Third Annual Conference of the Cognitive Science Society. Berkeley, CA, pp. 170–172.
McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart, Ed., *Attention and Performance XII: The Psychology of Reading.* London: Erlbaum, pp. 1–36.
McClelland, J. L. (1991). Stochastic interactive activation and the effect of context on perception. *Cognitive Psychology* 23: 1–44.
McClelland, J. L. (1994). The interaction of nature and nurture in development: a parallel distributed processing perspective. In P. Bertelson, P. Eelen, and G. D'Ydewalle, Eds., *International Perspectives on Psychological Science,* vol. 1: *Leading Themes.* Hillsdale, NJ: Erlbaum, pp. 57–88.
McClelland, J. L., and J. L. Elman. (1986). The TRACE model of speech perception. *Cognitive Psychology* 18: 1–86.
McClelland, J. L., B. L. McNaughton, and R. C. O'Reilly. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* 102: 419–457.
McClelland, J. L., and D. E. Rumelhart. (1981). An interactive activation model of context effects in letter perception: Part 1: an account of basic findings. *Psychological Review* 88: 375–407.
McClelland, J. L., and D. E. Rumelhart. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General* 114: 159–188.
Mozer, M. C. (1987). Early parallel processing in reading: a connectionist approach. In M. Coltheart, Ed., *Attention and Performance XII: The Psychology of Reading.* London: Erlbaum, pp. 83–104.
Mozer, M. C., and M. Behrmann. (1990). On the interaction of selective attention and lexical knowledge: a connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience* 2: 96–123.
Murre, J. M. (1997). Implicit and explicit memory in amnesia: some explanations and predictions of the tracelink model. *Memory* 5: 213–232.
Nystrom, L. E., and J. L. McClelland. (1992). Trace synthesis in cued recall. *Journal of Memory and Language* 31: 591–614.
Phaf, R. H., A. H. C. van der Heijden, and P. T. W. Hudson. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology* 22: 273–341.
Pinker, S. (1991). Rules of language. *Science* 253: 530.
Pinker, S., and A. Prince. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28: 73–193.
Plaut, D. C., J. L. McClelland, M. S. Seidenberg, and K. E. Patterson. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review* 103: 56–115.
Plunkett, K., and V. A. Marchman. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* 48: 21–69.
Rohde, D., and C. C. Plaut. (forthcoming). *Simple Recurrent Networks and Natural Language: How Important is Starting Small?* Pittsburgh, PA: Center for the Neural Basis of Cognition, Carnegie Mellon and the University of Pittsburgh.
Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic, Ed., *Attention and Performance VI.* Hillsdale, NJ: Erlbaum.
Rumelhart, D. E. (1989). Toward a microstructural account of human reasoning. In S. Vosniadou and A. Ortony, Eds., *Similarity and Analogical Reasoning.* New York: Cambridge University Press, pp. 298–312.
Rumelhart, D. E., and J. L. McClelland. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group, Eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* vol. 2. Cambridge, MA: MIT Press, pp. 216–271.
Rumelhart, D. E., P. Smolensky, J. L. McClelland, and G. E. Hinton. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group, Eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* vol. 2. Cambridge, MA: MIT Press, pp. 7–57.
Rumelhart, D. E., and P. M. Todd. (1993). Learning and connectionist representations. In D. E. Meyer and S. Kornblum, Eds., *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence and Cognitive Neuroscience.* Cambridge, MA: MIT Press, pp. 3–30.
Servan-Schreiber, D., A. Cleeremans, and J. L. McClelland. (1991). Graded state machines: the representation of temporal contingencies in simple recurrent networks. *Machine Learning* 7: 161–193.
Shastri, L., and V. Ajjanagadde. (1993). From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences* 16: 417–494.
Smolensky, P., G. Legendre, and Y. Miyata. (forthcoming). *Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition.* Hillsdale, NJ: Erlbaum. (Also available as *Technical Report CU-CS-600-92.* Boulder: Computer Science Department and Instutute of Cognitive Science, University of Colorado at Boulder.)
Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys and humans. *Psychological Review* 99: 195–231.
St. John, M. F. (1992). The story gestalt: a model of knowledge-intensive processes in text comprehension. *Cognitive Science* 16: 271–306.
St. John, M. F., and J. L. McClelland. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence* 46: 217–257.
Tanenhaus, M. K., M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268: 1632–1634.
Taraban, R., and J. L. McClelland. (1988). Constituent attachment and thematic role assignment in sentence processing: influences of content-based expectations. *Journal of Memory and Language* 27: 597–632.

