complexity of algorithms, complexity inherent to the task of specifying the deterministic or nondeterministic machine, and complexity of electric or logical circuits), physical limits of computing (that is, computation in the physical world requires the expenditure of energy, communication in space, and the passage of time), and limits of conceptualization (that is, finite, discrete concepts can never form a perfect model of a continuous world; the only things that can accurately be represented in concepts are manmade structures that once originated as concepts in some person's mind or systematic domains of distinctions created through the use of language) (Sowa 1984; White 1988). Furthermore, the main conceptual foundations of AI—namely, the knowledge representation hypothesis of Brian Smith (1982) and the physical symbol system hypothesis of Allen Newell (1980)—are not discussed at all. These hypotheses have been considered fundamental cornerstones of AI research, but they are now being questioned as posing strong limitations on AI (Dahlbäck 1989; Dreyfus 1972; Winograd and Flores 1986).

Given this perspective, the author concludes that AI's essential methodology is a continuous attempt to overcome the formal constraints of computer science and philosophy without sacrificing rigor. Although I liked the author's perspective, and I wholly agree with his main conclusion, both are just stated in the preface, and no further reference to them is given.

Let's get a feeling of what this first volume is really about. The organization of the text is clear and straightforward. This first volume is mainly expository in nature and comprises a detailed discussion of AI's formal constraints as a starting point to the discussion of AI's essential methodology that is promised for the second volume. The question of the possibility of machine intelligence is introduced in the first chapter through a detailed exposition on Turing's imitation game and an informal statement of its original objections. Turing's reformulation of the question of the possibility of machine intelligence into the imitation game is discussed after a brief exposition of two doctrines in the philosophy of mind—naive dualism and naive logical behaviorism—that provide the necessary background for an interpretation

of the reformulation. Each of the objections is then analyzed from a formal standpoint because the relevant elements of formal theory are introduced in subsequent chapters.

Theoretical computer science is dealt with in chapter 2. The chapter contains a brief description of function and automata theory as a basis for the detailed analysis of Lady Lovelace's objection. Despite the introductory character of the chapter, the omission of the theory of (primitive) recursive functions is surprising because Turing's and Church's theses on the limitations of computability, as well as Gödel's theorems, are strongly based on this theory (see, for example, Delong [1970]).

Formal philosophy is addressed in two chapters, one devoted to logic and another to semantics. Both propositional and predicate logic are covered in chapter 3 in addition to formal theories and systems, theorem proving, and logic-based knowledge representation. Gödel's theorems are exposed with clarity, although oversimplified, and the Mathematical objection in both its original and Lucas's version is thoroughly discussed. The presentation is far from complete, particularly with respect to theorem proving and logic-based knowledge representation (a more thorough treatment can be found, among others, in Genesereth and Nilsson [1987]), but is detailed enough to understand the metatheoretic view of logic systems.

Chapter 4 is much more comprehensive. It provides a detailed treatment of truth-conditional, model, and possible-world semantics. With the background in tensed modal logic given by the exposition of possible-world semantics, the distinction between the possibility and the necessity for machine intelligence is depicted with clarity. The argument from informality of behavior is also discussed.

In summary, this first volume provides a readable introduction to formal foundations of AI and gives a comprehensive analysis of the possibility for machine intelligence from this formal standpoint. Its significance is that it introduces a wide audience to the main issues surrounding thinking machines as well as provides a formal analysis of informal arguments for and against the possibility of machine intelligence. I highly recommend the book to anyone interested in the AI debate.

The second volume promises to draw on a characterization of AI's essential methodology as continuous attempts to overcome the formal constraints of computer science and philosophy by augmenting appropriate formal theories with nonformal yet rigorous models and approaches. It will also cover recent developments in neurocomputing. I hope to see my criticisms dissipate after reading the second volume.

References

Dahlbäck, N. 1989. A Symbol Is Not a Symbol. In Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, 8–14. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Delong, H. 1970. *A Profile of Mathematical Logic.* Reading, Mass.: Addison-Wesley.

Dreyfus, H. L. 1972. *What Computers Can't Do.* New York: Harper and Row.

Genesereth, M. R., and Nilsson, N. 1987. *Logical Foundations of Artificial Intelligence.* San Mateo, Calif.: Morgan Kaufmann.

Newell, A. 1980. Physical Symbol Systems. *Cognitive Science* 4:135–183.

Smith, B. 1982. Reflection and Semantics in a Procedural Language, Technical Report, MIT/LCS/TR-272, Massachusetts Institute of Technology.

Sowa, J. F. 1984. *Conceptual Structures: Information Processing in Man and Machine.* Reading, Mass.: Addison-Wesley.

White, I. 1988. The Limits and Capabilities of Machines: A Review. *IEEE Transactions on Systems, Man, and Cybernetics* 18(6): 917–938.

