Abstract of the Dissertation

Knowledge Acquisition Via Incremental Conceptual Clustering

by

Douglas Hayes Fisher Jr.

Doctor of Philosophy in Information and Computer Science

University of California, Irvine, 1987

Dr. Dennis Kibler, Chair

Concept learning and organization are much studied in artificial intelligence and cognitive psychology. Computational models of learning and memory that hope to be flexibly applied in real-world settings need to be incremental and improve an agent's ability to make predictions about the environment. While these are useful properties for purely artificial organisms, they also characterize much of human learning and memory.

This dissertation describes COBWEB, an incremental method of *conceptual clustering* that builds a classification hierarchy over a sequence of observations. These hierarchies are characterized in terms of their ability to improve prediction of unknown object properties. Computer experimentation and comparisons with alternate methods of classification show that COBWEB's approach effectively improves prediction ability. More generally, prediction of unknown object properties is forwarded as a performance task for all conceptual clustering systems. This opens the way for objective, not anecdotal, characterizations of and comparisons between concept formation systems.

A fundamental bias of this dissertation is that research on human learning and memory can usefully inspire directions for work on artificially intelligent systems and vice versa. Concept representations and measures of concept quality used by COBWEB are inspired by work in cognitive psychology on *typicality* and *basic level* effects. Conversely, COBWEB is the basis for a second system, COBWEB/2, that accounts for typicality and basic level effects in humans. Apparently, this is the first computational model that accounts for basic level effects. The account of typicality effects stresses the need to consider concepts in the context of a larger memory structure. This approach also facilitates speculation on possible interactions between basic level and typicality effects.

In summary, the dissertation presents an incremental method of conceptual clustering that is evaluated with respect to a prediction task. Concept representations and heuristics are borrowed from cognitive psychology, with repayment in the form of a cognitive model of basic level and typicality effects.

CHAPTER 1
Introduction

1.1 Contributions of the Dissertation

Classification is the basis of inferential capacity and is critical to the success of any intelligent organism. Artificial intelligence (AI) and cognitive psychology seek to explain the form and acquisition of classification structures and processes. This dissertation reports two systems for building classification schemes that have been influenced by principles of AI and cognitive psychology.

From an AI or machine learning standpoint this dissertation addresses the problem of learning under two assumptions. The first is that concept learning is *incremental*; objects are incorporated into a classification structure as they are observed. Second, concept learning should increase the correctness of predictions made about the environment. These assumptions are studied within the context of *conceptual clustering*, a machine learning task concerned with building classification structures.

From a cognitive psychology standpoint, constraints on human classification can illuminate principles of (human and machine) intelligence, generally. In particular, this dissertation takes advantage of research on *basic level* and *typicality* effects observed during human classification. These effects suggest principled ways of measuring concept quality, representing concepts, and classifying objects.

Two foci of interest, psychological and computational, represent apparently dichotomous objectives. However, these interests are cooperative and the interplay between them yields insights that are incorporated into two concept formation systems. COBWEB is an incremental conceptual clustering system that attempts

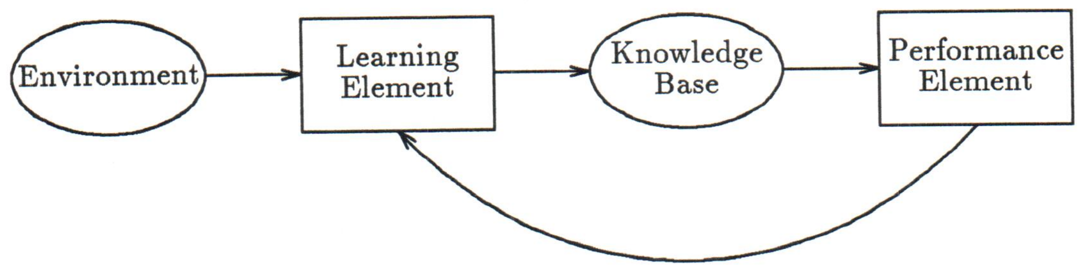

Figure 1

A model of learning and performance

to maximize the ability to correctly predict unknown object properties. To do this it uses a measure of concept quality inspired by psychological studies. Furthermore, incremental processing and inference ability characterize much of human learning and memory. From COBWEB, a second system, COBWEB/2 is derived. This system builds classification hierarchies that account for basic level and typicality phenomena. From a cognitive modeling standpoint, this work appears to be the first computational model to account for basic level effects, and its explanation of typicality effects has several advantages over previous accounts.

