Acquisition of hierarchical reactive skills in a unified cognitive architecture

Action editor: Ron Sun

Pat Langley [1], Dongkyu Choi[*], Seth Rogers [2]

Computational Learning Laboratory, Center for the Study of Language and Information, Stanford University, Stanford, CA 94305, USA

Abstract

In this paper, we review ICARUS, a cognitive architecture that utilizes hierarchical skills and concepts for reactive execution in physical environments. In addition, we present two extensions to the framework. The first involves the incorporation of means-ends analysis, which lets the system compose known skills to solve novel problems. The second involves the storage of new skills that are based on successful means-ends traces. We report experimental studies of these mechanisms on three distinct domains. Our results suggest that the two methods interact to acquire useful skill hierarchies that generalize well and that reduce the effort required to handle new tasks. We conclude with a discussion of related work on learning and prospects for additional research, including extending the framework to cover developmental phenomena.
© 2008 Elsevier B.V. All rights reserved.

Keywords: Incremental learning; Cognitive architecture; Reactive control; Problem solving; Hierarchical skills

1. Introduction and motivation

Research on cognitive architectures (Newell, 1990) attempts to understand the computational infrastructures that support intelligent behavior. A specific architecture characterizes the aspects of a cognitive agent that remain the same across time and over different domains, and typically makes strong commitments about the representation of knowledge structures and the processes that operate on them. Learning has been a central concern in most architectural research, with a variety of mechanisms having been proposed to model the acquisition of knowledge from experience. The learning methods embedded in most cognitive architectures are incremental, reflecting evidence that humans acquire knowledge in this manner, but there have been few accounts of the origin of hierarchical structures that appear crucial to complex cognition.

In this paper we review ICARUS (Langley & Choi, 2006), a candidate architecture that diverges from its predecessors on a number of dimensions. One important difference is that traditional architectures handle conceptual knowledge in a procedural manner, typically using production rules, whereas our framework contains separate memories for concepts and skills. Another distinctive feature is that most architectures are based on production systems, which encode knowledge as a 'flat' set of condition-action rules, whereas ICARUS makes an architectural commitment to the hierarchical organization of knowledge. One can certainly encode hierarchical structures in frameworks like ACT-R (Anderson, 1993) and Soar (Laird, Rosenbloom, & Newell, 1986), but this remains the modeler's choice rather than a strong theoretical claim. In addition, most cognitive architectures evolved from theories of human

---
* Corresponding author.
E-mail addresses: langley@csli.stanford.edu (P. Langley), dongkyuc@stanford.edu (D. Choi), srogers@dash.net (S. Rogers).
[1] Present address: School of Computing and Informatics, Arizona State University, Tempe, AZ 85287, USA.
[2] Present address: Dash Navigation, 888 Ross Drive, Sunnyvale, CA 94089, USA.

problem solving, which has led to subordinate roles for perception and action even in those frameworks that support them.[3] In contrast, ICARUS is primarily an execution architecture that perceives and reacts to external environments, which we view as more basic than problem solving.

However, ICARUS' reliance on hierarchical structures raises key questions about their origin. Moreover, the architecture's emphasis on execution does not mean that mental activities like problem solving are unimportant, since they can let an agent handle novel tasks for which stored knowledge is unavailable. The central hypothesis of this paper is that hierarchical skills arise, at least in many cases, from problem-solving behavior, and that, once learned, the agent can use these structures to support reactive execution in the environment. Moreover, this acquisition occurs in an incremental manner, with new skills being learned gradually as the agent encounters new problems it cannot handle without resorting to problem solving.

We refer to ICARUS as a 'cognitive architecture' in the same sense that the Soar community uses that expression. Both frameworks aim for consistency with general knowledge about human cognition and hope to support the same broad range of abilities that people demonstrate. However, our current research does not attempt to match at a fine-grained level the results of psychological experiments, as done with architectures like ACT-R. We may address such issues in future research, but for now we are concerned with coarse regularities that demand explanation, such as the apparent hierarchical nature of human skills and their incremental acquisition from experience.

In the sections that follow, we review ICARUS' representation and organization of concepts and skills, along with the inference and execution processes that utilize them. After this, we present a module that interleaves means-ends problem solving with execution when known skills are insufficient to solve a task. Next we describe a mechanism for creating generalized skills from traces of successful problem solving that supports incremental, hierarchical learning. We report experiments with this learning mechanism that demonstrate its ability to generalize to novel situations and reduce effort on new problems. In closing, we discuss earlier research on learning for problem solving and execution, along with some directions for future work. We have described this version of ICARUS in an earlier publication (Langley & Choi, 2006), but here we describe it from the perspective of cognitive architectures, and we report new experimental results that focus on cumulative effects of experience.