Cognitive Modeling, Symbolic

Symbolic cognitive models are theories of human cognition that take the form of working computer programs. A cognitive model is intended to be an explanation of how some aspect of cognition is accomplished by a set of primitive computational processes. A model performs a specific cognitive task or class of tasks and produces behavior that constitutes a set of predictions that can be compared to data from human performance. Task domains that have received considerable attention include problem solving, language comprehension, memory tasks, and human-device interaction.

The scientific questions cognitive modeling seeks to answer belong to *cognitive psychology,* and the computational techniques are often drawn from *artificial intelligence.* Cognitive modeling differs from other forms of theorizing in psychology in its focus on functionality and computational completeness. Cognitive modeling produces both a theory of human behavior on a task and a computational artifact that performs the task.

The theoretical foundation of cognitive modeling is the idea that cognition is a kind of COMPUTATION (see also COMPUTATIONAL THEORY OF MIND). The claim is that what the mind does, in part, is perform cognitive tasks by computing. (This does not mean that the computer is a metaphor for the mind, or that the architectures of modern digital computers can give us insights into human mental architecture.) If this is the case, then it must be possible to explain cognition as a dynamic unfolding of computational processes. A cognitive model cast as a computer program is a precise description of what those processes are and how they develop over time to realize some task.

A cognitive model is considered to be a *symbolic* cognitive model if it has the properties of a symbolic system in the technical sense of Newell and Simon's (1976) physical symbol system hypothesis (PSSH). The PSSH provides a hypothesis about the necessary and sufficient conditions for a physical system to realize intelligence. It is a reformulation of Turing computation (see CHURCH-TURING THESIS) that identifies symbol processing as the key requirement for complex cognition. The requirement is that the system be capable of manipulating and composing symbols and symbol structures—physical patterns with associated processes that give the patterns the power to denote either external entities or other internal symbol structures (Newell 1980; Newell 1990; Pylyshyn 1989; Simon 1996). One of the distinguishing characteristics of symbols systems is that novel structures may be composed and interpreted, including structures that denote executable processes.

The extent to which symbolic processing is required for explaining cognition, and the extent to which *connectionist* models have symbolic properties, has been the topic of ongoing debates in cognitive science (Fodor and Pylyshyn 1988; Rumelhart 1989; Simon 1996; see COGNITIVE MODELING, CONNECTIONIST). Much of the debate has turned on the question of whether or not particular connectionist systems are able to compose and interpret novel structures. In particular, Fodor and Pylyshyn argue that any valid cognitive theory must have the properties of productivity and systematicity. Productivity refers to the ability to produce and entertain an unbounded set of novel propositions with finite means. Systematicity is most easily seen in linguistic processing, and refers to the intrinsic connection between our ability to produce and comprehend certain linguistic forms. For example, no speaker of English can understand the utterance "John loves the girl" without also being able to understand "the girl loves John," or any other utterance from the unbounded set of utterances of the form "X loves Y." Both productivity and systematicity point to the need to posit underlying abstract structures that can be freely composed, instantiated with novel items, and interpreted on the basis of their structure.

A variety of empirical constraints may be brought to bear on cognitive models. These include: basic *functionality* requirements (a model must actually perform the task to some approximation if it is to be veridical); data from *verbal protocols* of human subjects thinking aloud while PROBLEM SOLVING (these reveal intermediate cognitive steps that may be aligned with the model's behavior; Newell and Simon 1972; Ericsson and Simon 1984); *chronometric data* (such data can constrain a cognitive model once assumptions are made about the time course of the component computational processes; Newell 1990); *eye movement* data (eye fixation durations are a function of cognitive, as well as perceptual, complexity; Carpenter and Just 1987; Rayner 1977); *error patterns;* and data on *learning rates* and *transfer of cognitive skill* (such data constrain the increasing number of cognitive models that are able to change behavior over time; Singley and Anderson 1989).

Though the problem of under-constraining data is a universal issue in science, it is sometimes thought to be particularly acute in computational cognitive modeling, despite the variety of empirical constraints described above. There are two related sides to the problem. First, cognitive models are often seen as making many detailed commitments about aspects of processing for which no data distinguish among alternatives. Second, because of the universality of computational frameworks, an infinite number of programs can be created that mimic the desired behavior (Anderson 1978).