1.2 Conceptual Clustering

Machine learning is concerned with improving performance by automating knowledge acquisition and refinement. This view is reflected by the simple model of learning and performance in Figure 1 [Diet82]. Learning organizes observations into a knowledge base that facilitates performance with respect to some task. Assumptions about environment, knowledge base, and performance all impact the design of a learning algorithm and delineate general learning tasks. For instance, *learning from examples* assumes that objects (states, events, etc.) come preclassified with respect to a number of 'teacher' defined classes. Under this environmental assumption a learner induces concepts for each object class. Learning to diagnose soybean disease from examples [Mich81] assumes that a 'teacher' identifies the

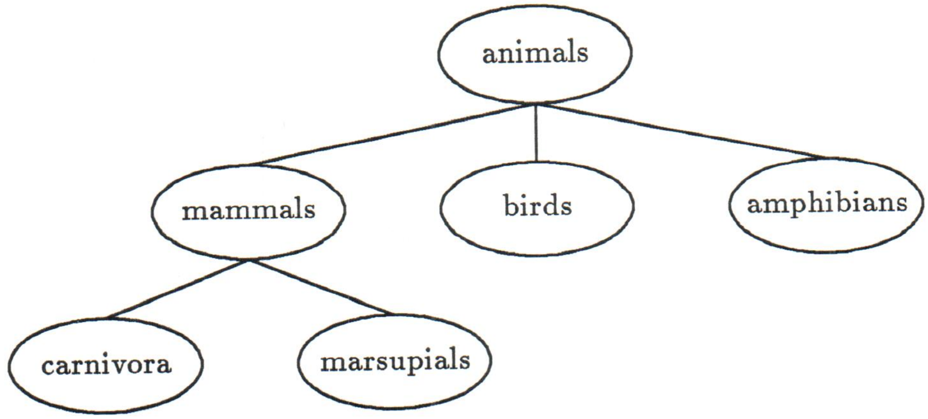

Figure 2

An example classification tree

disease (or lack of disease) of soybean plant case histories. Over several case histories the learner induces rules or concepts that allow it to independently identify diseases categories. Learning from examples has been applied in numerous domains [WINS75, HAYE78, VERE80, MITC83, PORT84, BRAD87, SCH86B], but in every system that learns from examples, performance reduces to matching previously unseen 'objects' against induced concepts, thus identifying their class membership (e.g., an example of a particular soybean disease).

In contrast to learning from examples, *conceptual clustering* systems [MICH80] accept a number of object descriptions and produce a classification scheme over the observed objects. For example, a conceptual clustering system might form a classification tree over a number of animal descriptions as shown in Figure 2. These systems do not require a 'teacher' to preclassify objects, but use an evaluation function to find classes with 'good' concept descriptions. Concept descriptions may be stored at classification tree nodes. For example, the 'mammals' node of Figure 2 might be characterized by the concept, **has-hair** $\wedge$ **bears-living-young**. Conceptual clustering is a type of *learning by observation* or *concept formation* (as opposed to *learning from examples*). However, the recency of conceptual clustering's definition has allowed little exploration of it in the context of environment and performance.

The most important contextual factor surrounding learning is the performance task that benefits from it. Unfortunately, this task is ill-defined or not discussed at all with respect to most conceptual clustering work (and thus the often asked question, "How do you know the classifications you get are any good?"). However, some attempts have been made to evaluate conceptual clustering with respect to a performance task. For example, Cheng and Fu [CHEN85] and Fu and Buchanan [FU85] use clustering techniques to facilitate disease diagnosis in expert systems. Generalizing (and clarifying) their use of conceptual clustering, classifications can be the basis for effective prediction of unseen object properties. The generality of classification as a means of guiding inference is manifest in recent discussions of problem-solving as classification [CLAN84]. For example, having recognized an animal with respect to the 'mammals' node of Figure 2 – say by virtue of it having hair – a prediction that it bears-living-young can be made. In a medical domain, a set of symptoms may suggest a particular disease, from which a treatment can be inferred. The first system described in this dissertation, COBWEB, is designed to form classification trees that are good predictive models of the environment.