2. Representation and organization

Like other cognitive architectures, ICARUS makes commitments to its representation of knowledge, the manner in which that knowledge is organized, and the memories in which it resides. Following most theories of human cognition, the framework distinguishes between long-term memories, which change only gradually due to learning, and short-term memories, which change rapidly as the agent revises its beliefs and goals. In this section, we discuss ICARUS' memories and the formalisms used to encode their contents.[4] We will take our examples from the Blocks World (Langley & Choi, 2006), since many readers should find this domain familiar. We have described these aspects of the framework in more detail elsewhere, including their use in other domains like urban driving (Choi, Kaufman, Langley, Nejati, & Shapiro, 2004) and multi-column subtraction (Langley, Cummings, & Shapiro, 2004).

2.1. Long-term conceptual memory

One of ICARUS' long-term memories stores concepts that describe generalized situations in the environment. These may involve isolated objects, such as individual blocks, but they can also characterize physical relations among objects, such as the relative positions of blocks. Long-term conceptual memory contains the definitions of these logical categories. Each element specifies the concept's name and arguments, along with fields which describe perceptual entities that must be present, lower-level concepts that must match, lower-level concepts that must not match, and numeric relations that must be satisfied. Table 1 presents some concepts from the Blocks World. For example, the relation *on* describes a perceived situation in which two blocks have the same *x* position and the bottom of one has the same *y* position as the top of the other. The concept *clear* instead refers to a single block, but one that cannot hold the relation *on* to any other.

Definitions of this sort organize ICARUS categories into a conceptual hierarchy. Primitive concepts are defined entirely in terms of perceptual conditions and numeric tests, but many incorporate other concepts in their definitions. This imposes a lattice structure on the memory, with more basic concepts at the bottom and more complex concepts at higher levels. The resulting hierarchy is similar in spirit to discrimination network models of human memory like EPAM (Richman, Staszewski, & Simon, 1995), as well as to frameworks like description logics (Nardi & Brachman, 2002). Structurally, this lattice bears a close resemblance to the Rete networks (Forgy, 1982) used for matching in production-system architectures.

2.2. Long-term skill memory

ICARUS also incorporates a second long-term memory that stores knowledge about skills it can execute in the

---

Table 1

Some ICARUS concepts for the Blocks World, with variables indicated by question marks. Percepts refer only to attribute values used elsewhere in the concept definition.

```
(on (?blockl ?block2)
  :percepts    ((block ?blockl xpos ?xposl ypos ?yposl)
                (block ?block2 xpos ?xpos2 ypos ?ypos2 height ?height2))
  :tests       ((equal ?xposl ?xpos2) (>= ?yposl ?ypos2) (<= ?yposl (+ ?ypos2 ?height2))))
(clear (?block)
  :percepts    ((block ?block))
  :relations   ((not (on ?other ?block))))
(unstack-able (?block ?from)
  :percepts    ((block ?block) (block ?from))
  :relations   ((on ?block ?from) (clear ?block) (hand-empty)))
(pickup-able (?block ?from)
  :percepts    ((block ?block) (table ?from))
  :relations   ((ontable ?block ?from) (clear ?block) (hand-empty)))
```

environment, including their conditions for application and their expected effects. Each skill clause includes a head (a name and zero or more arguments) and a body that specifies the concepts that must hold to initiate the skill and one or more components. A primitive skill clause indicates one or more ordered, executable actions, along with those concepts that, taken together, describe the situation the skill produces when done. A primitive skill may also state conditions that must hold throughout its execution, which may require multiple cycles to complete. For example, Table 2 shows the skill *pickup*, which must satisfy the start condition, (*pickup-able ?block ?from*), defined in Table 1, and invokes **grasp*, which grasps a block, and **vertical-move*, which moves the hand in the vertical direction. The skill's only stated effect is to make (*holding ?block*) true.

In contrast, a nonprimitive skill clause specifies how to decompose that activity further. For instance, Table 3 includes two clauses for the nonprimitive skill *clear*. Each indicates that executing the clause will achieve that goal, but they differ in their start conditions and in their subskills. Nonprimitive skill clauses do not specify either required conditions or effects, but their heads always corresponds to a concepts that the skill will achieve upon successful completion. This representational assumption

figures centrally in the learning mechanism we describe later. Because ICARUS concepts and skills utilize a syntax similar to that found in the programming language Prolog, we have referred elsewhere to sets of these long-term memory structures as *teleoreactive logic programs* (Choi & Langley, 2005). This phrase conveys both their structural similarity to traditional logic programs and their ability to behave reactively in a goal-driven manner, following Nilsson's (1994) notion of a teleoreactive system.