Theorists have responded to these problems in a variety of ways. One way is to adopt different levels of abstraction in the theoretical statements: in short, not all the details of the computer model are part of the theory. NEWELL (Newell 1990; Newell et al. 1991), MARR (1982), PYLYSHYN (1984) and others have developed frameworks for specifying systems at multiple levels of abstraction. The weakest possible correspondence between a model and human cognition is at the level of input/output: if the model only responds to functionality constraints, it is intended only as a sufficiency demonstration and formal task definition (Pylyshyn 1984, 1989). The strongest kind of correspondence requires that the model execute the same algorithm (take the same intermediate computational steps) as human processing. (No theoretical interpretation of a cognitive model, not even the strongest, depends on the hardware details of the host machine.)

An important method for precisely specifying the intended level of abstraction is the use of programming languages designed for cognitive modeling, such as PRODUCTION SYSTEMS. Production systems were introduced by Newell (1973) as a flexible model of the control structure of human cognition. The flow of processing is not controlled by a fixed program or procedure laid out in advance, as is the case in standard procedural programming languages. Instead, production systems posit a set of independent production rules (condition-action pairs) that may fire any time their conditions are satisfied. The flow of control is therefore determined at run time, and is a function of the dynamically evolving contents of the working memory that triggers the productions. A cognitive model written in a production system makes theoretical commitments at the level of the production rules, and defines a computationally complete system at that level. The particular underlying implementation (e.g., LISP or Java) is theoretically irrelevant.

A complementary approach to reducing theoretical degrees of freedom is to apply the same model with minimal variation to a wide range of tasks. Each new task is not an unrelated pool of data to be arbitrarily fitted with a new model or with new parameters. For example, a computational model of short-term memory that accounts for immediate serial recall should also apply, with minimal strategy variations, to free recall tasks and recognition tasks as well (Anderson and Matessa 1997).

Recent cognitive modeling research combines these approaches by building and working with cognitive architectures. A COGNITIVE ARCHITECTURE posits a fixed set of computational mechanisms and resources that putatively underlie a wide range of human cognition. As these cognitive architectures never correspond to the architectures of modern computers (for example, they may demand a higher degree of parallelism), the architectures must first be *emulated* on computers before cognitive models can be built within them for specific tasks. Such architectures, together with the variety of empirical constraints outlined above, place considerable constraint on task models.

Examples of the architectural approach include ACT-R (Anderson 1993), CAPS (Just and Carpenter 1992), SOAR (Newell 1990), EPAM (Feigenbaum and Simon 1984), and Epic (Meyer and Kieras 1997). (All are production systems, with the exception of EPAM.) These architectures have collectively been applied to a broad set of phenomena in cognitive psychology. For example, Anderson and colleagues (Anderson 1993; Singley and Anderson 1989) have demonstrated that a production rule analysis of cognitive skill, along with the learning mechanisms posited in the ACT architecture, provide detailed and explanatory accounts of a range of regularities in cognitive skill acquisition in complex domains such as learning to program LISP. ACT also provides accounts of many phenomena surrounding the recognition and recall of verbal material (e.g., the fan effect), and regularities in problem-solving strategies (Anderson 1993; Anderson and Lebiere forthcoming). EPAM is one of the earliest computational models in psychology and accounts for a significant body of data in the learning and high-level perception of verbal material. It has been compared in some detail to related connectionist accounts (Richman and Simon 1989). SOAR is a learning architecture that has been applied to domains ranging from rapid, immediate tasks such as typing and video game interaction (John, Vera and Newell 1994) to long stretches of problem-solving behavior (Newell 1990), building on the earlier analyses by Newell and Simon (1972). SOAR has also served as the foundation for a detailed theory of sentence processing, which models both the rapid on-line effects of semantics and context, as well as subtle effects of syntactic structure on processing difficulty across several typologically distinct languages (Lewis 1996, forthcoming). EPIC is a recent architecture that combines a parallel production system with models of peripheral and motor components, and accounts for a substantial body of data in the performance of dual cognitive tasks (Meyer and Kieras 1997). CAPS is a good example of recent efforts in symbolic modeling to account for individual differences in cognitive behavior. CAPS explains differences in language comprehension performance by appeal to differences in working memory capacity (Just and Carpenter 1992). Polk and Newell (1995) developed a constrained parametric model of individual differences in syllogistic reasoning that provides close fits to particular individuals by making different assumptions about the way they interpret certain linguistic forms (see also Johnson-Laird and Byrne 1991).

