you a chance to check our paper, but my portion was not finished until just before the deadline and I had to send it in. I hope it won't cause you any embarrassment."

Compare Marc Antony's soliloquy when first left alone with the corpse of Caesar:

O, pardon me, thou bleeding piece of earth,
That I am meek and gentle with these butchers!
Thou art the ruins of the noblest man
That ever lived in the tide of times.
Woe to the hand that shed this costly blood!
...
Cry 'Havoc!', and let slip the dogs of war,
That this foul deed shall smell above the earth,
With carrion men, groaning for burial.

This is far more powerful than explicit telling. When human emotions are expressed we don't usually "tell them the names". In fact, very often the emotions have no names, for they are complex and sometimes even unique combinations involving interactions between semantically rich coexisting motivational states, attitudes, evaluations, expectations, beliefs, etc. And even when they have names, having the emotion does not require knowing the name or being able to recognize the occurrence in oneself.[1]

5 Towards a modified version of Damasio's theory

Damasio wrote:

> "I see the *essence* of emotion as the collection of changes in body state that are induced in myriad organs by nerve cell terminals, under the control of a dedicated brain system, which is responding to <u>the content of thoughts relative to a particular entity or event</u>. (Damasio 1994, p.139, my underlining)

Emphasising the contents of thoughts is correct. Emphasising body state ignores central secondary emotions of type (a) involving only involuntary attention control mechanisms. Damasio shows later that he is fully aware that there can be such chains of causation at the cognitive level.

> "...consumed as we are by using the past to plan what-comes-next, a moment away or in the distant future. That all-consuming, ceaseless process of creation is what reasoning and deciding are about..." (p.165).

These are processes in a mental "virtual machine". We can modify his theory by allowing that some secondary emotions involve only such virtual machine processes, without any sentic modulation (though obviously there are physiological processes in the brain). He discusses a general requirement for such mental processes later:

> The personal and immediate social domain is the one closest to our destiny and the one which involves the greatest uncertainty and complexity. Broadly speaking, within that domain, deciding well is selecting a response that will be ultimately advantageous to the organism.... Deciding well also means deciding expeditiously, especially when time is of the essence, and, in the very least, deciding in a time frame deemed appropriate for the problem at hand. (Damasio p.169)

[1]It may be objected that it is often possible to read verbal utterances in more than one way: they do not uniquely express particular emotions. That is true, but equally facial expressions, posture, and other external signs can often be interpreted in more than one way. Moreover, such physical clues are often very deceptive, which is what makes possible being an actor, a spy, an undercover agent, or a con-man.

This emphasises what might be called "meta-level" control, i.e. combining the ability to solve problems and make plans with the ability to notice that such a process needs to be interrupted or redirected, for instance because time is too short to continue planning: action must start. Damasio assumed that only emotions can perform this kind of high level control. However, recent work in AI has produced alternative meta-level control mechanisms, including planners which use "anytime algorithms" (Boddy and Dean 1989) which allows an interrupted planning processes to produce useful initial partial plans in intelligent machines.

Some plan-based theorem provers use high level strategic knowledge to guide the search for a proof. Similarly, Luc Beaudoin's PhD thesis (Beaudoin 1994) discussed the notion of a meta-management layer in the architecture of an agent with reactive and deliberative mechanisms. This layer can monitor, evaluate, and to some extent control and redirect attention in deliberative mechanisms. There is much work still to be done exploring such mechanisms.

We therefore do not need to follow Damasio and Picard in assuming that rapid redirection of attention by a global alarm system *necessarily* works by invoking the primary emotion mechanism. Cognitive control mechanisms can directly redirect attention to new goals, new items of information, etc. (Simon 1967; Sloman 1997). In some cases this goes on even though such redirection is of little or no value, e.g. in obsessive jealousy or anger. Damasio comes close to accepting this (pp 197-8).