2.3. Short-term memories

In addition to long-term memories, which encode relatively stable knowledge about a domain, ICARUS follows standard psychological theory by incorporating short-term stores that change more rapidly. These contain the agent's temporary perceptions and beliefs about the environment, as well as its goals and intended activities. They include:

- a *perceptual buffer* that holds descriptions of physical entities which correspond to the output of sensors; for the blocks world, this includes structures like (*block B xpos 10 ypos 2 width 2 height 2*), which specify the position and size of individual blocks.

Table 2

Primitive skills for the Blocks World. Each clause has a head that specifies the skill's name and arguments, a set of typed percepts, a single start condition, a set of effects, and a set of executable actions (marked by asterisks).

```
(unstack (?block ?from)                          (pickup (?block ?from)
  :percepts  ((block ?block) (block ?from))         :percepts  ((block ?block) (table ?from))
  :start     ((unstack-able ?block ?from))          :start     ((pickup-able ?block ?from))
  :effects   ((clear ?from) (holding ?block))       :effects   ((holding ?block))
  :actions   ((*grasp ?block)                        :actions   ((*grasp ?block)
              (*vertical-move ?block)))                          (*vertical-move ?block)))
(stack (?block ?to)                              (putdown (?block ?to)
  :percepts  ((block ?block) (block ?to))           :percepts  ((block ?block) (table ?to))
  :start     ((stack-able ?block ?to))              :start     ((putdown-able ?block ?to))
  :effects   ((on ?block ?to)                       :effects   ((ontable ?block ?to)
              (hand-empty))                                      (hand-empty))
  :actions   ((*horizontal-move ?block ?xpos)       :actions   ((*horizontal-move ?block)
              (*vertical-move ?block)                            (*vertical-move ?block)
              (*ungrasp ?block)))                                (*ungrasp ?block)))
```

Table 3
Some nonprimitive skills for the Blocks World that involve recursion. Each skill clause has a head that specifies the goal it achieves, a set of typed percepts, one or more start conditions, and a set of ordered subskills. Numbers after the head distinguish different clauses that achieve the same goal.

```
(clear (?B) 1                              (unstack-able (?B ?A) 3
 :percepts  ((block ?C) (block ?B))         :percepts  ((block ?A) (block ?B))
 :start     ((unstack-able ?C ?B))          :start     ((on ?B ?A) (hand-empty))
 :skills    ((unstack ?C ?B)))              :skills    ((clear ?B) (hand-empty)))
(hand-empty ( ) 2                          (clear (?A) 4
 :percepts ((block ?C) (table ?T))          :percepts  ((block ?A) (block ?B))
 :start    ((putdown-able ?C ?T))           :start     ((on ?B ?A) (hand-empty))
 :skills   ((putdown ?C ?T)))               :skills    ((unstack-able ?B ?A)
                                                        (unstack ?B ?A)))
```

- a *short-term conceptual memory* that contains beliefs about the environment which the agent infers from items present in its perceptual buffer and long-term concept memory; for instance, this might contain the literal (*on B C*), which is an instance of the *on* concept in Table 1.
- a *short-term skill memory* that contains the agent's goals and associated skill instances it intends to execute; each goal literal specifies a concept's name and argumets, as in (*clear A*), whereas each associated intention gives a skill's name and its arguments, as in (*stack B C*), which is an instance of the skill *stack* in Table 2.

Unlike most cognitive architectures, every element in the short-term conceptual and skill memories must be an instance of some generalized structure in the long-term conceptual and skill memory, respectively; they cannot be arbitrary symbolic structures. We have discussed this strong correspondence assumption at more length elsewhere (Langley & Rogers, 2005).

3. Conceptual inference and skill execution

Like most cognitive architectures, ICARUS operates in distinct cycles. On each such iteration, the system updates its perceptual buffer by sensing objects in its field of view, with the specific sensors depending on the particular environment in which the agent is operating. This process produces perceptual elements, which are are deposited in the perceptual buffer and which initiate matching against long-term concepts. The matcher checks to see which primitive concepts (i.e., those defined entirely in terms of percepts) are satisfied, adds each matched instance to conceptual short-term memory, and repeats the process on nonprimitive concepts to infer higher-level beliefs.

This process aims to infer all instances of concepts that are implied by ICARUS' conceptual definitions and the contents of the perceptual buffer, but there are different ways to achieve this effect. Earlier versions recomputed all inferences from scratch on each cycle in an exhaustive manner,[5] but the current system uses a more efficient mechanism. ICARUS stores inferred beliefs with the general concept of

which they are instances and retains these beliefs across cycles until their perceptual support disappears, when it removes them. The Rete networks (Forgy, 1982) used in many production-system architectures rely on a very similar process; the main difference is that each node in the ICARUS hierarchy corresponds to a defined concept. This update mechanism, which also shares features with truth-maintenance systems (e.g. Doyle, 1979), greatly reduces unnecessary calculations across cycles.