A second factor surrounding learning is the environment. In particular, conceptual clustering systems have assumed that environmental inputs are indefinitely available for examination and thus the environment is amenable to nonincremental processing of observations. However, real world environments encourage incremental object assimilation [CARB86, LAN86A, SAMM86] and systems that process observations in this fashion are gaining prominence [REIN85, SCH86A, LEBO82, KOL83A]. In response to real world considerations, COBWEB has been constructed as an incremental system of conceptual clustering. Its underlying control mechanisms are abstracted from previous work on incremental concept formation, notably Lebowitz' UNIMEM [LEBO82] and Kolodner's CYRUS [KOL83A]. However, unlike these precursors, COBWEB is evaluated along a number of dimensions related to the cost and quality of learning.

This dissertation imposes the framework of conceptual clustering onto incremental concept formation systems like those developed by Lebowitz [LEBO82] and Kolodner [KOL83A]. This combination extends the traditional conceptual clustering literature to include incremental processing and clarifies the processing characteristics of these other incremental systems. In addition, this work suggests prediction of missing object properties as a performance task for conceptual clustering.

1.3 Basic Level and Typicality Effects

The processing strategies of COBWEB borrow from work in AI and machine learning. However, the AI influence is balanced with results from cognitive psychology. Many aspects of human intelligence demonstrate important principles of general intelligence. In the context of classification, two phenomena are of particular interest. The first is that members of a class are not regarded as equally representative, but vary along a dimension of *typicality* [MERV81, SMIT81]. For example, a *robin* is more quickly recognized as a *bird* than is a *penguin*. The observation that some instances are regarded as more typical of a class than others does not jibe with assumptions often associated with logical, typically conjunctive, concepts [SMIT81]. The limitations of logical representations motivates the use of *probabilistic concepts* in COBWEB. Probabilistic concepts associate probabilities or other confidence measures with object class properties. For example, a platypus is a mammal that lays eggs. A probabilistic concept for 'mammals' would indicate that a mammal has-hair with probability, 1.0, and bears-living-young with probability, 0.97, to accommodate platypi. Generally, probabilities add information to concept representations that can be exploited during classification and inference.

Other studies indicate the tendency of humans to prefer a particular conceptual level in hierarchical classification schemes [Ros76a, Joli84]. For instance, when asked to identify a *collie*, a subject will respond that it is a *dog* rather than a *collie* or *mammal*. This task, and a host of others [Merv81], indicate that in a hierarchy containing (*collie, dog, mammal, animal*), *dog* is the preferred or *basic level* concept. The identification of preferred concepts in humans suggests principled measures of concept quality in AI systems. COBWEB uses a measure of concept quality called *category utility* [Gluc85] that was inspired by basic level studies. Category utility assumes probabilistic information is known regarding class members, thus reinforcing the choice of probabilistic concepts. Moreover, category utility rewards concepts that facilitate prediction and is therefore compatible with COBWEB's performance goals.

Basic level and typicality effects motivate concept representation and evaluation in COBWEB. These psychological considerations do not interfere with the computational goals of incremental processing and utility of classifications for inference. Rather, probabilistic concepts and category utility are completely compatible with these goals.

Although its design is influenced by psychological concerns, COBWEB should not be regarded as a cognitive model *per se*. However, its environmental (i.e., incremental processing), performance (i.e., inference), and knowledge base (i.e., hierarchical classifications and probabilistic concepts) assumptions are consistent with much of human learning and memory. As a result, the memory structures of COBWEB are the basis of a second memory model that accounts for typicality and basic level effects observed during human classification. These hierarchies use an indexing scheme adapted from work by Lebowitz [LEBO82] and Kolodner [KOL83A] and they demonstrate how various pieces of partial evidence combine to produce the desired psychological effects. Indexed classification hierarchies are learned by COBWEB/2, a second system that demonstrates advantages and problems with indexed memory when learning.

1.4 An Overview of Computational and Psychological Antecedents

In summary, this dissertation draws upon work from AI and cognitive psychology. Work in conceptual clustering and incremental concept formation contributes to COBWEB's and COBWEB/2's control mechanisms, while work in cognitive psychology suggests concept representations and quality preferences. Specific antecedents and their contributions to this work are pictured in Figure 3. While interest in natural versus artificial intelligence traditionally divides research efforts in AI [HALL85, NEWE73], they are intertwined in this dissertation.

1.5 Methodological Biases

AI is an evolving discipline, amalgamating concepts from several fields, including computer science and psychology. As a result there is no consensus among AI practitioners as to which research problems are important, which methodologies are productive, and in general, what constitutes 'good research' [HALL85]. The burden of identifying important questions, productive methodologies, and evaluation criteria is placed on the individual researcher; these are not explicit and well understood constraints of the field as a whole. However, research communities have emerged within the field, the identification of which can aid in guiding and reporting research.