Building on Beaudoin's work, my colleagues and I have argued that humans have at least three important coexisting interacting control systems, an old reactive layer, a newer deliberative layer, and a still more recent meta-management layer concerned with monitoring, evaluating and redirecting internal processes. Control by the third layer is generally only partial since there are automatic mechanisms which can redirect attention, including one or more reactive "global alarm" mechanisms as discussed above. We have suggested that different sorts of emotions correspond to processes in these three layers. Picard summarises these ideas about meta-management and perturbance on pages 211–213. However this theory is in its early stages of development. There is still no adequate account of pleasure and pain.

If computers are to be able to detect, respond to, or model these types of affective states they will need rich and subtle linguistic abilities, and a deep understanding of the structure of human minds. Chapter 7 of the book discusses various recent attempts to model cognitive affective processes, which are totally different from the sentic modulation capabilities that are the main focus of chapters 5 and 6. This is an important topic for future AI research on self-aware and socially-aware agents.

My feeling is that looking back in years to come we'll find that all this work is very shallow and inadequate, especially models involving explicitly labelled emotional states and special emotion generating rules.

6 Emotions and intelligence

As Picard points out, most AI researchers ignore motivational and emotional mechanisms. There are a few exceptions, notably Simon's pioneering 1967 paper, responding to criticisms of AI by the psychologist Neisser. Randall Davis, in his AAAI96 presidential address, views AI as exploration of the space of designs for intelligent systems, including both natural and artificial systems. He discusses at length the diversity of capabilities and mechanisms which have evolved naturally, yet there is no mention of motivation and emotion.

If emotions are not *required* for meta-level control, is affect necessarily relevant to intelligence? A partial answer is one of the Picard's themes mentioned above, with which I agreed: if intelligent machines are to communicate effectively with humans in a wide range of situations for a wide range of purposes (including teaching, advising, etc.), then they will often need to take account of actual and likely motivational and emotional states.

However, must intelligent systems also *have* emotions? I think they almost certainly will, but not for the reasons given by Picard and Damasio. I have argued (Sloman and Croucher 1981) that certain kinds of emotions will be side effects of mechanisms designed to overcome resource limits in intelligent systems. They argue that emotional mechanisms are required for intelligence.

6.1 Must a skilled communicator have emotions?

It may be possible for a robot without any emotions of its own to learn a great deal about human emotions, including learning how to recognise them on the basis of both physical states and, more generally, on the basis of understanding what the pupil or client is saying.

It is often suggested that we reason about mental processes of others by simulating them in our own brain. If this were the *only* way to predict emotional reactions, it would be impossible for a completely unemotional agent. However, we can also use general knowledge to reason about reactions of other people. Nevertheless, it is likely, that one source of information about emotions, is one's own emotional experiences. So perhaps a Spock without emotions might find it difficult to deal adequately with humans, though not impossible.

6.2 Are emotions required for general intelligence?

Picard, like Damasio, makes a much stronger claim, namely that the ability to have emotions is *required* not only for communicating about emotions, but more generally.

> "We all know that too much emotion can wreak havoc on reasoning, but now there is evidence that *too little* emotion also can wreak havoc." (p10)
> "Apparently, a balance is needed – not too much emotion, and not too little emotion." (p11)
> "Damasio's findings point to an essential role of emotion in rational thinking." (p12)
> "Nevertheless when a system faces problems where the possibilities cannot be enumerated and evaluated in the available time, I suggest that affective decision making provides a good solution. Humans use feelings to help them navigate the oceans of inquiry, to make decisions in the face of combinatorial complexity." (p222)

I fear this is an over-generalisation from how a subset of humans work. There are two problems with these claims: mis-construal of the expertise involved in avoiding search, and misinterpretation of the evidence from brain damage.

There are many areas of expertise which potentially involve massive searches, but where humans somehow manage to avoid the searches. Anyone who has worked on natural language processing, or image understanding will know that both utterances and retinal images have huge amounts of local ambiguity regarding segmentation, grouping, selection among possible meanings, and, in the case of images, possible occlusions. Typically the problem of resolving the local ambiguity subject to global constraints can involve enormous search spaces. Yet humans seem to home in rapidly on a unique interpretation without searching, except in the case of garden path sentences (e.g. "the horse raced past the barn fell") and puzzle pictures.