On each cycle, the architecture also examines the agent's goals and their associated intentions in short-term skill memory to determine which, if any, apply to the current situation.[6] For each intended skill instance, ICARUS accesses all clauses of the general skill to see if they are applicable. Since variables can be bound within a skill's body, this set may include multiple variants of each skill clause stored in long-term memory. A primitive skill clause is applicable if, for its current variable bindings, its effects do not yet hold, its requirements are satisfied, and, if the system has not yet started executing it, the start conditions match the current situation. A higher-level skill clause is applicable if its head is not satisfied, the start conditions are satisfied if it has not been initiated, and at least one subskill is applicable. Because this latter test is recursive, a skill is applicable only when ICARUS can find at least one acceptable path downward to executable actions, which the architecture returns for invocation.

For example, suppose an ICARUS agent has the goal (*clear A*) in a situation where block A is on the table, block B is on A, block C is on B, and the hand is empty. Suppose further that the agent has access to the primitive skills in Table 2 and the nonprimitive ones in Table 3. In this case, the system would find an applicable path through the skill hierarchy that is relevant to its goal, as shown in Fig. 1: [(*clear A*), (*unstack-able B A*), (*clear B*), (*unstack C B*)]. This holds because the instantiated start conditions of each skill along the path (e.g., (*on B A*) and (*hand-empty*) for the topmost skill) are present in conceptual short-term memory. If selected, (*unstack C B*) would alter the environment, making the path [(*clear A*), (*unstack-able B A*), (*hand-empty*), (*putdown C T*)] acceptable on the next cycle. This would

---

[5] Asgharbeygi, Nejati, Langley, and Arai (2005) report an alternative mechanism that uses expected values to makes inferences more selectively.

[6] ICARUS' first step in a run typically involves selecting a relevant and applicable instance of a nonprimitive skill that it believes will achieve one of its goals.
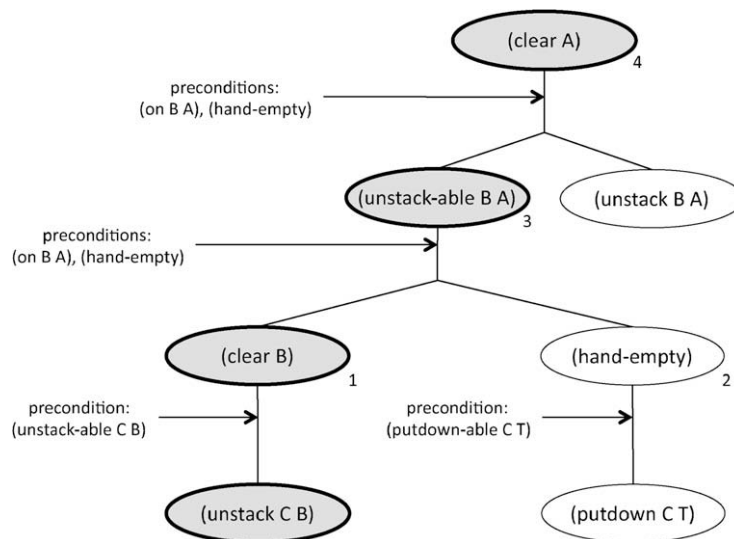
Fig. 1. An executable path through ICARUS' skill hierarchy for the Blocks World, when the top-level goal is (*clear A*). Ellipses denote the goals of the skills at different levels, while the numbers associated with them show the skill labels from Table 3.

produce a belief state that enables the next step in the procedure, which would continue until the agent had satisfied its top-level goal, (*clear A*).

During skill selection, ICARUS incorporates two preferences that provide a balance between reactivity and persistence. When confronted with a choice between two or more subskills, it selects the first alternative for which the head is not satisfied. This supports reactive control, since the system reconsiders previously completed subskills and, if their effects no longer hold for some reason, reexecutes them to remedy the problem. On the other hand, when encountering two or more applicable skill paths, ICARUS selects the one that shares the most elements from the start of the path executed on the previous cycle. This encourages the system to continue executing a high-level skill it has already started until that skill achieves its associated goal or until it becomes inapplicable.

4. Means-ends problem solving

As just explained, ICARUS can execute complex hierarchical skills in a reactive manner, but our initial studies (e.g. Choi et al., 2004; Langley et al., 2004) assumed that these skills are already present in long-term memory. Although much human behavior appears to involve the application of such routine skills, people can also solve novel tasks that require the dynamic combination of existing knowledge elements through some form of heuristic problem solving.