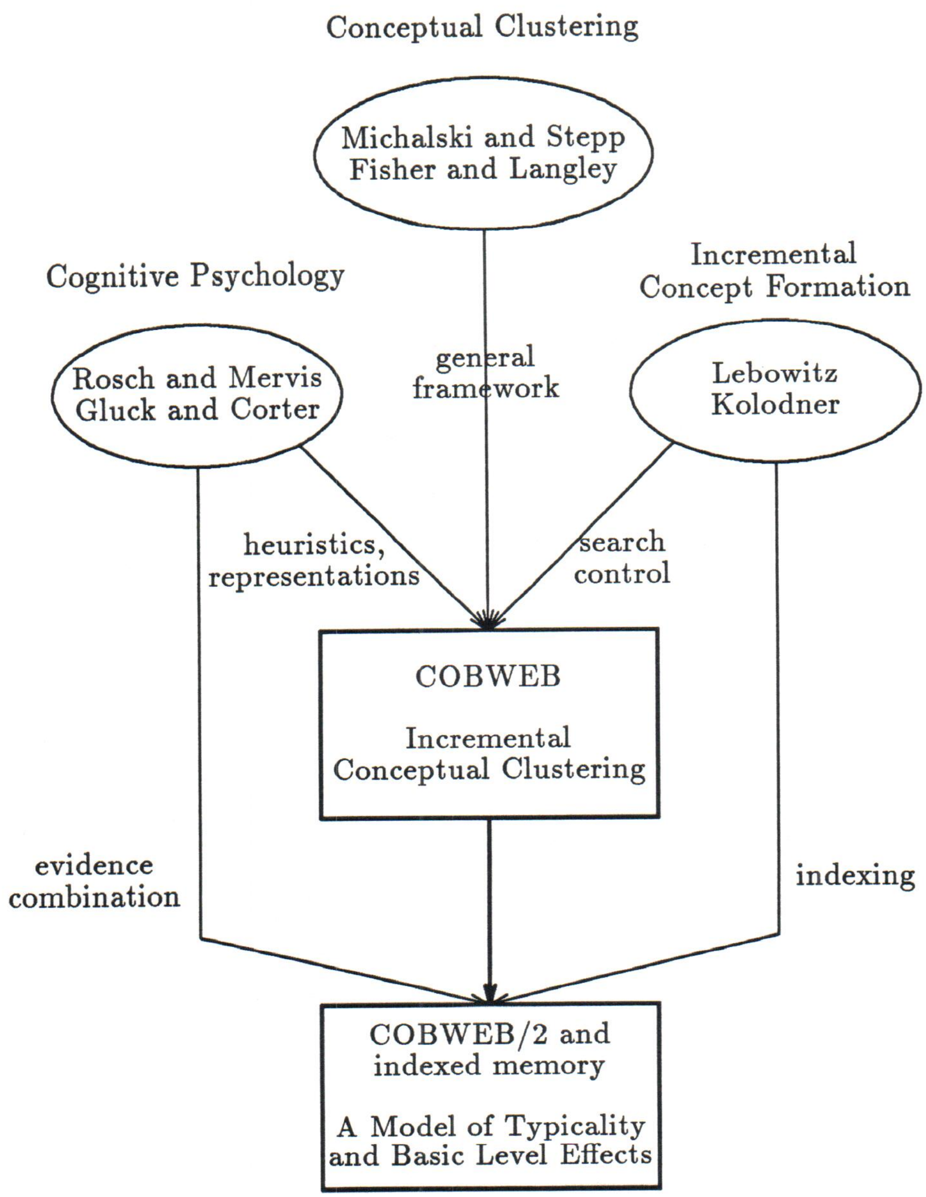

Figure 3

Antecedents to COBWEB

1.5.1 Methodological Perspectives in AI

Hall and Kibler [HALL85] have recently forwarded a taxonomy of methodological perspectives in AI. They initially divide perspectives by interest in natural intelligence versus purely artificial intelligence. Natural approaches are further broken down by the specificity of the natural behavior that a computer system is expected to model.

Empirical approaches are characterized by a validation of system behavior with respect to tightly constrained experimental evidence. Experiments need not be performed by the 'cognitive modeler', but may be performed previous to and independent of the computational model. An existing database of experimental findings can be used for comparative purposes. Hall and Kibler cite GPS [Newe63] as an example of this perspective. Feigenbaum's EPAM [Feig63] is another.