The obvious explanation is that expertise is based on a very large collection of slightly generalised special cases stored in some kind of content-addressable memory. The same can be said about expertise in more abstract domains, such as logic, algebra, programming, and games such as chess. In a culture the experience of many generations can be transmitted in a compressed time-scale to new members. Likewise it is now commonplace to allow chess programs to make use of a great deal of "book learning" to avoid massive combinatorial searches.

Although the full workings of human associative learning are neither well understood nor replicated in current AI systems, it is clear that they need not involve emotions (though in some cases they do, sometimes with bad results). For example, a young child picks up a huge vocabulary and many subtle grammatical rules, simply by being exposed to speakers of the language.

Another thing we learn is control information. Experts learn to detect patterns in a situation which suggest that a different approach is needed, or that the problem is insoluble, or that some more important and urgent problem has arisen. So, besides hierarchically structured goal-directed processes, an animal or robot with multiple independent sources of motivation inhabiting a dynamically changing and partly unpredictable environment needs mechanisms which can redirect attention away from the *current* goal and its subgoals. (Compare chapter 6 of (Sloman 1978).)

These mechanisms need not be emotional, though sometimes they will be, namely when the intrusions involve highly positively or negatively charged evaluations and cannot be prevented by meta-management decisions of the agent. I have called these "perturbant" states.[2] (A perturbant state can become dormant when attention is grabbed by something more powerful. Dormant emotions don't go away: they simply await their chance to regain control, for instance when grief is temporarily forgotten because an urgent and important task grabs one's attention.)

In summary: the heuristic control powers Picard and Damasio attribute to emotions can occur without emotional mechanisms, though, as argued in (Simon 1967; Sloman and Croucher 1981) such control mechanisms may be capable of generating central secondary emotions.

6.3 How to interpret Damasio's evidence

Damasio, Picard and others have misinterpreted the evidence about brain damage in Damasio's book as implying that emotions are essential to intelligence. There's a simple non-sequitur.

Certain sorts of frontal lobe damage produce two effects: (1) the patients lose the ability to have certain kinds of (secondary) emotional reactions, or to *care* about things that previously mattered to them (including physical pains in some cases), (2) the patients become less creative and decisive, and less able to take strategic decisions. This can render them totally unable to manage their own lives, even though they retain normal functions of perception, memory, language, motor control, perform well on all standard intelligence tests, and even have explicit knowledge about how they *ought* to behave in various circumstances.

It is fallacious to infer from this that (1) is the cause of (2), For there may be mechanisms for controlling and redirecting attention in the cognitive virtual machine, which are essential for intelligence, and which *also* produce secondary emotions. When they are damaged there is a loss of secondary emotions and *also* a loss of balanced judgement and control of thought processes. It does not follow that emotions are necessary for intelligence. Rather, mechanisms required for intelligence sometimes produce emotions. Such emotions are "emergent". Compare: disconnecting the car battery will prevent the radio working and the car starting. It doesn't follow that the radio is required for the car to start.

Of course, Damasio's evidence does support the hypothesis that in addition to the mechanisms studied in particular sub-fields of AI and cognitive psychology, an intelligent agent requires more global control mechanisms which attempt to ensure that those mechanisms are deployed appropriately. This is what meta-management is about. This need is obvious to any software

[2] In (Sloman 1998) and more recent papers they are referred to as "tertiary emotions".

engineer accustomed to designing systems with multiple capabilities, e.g. operating systems, or plant control systems.

Damasio's patients reveal that the high level control mechanism may be incapacitated while many more specific aspects of intelligence remain intact, which should not be very surprising to software engineers.

7 Conclusion

This wide ranging, ambitious, book presents work from several disciplines, including empirical results, theoretical analysis and practical applications along with some of the ethical issues they raise.

The majority of the discussion of emotions is based on the widely held assumption that they *always* involve externally detectable sentic modulation, as primary emotions do. This assumption, however, ignores the possibility of central secondary emotions, which I have claimed are the most important emotions in our (adult) lives and certainly of most interest in much of our thinking about one another. From this viewpoint the emphasis on externally detectable patterns of physiological processes is unfortunate.