Winograd, T., and Flores, F. 1986. *Understanding Computers and Cognition: A New Foundation for Design.* Norwood, N.J.: Ablex.

Gabriel Valiente is a research associate in the Center of Advanced Studies of Blanes, Spain, and a graduate student at the Autonomous University of Barcelona. His research interests include the formal development and validation of expert systems.

The Cognitive Structure of Emotions

Roland J. Zito-Wolf

The Cognitive Structure of Emotions, Andrew Ortony, Gerald L. Clore, and Allan Collins, Cambridge University Press, Cambridge, United Kingdom, 207 pages, $14.95, ISBN 0-521-38664-0.

The Cognitive Structure of Emotions is an attempt to explain "how people's perceptions of the world—their construals—cause them to experience emotions" (p. 12). This work is one of theoretical psychology, striving to develop a cognitive framework for our emotional repertoire that will structure the existing empirical data and guide future research. The authors propose a taxonomy of emotion types differentiated by cognitive criteria and discuss the major cognitive factors that determine the type and intensity of emotional experience. They provide detailed descriptions of each emotion class and the specific emotions within it.

This book was published in hardcover in 1988; the occasion for this review is its publication in paperback form, an event that should greatly expand its audience. The ideas in this book deserve a wide reading. My discussion is from the point of view of an AI researcher rather than an emotion theorist. One particularly pleasant aspect of the book is that it is written to be generally accessible to the cognitive science and AI communities. The issues addressed are relevant to a wide variety of endeavors in these fields.

The Cognitive Structure of Emotions views emotions as elicited by specific types of *appraisals*—positive or negative judgments—made by an individual. Emotions can be differentiated in strictly cognitive terms according to the aspect on which the eliciting judgment focuses: *Events* are judged by their consequences, *agents* by their actions, and *objects* by their (presumably) intrinsic properties. These aspects are not exclusive; for example, people can be judged as agents or as objects, depending on whether one focuses on their behavior or their characteristics as individuals. The three aspects lead to three major classes of emotions: Appraisals of objects correspond to likes and dislikes of various kinds, appraisals of agents to pleasure and displeasure, and appraisals of events to approval and disapproval. The latter two classes have a rich internal structure. For example, the class of event appraisals bifurcates according to whether the event in question affects oneself (fortunes-of-self) or someone else (fortunes-of-others). Fortunes-of-self emotions are further differentiated by whether the appraised event has already occurred (joy, distress) or is prospective (hope, fear). Fortunes-of-others subdivides

according to the desirability of the event from the appraiser's point of view, yielding the emotion types happy-for, resentment, gloating, and pity. Altogether, 28 emotion types are distinguished. (These emotion terms are intended to define a technical vocabulary for emotion types and, thus, correspond only roughly to their vernacular senses.)

The research goals and emotion taxonomy are laid out in the initial two chapters. The reader is referred elsewhere for issues outside the major focus, such as how emotions are expressed, how they affect behavior, and how specific emotion words are used. Chapters 3 and 4 discuss the variables that distinguish subclasses of emotions and contribute to particular appraisals. Chapters 5 through 8 discuss the emotion taxonomy in detail, showing how and where some 130 (English) emotion terms fit into it. Chapter 9 summarizes the research and discusses limitations and possible applications.

This work is directly relevant to AI systems because it challenges the widely held belief that emotions resist computational modeling. This work suggests that although the experience (qualia) of emotion might be species (or entity) specific, the cognitions underlying these states can be broken into component cognitions more amenable to computational modeling, namely, judgments and inferences based on goals, plans, and other knowledge. Because appraisals are a fundamental part of making sense of the interactions between intelligent agents, the ability to reason about appraisals is important for systems aspiring to understand their users' responses in a deep sense or to model human behavior. The emotion taxonomy provides a basis for reasoning about emotions in general; for example, the taxonomy can be used as an organizing (indexing) structure for case-based reasoning (Wentworth 1991).

This book is interesting for several reasons. First, it makes explicit the large amount of cognitive processing underlying an occurrence of an emotional state. Appraisal is clearly a cognitive (although not necessarily conscious) process. Walking through a dimly lit alley, I experience fear because I have assessed the situation,

projected a potential consequence, and evaluated it as undesirable. In turn, appraisal depends on construal: one's choice of what to focus on in a given situation, how to view what is perceived, and what experience to bring to bear. My emotional response to the previous circumstances would be different if I judged the lighting adequate to deter attack, I had previously negotiated the same alley without incident, or I believed I was walking through a movie set. All these processes involve cognition. (One wonders what role the emotional experience itself plays in this account.)