In contrast, *speculative* approaches look to natural behavior for initial inspiration, introspect as to the rules guiding this behavior, and validate the resultant computer system by gross comparison of system and natural behavior. Speculative methods are not constrained by specific experimental evidence, but seek general principles by looking to 'general' behavior. Hall and Kibler give Schank and Abelson's [Scha77] theory of scripts as an exemplar of the speculative approach.

Empirical and speculative approaches can be viewed as differing in the 'grain size' of the natural phenomena that are used to validate the cognitive model. The empirical approach dictates validation with respect to tightly constrained behavior, while research efforts following the speculative approach are compared with natural behavior of less specificity. Importantly, this distinction does not imply that the *mechanisms* of an empirical artifact be special purpose. In fact, one property (intended or not) of many empirically motivated studies is that the cognitive model's mechanisms move beyond the experimental evidence and allow predictions about natural behavior that was not the original focus of study. More generally, mechanisms suggested by either perspective may be transported outside the realm of psychological interest entirely. The means-ends strategy of GPS and semantic nets [Quil68] are well-known examples of formalisms that were initiated for cognitive modeling, but that have been adopted by artificial intelligence generally.

In contrast to studies of natural behavior, Hall and Kibler propose three perspectives interested in strictly 'artificial' intelligence. *Constructive* AI forces general principles of intelligence to emerge by designing and building computer systems that address complex but specific real-world problems. For example, DENDRAL [Buch69] illuminated general issues of knowledge-intensive or expert systems while focusing on the specific task of identifying molecular structure.

Analysis of heuristic search (A*) [Hart68] is an example of work in the *formal* perspective of AI. In general, formal work seeks to unify a body of disparate work under a single, generalizing framework. Additionally, Hall and Kibler stipulate that this unifying framework be characterized formally or analytically (e.g., by proofs of correctness).

Finally, *performance* AI seeks to achieve expert behavior, with little concern towards extracting important processing principles that underlie performance. Performance AI should not be identified with every system concerned with a performance task, but only with systems that are concerned with performance to the exclusion of underlying processing principles.

1.5.2 The Dissertation in Perspective

This dissertation reflects several of the approaches outlined by Hall and Kibler. With important qualifications, the development of COBWEB resembles a formal study. The conceptual clustering framework proposed by Michalski and Stepp [Mic83a, Mic83b], and elaborated by Fisher and Langley [Fis85a, Fis86a], clarifies the basic control mechanisms of existing incremental concept formation systems [Lebo82, Kol83a]. This inspires the basic processing assumptions of COBWEB. In addition, the system uses probabilistic concepts and a principled measure of concept quality, as opposed to logical representations or the statistically-based, but *ad hoc*, representations and measures found in [LEBO82, KOL83A]. This approach highlights issues of representation and evaluation that are difficult to extract from more *ad hoc* techniques and also suggests several dimensions for evaluating similar systems. Prediction of unseen object properties is implied as a performance task for conceptual clustering systems and criteria relating to the cost and quality of learning are suggested as dimensions for evaluating incremental concept formation systems.[1]

Like formal approaches, work on COBWEB seeks to clarify and cast new light on existing work. However, the characterization of COBWEB is not analytical, but relies instead on empirical validation via extensive computer experimentation. This type of process characterization is influenced by Quinlan [QUIN86] and others [HAMP83, SCH86A] and finds it roots in work on pattern recognition and data analysis [DUDA73, EVER80]. However, COBWEB's empirical characterization is novel in several respects, most notably as it relates to prediction ability. The system's ability to make accurate predictions is compared to two alternative methods: a 'straw man' and a better known system for learning from examples.

Furthermore, COBWEB is not only characterized in a number of domains, but a measure for characterizing the domain itself is forwarded. In general, little attention has been paid to Simon's point [SIMO69] that domains must be characterized before the advantage of a learning system can be evaluated. Collectively, computer experiments are used to address the same issues as more formal methods, e.g., system behavior under varying conditions. There is no debate that when possible, a formal analysis is better than an empirical one. But when a system (or

[1] Importantly, control strategies and representations used by COBWEB were abstracted from or inspired by existing systems; they did not emerge (in this study) as the result of exploring concept formation in a highly constrained domain. Thus, COBWEB's development is not constructive. However, the development of some of COBWEB's precursors, particularly UNIMEM and CYRUS, might be so classified.

researcher) is not amenable to formal analysis, empirical studies may enumerate the dimensions along which some future, more formal analysis can proceed.