To model this capability in ICARUS, we have introduced a variant of means-ends analysis (Newell, Shaw, & Simon, 1960) that operates over the architecture's knowledge structures, including both long-term concepts and skills provided by the programmer and short-term beliefs and goals produced by the architecture. Traditional means-ends problem solving selects some unsatisfied aspect of the goal state to achieve, then selects an operator that would achieve it. If that operator's preconditions match the current state, it is applied; otherwise, the method selects an unsatisfied precondition to achieve, selects an operator that would achieve it, and so on. Once a condition is met, the process is repeated until the original goal description is satisfied. This may require search, which is often pursued in a depth-first manner. Means-ends analysis has been implicated repeatedly in human problem solving on novel tasks.

To support this mechanism, the current version of ICARUS augments the short-term skill memory with a goal stack. Each element in this stack specifies a goal (a desired concept instance), whether the agent intends to achieve it by backward chaining off a concept definition or a skill clause, and, in the latter case, the skill instance that, if executed, should achieve it. Each goal element also specifies subgoals that have already been achieved, along with skill and/or concept instances that it has tried in reaching this goal but that have failed. The first are needed to keep the system from considering skills that would undo its previous accomplishments, whereas the second ensures it does not repeat earlier mistakes. We also assume that both the start conditions of primitive skills and top-level goals must be cast as single relational literals, which causes no loss in generality, since either may be defined concepts.

The ICARUS interpreter takes advantage of these new memory structures. On each cycle, the system takes the first true alternative of the following five statements:

- If the current goal G of the goal stack GS is satisfied, then pop G from GS and store information about the success with G's parent.
- If the goal stack GS does not exceed the depth limit and there are applicable skill paths that start from a skill instance with the current goal G as its head, then select one such path and execute it.

- If there is a nonempty set of primitive skill instances in which the current goal G is an effect that have not already failed, then select a skill instance from this set and push its start condition (which we assume subsumes any required conditions) onto the goal stack GS.
- If the current goal G is an instance of a complex concept with unsatisfied subconcepts H and with satisfied subconcepts F, then if there is a subconcept I in H that has not yet failed, push I onto the goal stack GS.
- Otherwise pop the current goal G from the goal stack GS and store information about the failure with G's parent.

We assume that each of these activities takes a single cycle of the architecture, with the initial situation being a special case of the third item that triggers the process. Because reasoning about how to achieve an objective can require many manipulations of the goal stack, it takes more cycles than executing a stored hierarchical skill for that goal, even when the agent finds a solution on its first attempt and does not have to backtrack.

Fig. 2 shows a successful trace of the problem solver's behavior on a simple urban driving task when the goal is to turn right and stop in the street segment on the right-hand side, thus making (*in-segment me g998*) true. Initially, the agent's car is stopped in the leftmost lane, *g1020*, in a

street segment, *g994*. In this situation, the system looks for skills with the goal as its head but, when this fails, it considers skills that have the goal as one of their effects. In this case, invoking the primitive skill instance (*steer-for-right-turn me g978 g998*) would produce the intended result, but it cannot be applied because its instantiated start condition, (*in-intersection-for-right-turn me g978*), does not hold. In response, the problem solver stores the skill instance with the initial goal and pushes the subgoal onto the goal stack. Then the system looks for skills that will achieve the subgoal and finds the skill instance, (*in-intersection-for-right-turn me g978*). However, it also has an unsatisfied start condition, (*in-rightmost-lane me g1021*), and in turn the problem solver pushes the new subgoal onto the goal stack.