In short, modern symbolic cognitive modeling is characterized by detailed accounts of chronometric data and error patterns; explorations of the explanatory role of the same basic architectural components across a range of cognitive tasks; attempts to clearly distinguish the contributions of relatively fixed architecture and more plastic task strategies and background knowledge; and attempts to explicitly deal with the problem of theoretical degrees of freedom. The underlying goal of all these approaches is to produce more unified accounts of cognition explicitly embodied in computational mechanisms.

See also KNOWLEDGE-BASED SYSTEMS; KNOWLEDGE REPRESENTATION; RULES AND REPRESENTATIONS

—Richard L. Lewis

References

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review* 85: 249–277.

Anderson, J. R. (1993). *The Adaptive Character of Thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R., and M. Matessa. (1997). A production system theory of serial memory. *Psychological Review* (forthcoming). AuQ: Update?

Anderson, J. R., and C. Lebeire. (Forthcoming). *ACT: Atomic Components of Thought.* Hillsdale, NJ: Lawrence Erlbaum.

Ericcson, K. A., and H. A. Simon. (1984). *Protocol Analysis: Verbal Reports as Data.* Cambridge, MA: MIT Press.

Feigenbaum, E., and H. A. Simon. (1984). EPAM-like Models of Recognition and Learning. *Cognitive Science* 8:305–336.

Fodor, J. A., and Z. W. Pylyshyn. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition* 28:3–71.

John, B. E., A. H. Vera, and A. Newell. (1994). Toward real-time GOMS: A model of expert behavior in a highly interactive task. *Behavior and Information Technology* 13: 255–267.

Johnson-Laird, P. N., and R. M. J. Byrne. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum.

Just, M. A., and P. A. Carpenter. (1987). *The Psychology of Reading and Language Comprehension.* Boston: Allyn and Bacon.

Just, M. A., and P. A. Carpenter. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99 (1): 122–149.

Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research.*

Lewis, R. L. (Forthcoming). *Cognitive and Computational Foundations of Sentence Processing.* Oxford: Oxford University Press.

Marr, D. (1982). *Vision.* New York: Freeman.

Meyer, D. E., and D. E. Kieras. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1: basic mechanisms. *Psychological Review* 104(1): 3–65.

Newell, A. (1973). Production systems: models of control structures. In W. G. Chase, Ed., *Visual Information Processing.* San Diego, CA: Academic Press, pp. 463–526.

Newell, A. (1980). Physical symbol system. *Cognitive Science* 4; 135–183.

Newell, A. (1990). *Unified Theories of Cognition.* Cambridge, MA: Harvard University Press.

Newell, A., and H. A. Simon. (1972). *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.
Newell, A., and H. A. Simon. (1976). Computer science as empirical inquiry: symbols and search. *Communications of the ACM* 19: 113–126.
Newell, A., G. Yost, J. E. Laird, P. S. Rosenbloom, and E. Altmann. (1991). Formulating the problem-space computational model. In R. F. Rashid, Ed., *CMU Computer Science: A 25th Anniversary Commemorative.* Reading, MA: Addison-Wesley.
Polk, T. A., and A. Newell. (1995). Deduction as verbal reasoning. *Psychological Review* 102 (3): 533–566.
Pylyshyn, Z. W. (1984). *Computation and Cognition.* Cambridge, MA: Bradford/MIT Press.
Pylyshyn, Z. W. (1989). Computing in cognitive science. In M. I. Posner, Ed., *Foundations of Cognitive Science.* Cambridge, MA: MIT Press.
Rayner, K. (1977). Visual attention in reading: eye movements reflect cognitive processes. *Memory and Cognition* 4: 443–448.
Richman, H. B., and H. A. Simon. (1989). Context effects in letter perception: comparison of two theories. *Psychological Review* 96 (3): 417–432.
Rumelhart, D. E. (1989). The architecture of mind: a connectionist approach. In M. I. Posner, Ed., *Foundations of Cognitive Science.* Cambridge, MA: MIT Press.
Simon, H. A. (1996). The patterned matter that is mind. In D. Steier and T. Mitchell, Eds., *Mind Matters: Contributions to Cognitive and Computer Science in Honor of Allen Newell.* Hillsdale, NJ: Erlbaum.
Singley, M. K., and J. R. Anderson. (1989). The transfer of cognitive skill. Cambridge, MA: Harvard University Press.