Although COBWEB is of computational interest, the work reported in the following chapters also reflects empirical and speculative approaches to cognitive modeling. In particular, COBWEB classification trees are the basis for an indexed memory structure that accounts for certain basic level and typicality effects. While this dissertation does not include psychological experimentation, a significant body of existing research supports the existence of these phenomena. However, there are difficulties with using all of this data for comparative purposes. Many experimental studies use natural domains (e.g., animals), but in such domains there is no way of knowing the properties that human subjects use to represent instances and therefore no way of assuring equivalent encodings in the computer model. Nonetheless, comparisons between human subjects and the computer model are made with respect to two experimental studies of basic level effects and one study of typicality effects, each using artificially constructed domains (e.g., nonsense strings). Artificial domains allow some experimental control over the properties to which subjects attend.

Besides the three experimental studies referred to above, the cognitive model is also characterized with respect to other tasks and domains, but these comparisons are hypothetical in nature. For example, computer experiments using a classification hierarchy over objects of the 'natural' domain of congressional voting records suggest several properties of human memory that cannot currently be verified as consistent with human behavior.

Additionally, the cognitive model is characterized with respect to some of the same tasks as COBWEB classification trees. Incremental learning and accurate prediction are important to human behavior and the ability to do these well is the classic sort of evidence admitted by speculative studies for the legitimacy of a cognitive model. However, this dissertation avoids using such an analysis as confirming evidence for the psychological validity of a model. Rather, the bias is that these analyses supply disconfirming evidence, if they supply any evidence at all. If indexed memory can not be effectively modified as learning occurs this must impact claims for its psychological validity, as well as its computational utility. In particular, the learnability of indexed classification hierarchies is investigated in the context of COBWEB/2, a system derived from COBWEB. Analysis of COBWEB/2's behavior generally indicates good learning and prediction abilities, but also uncovers a weakness of indexing. The impact of this finding on the validity of the memory model is discussed.

The methodological biases exhibited in this dissertation have been related to three of the approaches outlined by Hall and Kibler: formal, empirical, and speculative. However, problems arise when one uses their taxonomy to classify the biases of the dissertation. For example, while work on COBWEB reflects the intent of the formal approach, empirical, rather than formal characterization distinguishes it from this approach. Furthermore, the dissertation addresses computational as well as psychological concerns. This dichotomy is magnified by Hall and Kibler's initial division of methods by the intention of the researcher (i.e., interest in natural versus artificial intelligence). This division is common to other commentaries on methodological biases in AI (e.g., [NEWE73]) as well. A taxonomy that generalizes the formal approach and lessens the apparent schism between computational and psychological research is motivated and developed next.

1.5.3 A Taxonomy of AI Research

Classification schemes are rarely useful if developed in a vacuum; typically, they are motivated by some intention or goal. A fundamental bias of this dissertation is that AI depends on demonstrations of natural intelligence to supply specifications for its systems, whether this is inspirational or is more constrained

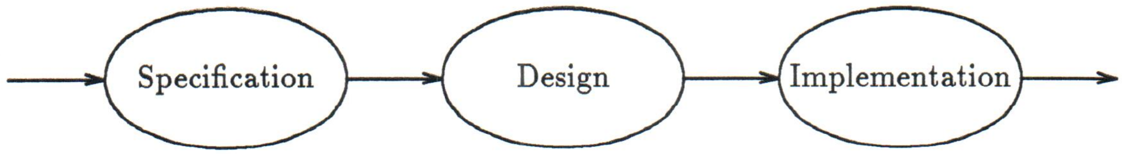

Figure 4

A view of system development

in nature. Therefore, division of research based on artificial versus natural intelligence is somewhat illusory. It may also be counterproductive since it may mask insights into methods and representations that are true across natural and artificial systems (e.g., the utility of probabilistic concepts). An alternate to the intention-centered taxonomy of Hall and Kibler is a method- or design- centered taxonomy that emphasizes the portability or generality of an *information-processing system.*

A popular information-processing view of system development distinguishes *specification*, *design*, and *implementation* (Figure 4) [Page80]. The specification is a statement of a system's function. Whether the system is a cash register, a library access system, or an expert system for medical diagnosis, system specification describes *what* the system is supposed to do; specification defines a 'black box'. The objective of system design is to outline the procedures and data representations necessary to satisfy the functionality of a black box. That is, design is concerned with *how* a system performs. Finally, implementation is concerned with realizing procedures and representations on a physical device (e.g., a computer). Within AI proper, Marr's *computational theory*, *algorithm*, and *implementation* levels [Marr82] represent a similar view of information processing systems.[2] Analogs to Hall and Kibler's perspectives can be understood in terms of how they differ along dimensions suggested by this view of information processing systems.