However as an account of how primary emotions and some peripheral secondary emotions are expressed and how they might be detected it may be a good beginning.

The book includes many topics I have not had space to discuss, including several challenging and potentially extremely interesting and probably very difficult applications of affective computing, e.g. automating the process of searching a library for a picture or a piece of music with a specific type of mood, for use in an advertisement or as background to a film.

Although I have been critical of some major themes, there is much of interest and value and I believe the book has no competitors. As a wide-ranging and provocative ground-breaker it can be recommended to students of AI who need to have their minds stretched. But they should be warned not to believe everything they read!

Readers will find in the text and bibliography pointers to much relevant literature, though I suspect the best literature on this topic has yet to be written, perhaps by readers stimulated and challenged by this book. Such work requires a broad multidisciplinary background. Unfortunately there are still too few researchers like Picard willing to combine psychology, ethology, neuroscience, evolution, computer science, software engineering, AI and philosophical insight, in the context of creative engineering design.

Maybe one day their numbers will reach a critical mass, they will discover a common conceptual framework within which to communicate, and the subject will really take off. A regular section in the *AI Magazine* on "Affective Computing" might help to accelerate this process.

Acknowledgements

I gratefully acknowledge ideas gained from or sharpened by colleagues in the Cognition and Affect group in Birmingham, and also useful comments and suggestions on an earlier draft, by Adele Howe, Fanya Montalvo and Rosalind Picard.

References

Magda B Arnold, editor. *The Nature of Emotion*. Penguin Books, Harmondsworth, England, 1968.

L.P. Beaudoin. *Goal processing in autonomous agents*. PhD thesis, School of Computer Science, The University of Birmingham, 1994.

M. Boddy and T. Dean. Solving time-dependent planning problems. In *Proc 11th International Joint Conference on AI*, Detroit, MI, August 1989.

Antonio R Damasio. *Descartes' Error, Emotion Reason and the Human Brain*. Grosset/Putnam Books, 1994.

Randall Davis. What are intelligence? and why? *AI Magazine*, 19(1):91–110, 1998. (Presidential Address to AAAI96).

Daniel Goleman. *Emotional Intelligence: Why It Can Matter More than IQ*. Bloomsbury Publishing, London, 1996.

Joseph E LeDoux. *The Emotional Brain*. Simon & Schuster, New York, 1996.

H. A. Simon. Motivational and emotional controls of cognition, 1967. Reprinted in *Models of Thought,* Yale University Press, 29–38, 1979.

A. Sloman and M. Croucher. Why robots will have emotions. In *Proc 7th Int. Joint Conference on AI*, pages 197–202, Vancouver, 1981.

A. Sloman. *The Computer Revolution in Philosophy*. Harvester Press (and Humanities Press), Hassocks, Sussex, 1978.

A. Sloman. Motives mechanisms and emotions. *Cognition and Emotion*, 1(3):217–234, 1987. Reprinted in M.A. Boden (ed), *The Philosophy of Artificial Intelligence*, 'Oxford Readings in Philosophy' Series, Oxford University Press, 231–247, 1990.

A. Sloman. Prolegomena to a theory of communication and affect. In A. Ortony, J. Slack, and O. Stock, editors, *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues*, pages 229–260. Springer, Heidelberg, Germany, 1992.

A. Sloman. Explorations in design space. In A.G. Cohn, editor, *Proceedings 11th European Conference on AI*, pages 578–582, Amsterdam, 1994. John Wiley.

A. Sloman. What sort of control system is able to have a personality. In Robert Trappl and Paolo Petta, editors, *Creating Personalities for Synthetic Actors: Towards Autonomous Personality Agents*, pages 166–208. Springer (Lecture Notes in AI), Berlin, 1997.

A. Sloman. Damasio, descartes, alarms and meta-management. In *Proceedings International Conference on Systems, Man, and Cybernetics (SMC98)*, pages 2652–7. IEEE, 1998.

I.P. Wright, A. Sloman, and L.P. Beaudoin. Towards a design-based analysis of emotional episodes. *Philosophy Psychiatry and Psychology*, 3(2):101–126, 1996.