Color Categorization

Lexical color categorization consists of the division of color sensations into classes corresponding to the significata of the color words of a particular language. Perceptual color categorization consists of the division of the color sensations into classes by the perceptual processes of an organism—human or nonhuman, adult or neonate, possessed of knowledge of a language or not so possessed. Conflict among views on the relationship of lexical to perceptual color categorization has prevailed for over a century. Nineteenth-century classicists, anthropologists, and opthalmologists were aware that all languages do not reflect identical lexical classifications of color. The classicist (and statesman) William Gladstone concluded that differences in color lexicons reflect differences in perceptual abilities, for example, "that the organ of color and its impressions were but partially developed among the Greeks of the heroic age" (see Berlin and Kay 1969). The opthalmologist Hugo Magnus recognized that failure to distinguish colors lexically need not indicate inability to distinguish them perceptually (see Berlin and Kay 1969: 144ff). These and other late nineteenth-century scholars strongly tended to view differences in color lexicons in evolutionary terms.

In the 1920s, 1930s, and 1940s, Edward SAPIR (e.g., 1921: 219) and B. L. Whorf (e.g., 1956 [1940]: 212ff) rejected evolutionism for the doctrine of radical linguistic and cultural relativity. The favorite field for the empirical establishment and rhetorical defense of the relativist view, which became established doctrine in the 1950s and 1960s, was the lexicon of color. With respect to color categorization, there have been two major traditions of research stemming from the relativity thesis: a within-language, correlational line of research and a cross-language, descriptive one.

Early work in the former tradition (e.g., Brown and Lenneberg 1954; Lenneberg and Roberts 1956) is primarily concerned with establishing a correlation between a linguistic variable distinguishing colors (for example, how easy different colors are to name or how easy they are to communicate about) and a nonlinguistic cognitive variable over colors: memorability. Discovery of such a correlation was interpreted as support for the Sapir-Whorf view that linguistic categorization can influence nonlinguistic perception/cognition. In the 1950s and 1960s, such correlations were reported within English and, to a limited extent, in other languages (Stefflre, Castillo Vales, and Morley 1966). Because it was assumed at the time that the linguistic variable (codability or communication accuracy) would vary across languages, correlation between a linguistic and nonlinguistic variable within a single language (almost always English) was taken to validate the docrine that the coding systems of different languages induce differences in the nonlinguistic cognition of their speakers. Eleanor Rosch (e.g., Heider 1972) challenged this assumption on the basis of the apparent universal lexical salience of certain "focal" colors (identified by Berlin and Kay 1969). Rosch showed that universal perceptual salience determines both the nonlinguistic and the linguistic variables of the correlational approach, thus undercutting the logic of this line of research. Rosch's view was criticized by Lucy and Shweder (1979), who also challenged her experimental procedure; Lucy and Shweder's experimental procedure was in turn challenged by Kay and Kempton (1984), who supported Rosch's view of the matter. (Kay and Kempton, using a noncorrelational, cross-linguistic experimental procedure, showed that certain nonlinguistic color classification judgments may be influenced by the lexical classification of color in a language, while others are not so influenced, thus re-establishing limited Whorfian effects in the color domain.)

In the tradition of cross-language description, the studies of the 1950s and 1960s likewise reflected the dominance of radical linguistic relativism (Ray 1952; Conklin 1955; Gleason 1961: 4). These studies sought to discover and celebrate the differences among color lexicons. In 1969, using the original stimulus set of Lenneberg and Roberts (1956), Berlin and Kay compared the denotation of basic color terms in twenty languages and, based on these findings, examined decriptions of seventy-eight additional languages from the literature. They reported that there are universals in the semantics of color: the major color terms of all languages are focused on one of eleven landmark colors. Further, they postulated an evolutionary sequence for the development of color lexicons according to which black and white precede red, red precedes green and yellow, green and yellow precede blue, blue precedes brown, and brown precedes purple, pink, orange and gray. These results were challenged on experimental grounds, mostly by anthropologists (e.g., Hickerson 1971; Durbin 1972; Collier 1973), and largely embraced by psychologists (e.g.,