[2] Of course there are problems with an exact mapping between these views. In particular, Marr is intimately concerned with 'why' a computation occurs. This issue typically arises in a *requirements analysis* phase of system development, which precedes specification.

Four perspectives seek to uncover general principles or descriptions of intelligence that best fit the design level of information-processing systems. Heuristic search, evidence combination, scripts, and means-ends analysis are general processing principles that emerged from work in the formal, constructive, speculative, and empirical perspectives, respectively. Only performance AI is unconcerned with forwarding general processing principles.

Constructive, empirical, and performance AI each assume domain-dependent specifications. This assertion rests on the assumption that an 'expert' identifying molecular structures and a subject solving the eight-tile puzzle exhibit behavior of roughly the same granularity or level of specificity, despite differences in the overall complexity of these tasks. GPS and much of the work on expert systems move from specific (specifications) to general (designs).

Formal and speculative approaches are distinguished by their use of general, domain-independent specifications. Note that this does not imply ill-defined specifications. For example, A* is precisely specified. Furthermore, objections to (apparently) speculative approaches [OHLS83] may be symptomatic of 'bad speculation' and not of the speculative approach itself.

Figure 5 gives a revised taxonomy of AI perspectives. This taxonomy is heavily influenced by, but differs in a number of respects from Hall and Kibler's framework. At the top level, perspectives are divided in terms of the generality of system design or principles. Performance AI is distinguished from the others along this dimension. The remaining four perspectives are distinguished by the generality of specification or problem statement. Constructive and empirical approaches move from specific problem statements to general principles, while formal and speculative approaches assume that general mechanisms/representations are derived from general, domain-independent specifications. Finally, approaches are distinguished by the interest of the researcher in natural versus purely computational intelligence.

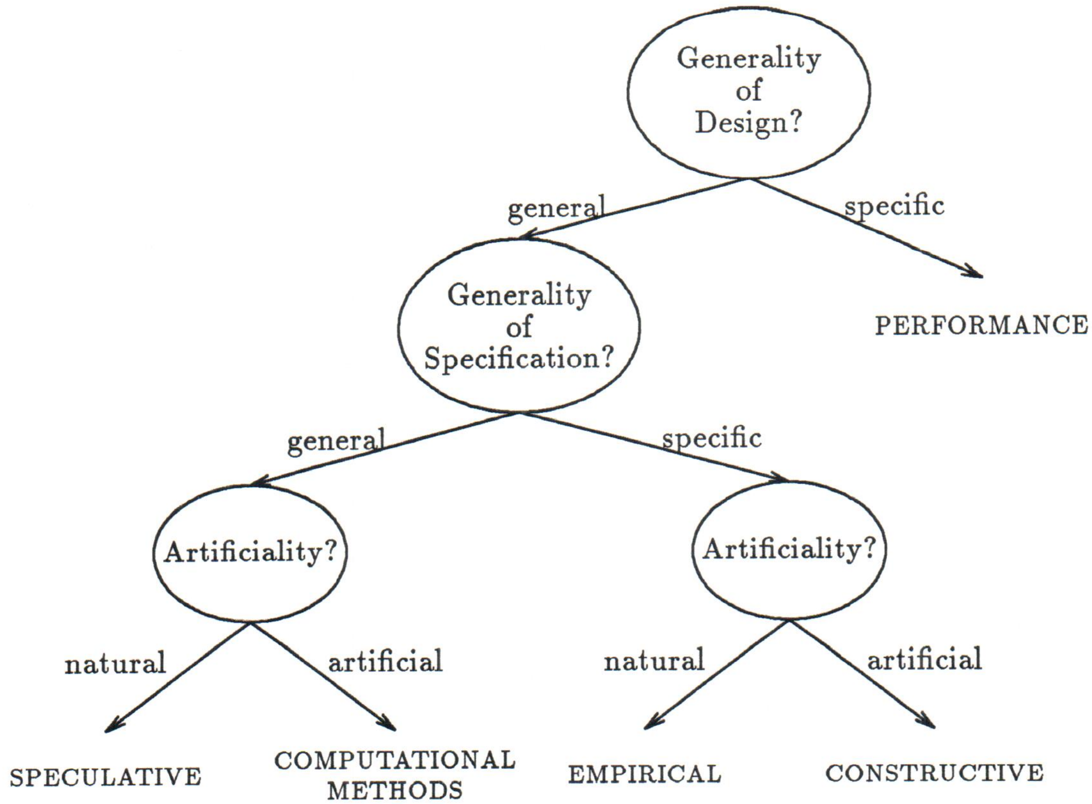

Figure 5

An information processing view of methodological perspectives in AI

At the leaves of this taxonomy, the formal approach has been generalized to one that roughly corresponds to Newell's view of AI as a quest for general 'methods' [NEWE73]. This approach shares the intent of Hall and Kibler's formal approach, but leaves the form of characterization (i.e., formal or empirical) unspecified.