Next, the problem solver attempts to retrieve skills that would achieve (*in-rightmost-lane me g1021*). However, because it has no such skills in memory, it resorts to chaining off the definition of *in-rightmost-lane*. This involves two instantiated subconcepts – (*driving-in-segment me g994 g1021*) and (*last-lane g1021*) – but only the first of these is unsatisfied, so the system pushes it onto the goal stack. This in turn leads it to consider skills that would produce this literal as an effect, but when it finds that there are no such skills, it chains off the definition of *driving-in-segment* again. Among the unsatisfied subconcepts, (*centered-in-*
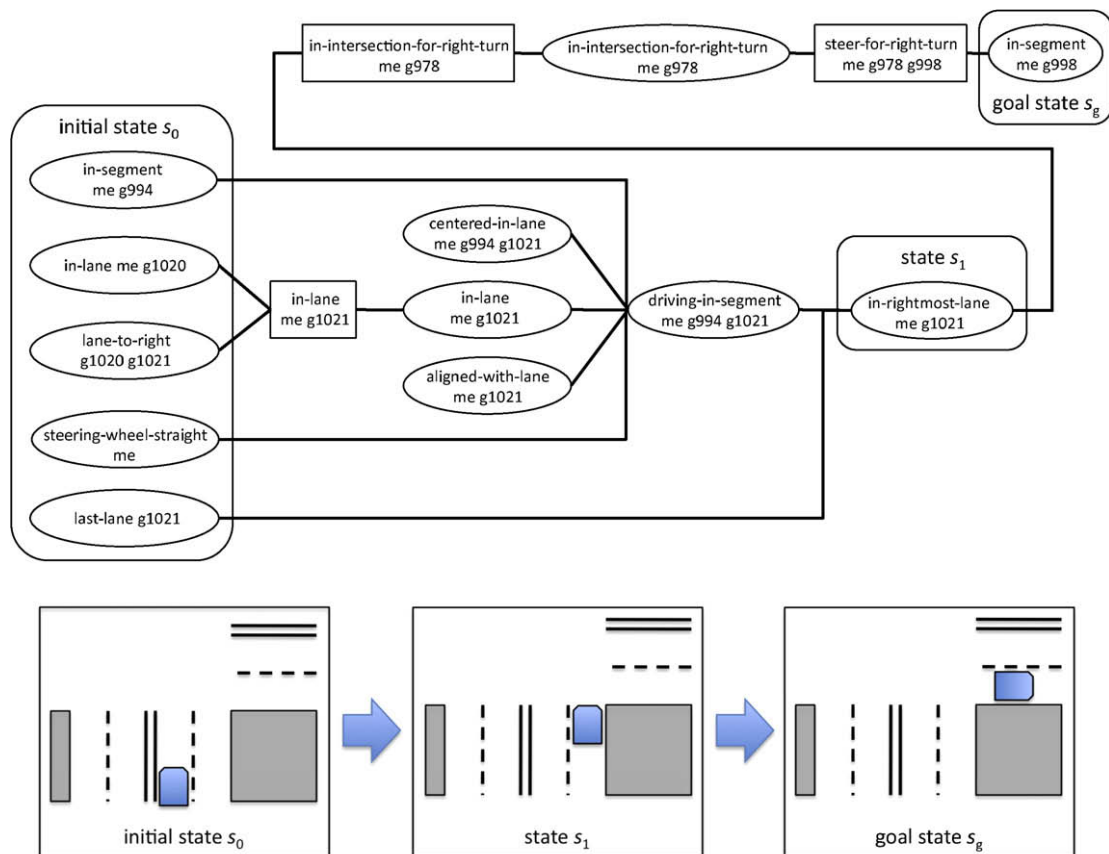

Fig. 2. A trace of successful problem solving in the urban driving domain, with ellipses indicating (sub)goals and rectangles denoting primitive skills.

lane me g994 g1021), (*in-lane me g1021*), and (*aligned-with-lane me g1021*), the problem solver chooses the second one, possibly after some search. It then retrieves the corresponding skill instance (*in-lane me g1021*), whose start conditions, (*in-lane me g1020*) and (*lane-to-right g1020 g1021*) already hold, and executes the selected skill. The associated actions alter the environment over several cycles and cause the agent to infer (*in-lane me g1021*) from its percepts. In response, the system pops this goal from the stack and reconsiders its parent, (*driving-in-segment me g994 g1021*).

However, this has not yet been achieved because two other subconcepts, (*centered-in-lane me g994 g1021*) and (*aligned-with-lane me g1021*), are still not satisfied in the environment. Moreover, executing the previous skill has caused another of its component concept instances, (*steering-wheel-straight me*), to become false. Thus, the system pushes one of these onto the stack and continues working at this level in a similar fashion until all the instantiated subconcepts are satisfied. Once all the components, (*in-segment me g994*), (*centered-in-lane me g994 g1021*), (*in-lane me g1021*), (*aligned-with-lane me g1021*), and (*steering-wheel-straight me*) are achieved, their parent (*driving-in-segment me g994 g1021*) becomes true in the environment, and the system pops it from the goal stack.

This step also achieves the subgoal (*in-rightmost-lane me g1021*), thanks to the other component, (*last-lane g1021*) being true from the beginning of execution. Thus, the problem solver pops this element from the goal stack and executes the skill instance it had originally selected, (*in-intersection-for-right-turn me g978*), in the new situation. Upon completion, the system perceives that the altered environment satisfies the start condition for the skill, (*steer-for-right-turn me g978 g998*), which it selected for its goal, (*in-segment me g998*). After executing this skill, it then detects that it has solved the problem by achieving the top-level goal and halts.