The emotion taxonomy, besides being novel, has several benefits. It provides a coherent account of the relations among a wide variety of emotions. The theory postulates some emotion categories that do not correspond to any English emotion word, such as fears-confirmed (confirmation of a negative prospect) and happy-for (pleasure over a positive event occurring to another). Nearness in the taxonomy can be used as a measure of the relatedness of emotions. One consequence is that anger turns out to be the complement of gratitude. Similarly, depth in the taxonomy suggests a measure of the cognitive complexity of individual emotion types. Although the authors do not wish to be drawn into the extensive debate over how basic particular emotions are, their theory provides an additional perspective on this issue. For example, although most emotion theories take anger to be a (relatively) basic emotion, the framework presented here suggests that it is one of the most complex. A detailed discussion of this issue can be found in Ortony and Turner (1990).

Having shown that many emotions can be classified using their taxonomy of appraisal types, the authors further suggest that emotions be defined as precisely those items that fit into taxonomies so constructed: necessarily *valence* reactions (that is, positive or negative judgments) to specific events, agents, or objects. This proposal is suggestive but introduces some complications.

As a first pass at a general cognitive theory of emotion, given its ambitious scope, *The Cognitive Structure of Emotions* necessarily makes certain assumptions and simplifications. The classification of emotions using a taxonomy of appraisals works best where there is an identifiable target

for appraisal and where the resulting appraisal has a characteristic valence. Problems arise with states that one would intuitively call emotional but do not satisfy both requirements. Two examples are discussed here.

First, certain cognitive states, such as surprise and shock, can occur with either positive or negative valence and, possibly, with no valence at all. Confronted by an unexpected event that evokes a visceral reaction, must this experience necessarily involve a (significant) positive or negative feeling? The authors suggest that a cognitive state does not qualify as an emotional one if it lacks a characteristic valence (see also Ortony [1987]). As they note, this is in contrast to many other studies of emotion, not to mention my intuitive judgment. What is one to make of cognitive states combining emotions with opposed valences: If in some particular situation, the resulting state has no significant net valence, is it not, therefore, an emotional state? For example, lovesickness might mix the positive emotion of liking and the negative emotion of frustration, yielding a neutral or even oscillating net valence. My reading of the authors' position is that such mixtures will always have a dominating component, so the issue will not arise. Although the use of valence as a defining attribute of emotion is simple and intuitively plausible, it leaves a large gray area of states having visceral content but lacking characteristic valence: shock, confusion, ambivalence, desire, interest, anticipation.

A second difficulty is raised by states that do not have a specific evaluative focus: euphoria and anxiety as opposed to pleasure and fear. The authors treat such emotions as representing extremes of valence (that is, euphoria = extreme joy) or as being cumulative or diffuse reactions. However, such states can only in a weak sense be said to result from evaluations of specific events or actions because it is typically difficult to identify a specific appraisal that caused the state. In addition, such states have content beyond the merely evaluative, an experiential character distinct from their cause— depression and euphoria immediately come to mind. The authors define *moods* as cognitive states that predispose one to certain kinds of emotions, which suggests that items such as anxiety should be so treated. How-ever, this definition puts one in the position of asserting that anxiety, euphoria, and moods of all kinds are not emotional states at all.

The text does not address the noncognitive aspects of emotions. The cognitive analysis does not account, for example, for the similarity of physiological responses to particular emotions across individuals or the possibility of chemical influences on emotional states. If emotional states can be induced and modified by noncognitive factors, they cannot be defined purely in cognitive terms. These questions do not by any means negate the contribution of their analysis. However, they do suggest that there is perhaps a larger landscape remaining to be explored and that a comprehensive definition of emotion is still being negotiated.

I found *The Cognitive Structure of Emotions* intriguing, ingenious, and thought provoking. It is well organized, clearly written, and thorough. The authors state that a major goal is to show "that a systematic and comprehensive account of the cognitive antecedents of the emotions is possible" (p. ix). They certainly achieved this goal and in so doing opened fertile new areas for research. The issues are by no means settled; the authors acknowledge that (as of 1988) empirical verification of their proposal is just beginning.

References

Ortony, A. 1987. Is Guilt an Emotion? *Cognition and Emotion* 1:283–298.

Ortony, A., and Turner, T. 1990. What's Basic about Basic Emotions? *Psychological Review* 97:315–331.

Wentworth, M. 1991. HomeRemedy: A Case-Based Parenting Skills Tutor. Master's Thesis, CS-91-158, Computer Science Dept., Brandeis Univ.

Roland Zito-Wolf is a Ph.D. student in AI at Brandeis University, Waltham, Massachusetts. He received his B.S. in electrical engineering and computer science from the Massachusetts Institute of Technology in 1980 and his M.S. in computer science from Brandeis in 1990. His areas of interest include commonsense learning, knowledge representation, and information systems.