The taxonomy of Figure 5 demotes the importance of the natural/artificial distinction. Rather it emphasizes the importance of design and specification generality, and thus the portability of ideas across domains and tasks. Importantly, few researchers fit precisely within one perspective. However, this taxonomy predicts that a researcher's differing perspectives will share method; constructive and empirical approaches will intermingle, as will the computational methods and speculative approaches. This seems to more accurately reflect the sort of methodological shifting that occurs than does a taxonomy that initially differentiates based on artificial and natural orientations. Hall and Kibler cite Feigenbaum's early work on EPAM [FEIG63] as an example of work in the empirical perspective, while his later work on expert systems is constructive. Closer to home, this taxonomy lessens the schism between interest in natural and computational mechanisms that is exhibited in this dissertation. Hall (personal communication) suggests that combining interests in natural and artificial intelligence can be problematic, since it facilitates confusions between claims of psychological validity with claims for computational utility. That this happens frequently can be taken as evidence for the descriptive accuracy of the taxonomy of Figure 5. Prescriptively speaking though, this sort of confusion is a flaw that the dissertation seeks to avoid.

1.6 Overview of the Dissertation

This dissertation describes COBWEB, COBWEB/2, and the classification structures formed by these systems. The presentation focuses on the design (or algorithmic) level, as opposed to implementation level descriptions. The emphasis on design-level issues clarifies the connection between these systems. This descriptive level also maximizes the 'portability' of these systems and facilitates *rational reconstruction* [BUND84].

Chapter 2 gives relevant background from machine concept learning. While this chapter describes particular systems, the goal is to present a general framework for incremental conceptual clustering. This framework is described in terms of a predominant AI paradigm: *search*. In addition, the chapter motivates and describes a performance task for conceptual clustering.

Chapter 3 describes important background from cognitive psychology on typicality and basic level effects. Results presented in this chapter are important for validating the psychological consistency of an indexing scheme presented in chapter 7 and for motivating the concept representation and evaluation measure used during conceptual clustering.

Chapter 4 describes COBWEB, an incremental and domain independent system of conceptual clustering. This system instantiates the general search framework of chapter 2.

Chapter 5 evaluates classification schemes produced by COBWEB in terms of prediction. In particular, experiments with soybean and thyroid disease diagnosis demonstrate the cost effectiveness of the approach as opposed to selected alternative methods.

Chapter 6 characterizes COBWEB along dimensions that are relevant to incremental learning systems. This chapter demonstrates that the system is computationally economical, while still robust in the sense that 'high quality' classification schemes are typically constructed.

Chapter 7 shows how COBWEB classification schemes can be modified to account for basic level and typicality effects. While results from three (human) experimental studies are explained by the classification model, support of a more hypothetical nature is garnered from (computer) experiments in the domain of congressional voting records.

Chapter 8 describes COBWEB/2, a derivative of COBWEB that incrementally builds the classification structures of chapter 7. The system's economy, robustness, and inference ability are characterized in relation to the 'ideal' COBWEB system. The fact that classification structures of this type can be learned and perform reasonably along a number of computationally important dimensions is not taken as confirming evidence for the psychological validity of the indexing scheme. Rather, problems during learning motivate a discussion of some possible weaknesses of the indexing scheme as a psychological model.

Chapter 9 concludes the dissertation with a summary of results and a prospectus of future research.

Chapter Acknowledgements

Organizing frameworks are very general, but nonetheless useful for thought and communication. Dietterich's context for learning is one such framework. So is Hall and Kibler's discussion of methodological perspectives in AI. Both frameworks were important in organizing the chapter and the dissertation as a whole. Discussions with Rogers Hall, Dennis Kibler, Pat Langley, and Jack Beusmans clarified my understanding of existing methodological perspectives in print, but they should not be held responsible for my views on these matters. Discussions with Dennis Kibler are responsible for linking conceptual clustering and performance.