For the sake of clarity, we have presented the trace of successful problem solving, but finding such a solution may involve search. When backward chaining off skills that would achieve the objective of the current stack entry, ICARUS considers only skill instances that have not yet failed. The system also prefers candidates that have the fewest expanded start conditions that are unmet by the current environmental state, with fully matched conditions being most desirable. If candidates tie on this criterion, it selects an alternative at random. When backward chaining off the unmatched elements of a concept definition, the system selects subgoals at random after eliminating those which have failed in the past.

Taken together, these biases produce a heuristic version of means-ends analysis. However, this problem-solving method is tightly integrated with the execution process. ICARUS backward chains off concept or skill definitions when necessary, but it executes the skill associated with the top stack entry as soon as it becomes applicable. Moreover, because the architecture can chain over hierarchical reactive skills, their execution may continue for many

cycles before problem solving is resumed. In contrast, most models of human problem solving and most AI planning systems focus on the generation or the execution of plans, rather than interleaving the two processes.

Of course, executing a component skill before constructing a complete plan can lead an agent into difficulties, since one cannot always backtrack in the physical world. This strategy may well lead to suboptimal behaviors, but human intelligence is more about satisficing than optimizing, and interleaving problem solving with execution requires far less memory than constructing a full plan before executing it. However, it can produce situations from which the agent cannot recover without starting the problem over.

In such cases, ICARUS stores the goal element for which its executed skill caused a problem, along with everything below it in the stack. The system begins the problem again, this time avoiding the skill and selecting another option. If it makes a different execution error this time, it again stores the problematic skill and its context, then starts over once more. ICARUS also starts over if it has not achieved the top-level objective within a specified number cycles. Such repeated attempts at solving a task, with selected memory about previous passes, seems a better model of human problem solving than systems that construct a complete plan before execution. Jones and Langley's (2005) model of means-ends problem solving, EUREKA, used a similar restart strategy, but it kept no explicit record of previous failed paths.

5. Learning hierarchical skills from problem solving

In the previous pages, we described two facets of ICARUS: its execution of hierarchical skills on familiar tasks and its use of problem solving to handle novel ones. The first lets the system operate efficiently, but skills are tedious to construct manually, whereas the second gives the system flexibility but requires reasoning and means-ends search. We believe that humans also have both capabilities, but that they use learning to transform the results of successful problem solving into hierarchical skills. We would like to incorporate a similar capability into ICARUS.

However, we want our learning mechanisms to reflect certain properties that appear to hold for human skill acquisition. One is that learning should take advantage of existing knowledge, such as the definitions of current skills and concepts. In addition, acquisition should be incremental, in that it learns from each new experience, and interleaved with the problem-solving process. The recent literature on computational learning contains few cases of such knowledge acquisition, although in Section 7 we discuss some older work that has this character.

Our extension of ICARUS achieves this effect through a form of impasse-driven learning that is tied closely to its problem-solving and execution processes. For this reason, the learning mechanisms require no additional inputs beyond those required for these basic performance processes. As in Soar (Laird et al., 1986), the purpose of skill

Table 4
Four skill clauses learned for urban driving, along with a primitive skill that is used in a learned skill recursively.

```
; skill clause for situations that require a lane change
(driving-in-segment (?me ?g994 ?gl02l) 20
  :percepts  ((segment ?g994) (lane-line ?gl02l) (self ?me))
  :start     ((in-segment ?me ?g994) (steering-wheel-straight ?me))
  :ordered   ((in-lane ?me ?gl02l) (centered-in-lane ?me ?g994 ?gl02l)
             (aligned-with-lane-in-segment ?me ?g994 ?gl02l)
             (steering-wheel-straight ?me)))
; skill clause that prepares the agent for a right turn
(in-rightmost-lane (?me ?gl02l) 21
  :percepts  ((self ?me) (lane-line ?gl02l))
  :start     ((last-lane ?gl02l))
  :ordered   ((driving-in-segment ?me ?g994 ?gl02l)))
; recursive skill clause that takes first step needed for right turn
(in-intersection-for-right-turn (?me ?g978) 22
  :percepts  ((lane-line ?gl02l) (self ?me) (intersection ?g978))
  :start     ((last-lane ?gl02l))
  :ordered   ((in-rightmost-lane ?me ?gl02l)
             (in-intersection-for-right-turn ?me ?g978)))
; skill clause for handling right turns
(in-segment (?me ?g998) 23
  :percepts  ((self ?me) (intersection ?g978) (segment ?g998))
  :start     ((last-lane ?gl02l))
  :ordered   ((in-intersection-for-right-turn ?me ?g978)
             (steer-for-right-turn ?me ?g978 ?g998)))
; primitive skill that is called recursively by learned clause
(in-intersection-for-right-turn (?self ?int) 1
  :percepts  ((self ?self) (segment ?sg) (intersection ?int)
             (lane-line ?lane segment ?sg))
  :start     ((in-rightmost-lane ?self ?lane))
  :requires  ((in-segment ?self ?sg) (intersection-ahead ?int)
             (last-lane ?lane))
  :actions   ((*cruise))
  :effects   ((in-intersection-for-right-turn ?self ?int)))
```

learning is to avoid such impasses in the future. Thus, whenever the architecture achieves an objective that is associated with an entry in the goal stack, this success provides an opportunity for learning. The system acquires two distinct forms of skill that are tied to different aspects of problem solving.

The first class of skills result from situations in which the problem solver cannot find a skill to achieve a goal *G*, and thus pursues subgoals based on the unsatisfied conditions of *G*'s conceptual definition. If the agent achieves these subgoals in the order $\{G_1, G_2, \ldots, G_n\}$, thus satisfying the parent goal *G*, ICARUS constructs a new skill clause that has *G* as its head and that has $\{G_1, G_2, \ldots, G_n\}$ as its ordered subskills.[7] The start conditions of the new clause are simply those subconcepts of *G* that were satisfied when it was pushed onto the goal stack. The head, conditions, and subskills have their arguments replaced by variables in a consistent manner, ensuring applicability to analogous situations that involve different objects.

For example, upon achieving the subgoal (*driving-in-segment me g994 g1021*) in Fig. 2, the system constructs the

driving-in-segment skill clause labeled 20 in Table 4. The head (*driving-in-segment ?me ?g994 ?g1021*) is a generalized version of the goal (*driving-in-segment me g994 g1021*), whereas the ordered subskills (*in-lane ?me ?g1021*), (*centered-in-lane ?me ?g994 ?g1021*), (*aligned-with-lane ?me ?g1021*), and (*steering-wheel-straight ?me*) are generalized versions of its four subgoals. The start conditions are (*in-segment ?me ?g994*) and (*steering-wheel-straight ?me*), which are generalized versions of the subconcepts that held when the goal was established. Finally, the :percepts field specifies the types for objects that serve as the head's arguments. This mechanism constructs different variants of a skill, with separate start conditions and distinct subgoals, from subproblems that involve different initial conditions.

The second category of skills results from situations in which ICARUS has selected a primitive skill instance *S*2 in order to achieve a goal *G*, but found its single start condition *G*2 unsatisfied and selected another skill instance, *S*1, to achieve it. Once the agent has executed both skills successfully and it has reached the goal, the system constructs a new skill clause that has *G* as its head and that has *G*2 (rather than the specific clause *S*1) and *S*2 as ordered subskills. The start conditions are simply the start conditions of the *S*1 clause used in the subproblem solution, which are

---

[7] Note that the system refers to subskills by the goals they achieve, rather than to specific clauses, which lets the parent skill take advantage of other clauses for these goals that are learned later.

sufficient because the problem solver *S*1 selected it to achieves the start condition of *S*2, which in turn achieves the goal *G*. Again, specific arguments are replaced consistently by variables.

For instance, upon achieving the top-level goal (*in-segment me g998*) in Fig. 2, ICARUS creates the *in-segment* skill clause labeled 23 in Table 4. This incorporates a generalized version of (*in-segment me g998*) as its head, along with variablized versions of (*in-intersection-for-right-turn me g978*) and (*steer-for-right-turn me g978 g998*) as its two ordered subskills. The start condition, (*last-lane ?g1021*), is the same as that for *in-right-most-lane* clause 21, since the latter was created to achieve the start condition of *in-intersection-for-right-turn*, which in turn satisfies the start condition of *steer-for-right-turn* that is used to achieve the top-level goal.

Both learning mechanisms are fully incremental, in that each learning event draws on a single problem-solving experience and thus requires no memory of previous ones. They support within-trial learning, since skills acquired on one subproblem may be used to handle later subproblems. The processes also build on existing knowledge, since the construction of new skill clauses involves the composition of those used in a training problem's solution. Taken together, these support a form of cumulative learning, in which ICARUS learns skills on one problem, uses them to solve a later problem, and incorporates them into still higher-level structures.

As suggested by our examples, these learning methods can acquire both disjunctive and recursive skills. The key to this ability lies in the assumption that acquired skill clauses which achieve the same goal should be given the same head. By indexing skills in this manner, ICARUS knows when two or more clauses should be stored together, which leads in turn to the creation of skills that call on themselves, either directly or through intermediate skills. This makes the architecture's learned skills considerably more flexible and general than traditional 'macro-operators' (e.g., Iba, 1989) or composed production rules (e.g., Neves & Anderson, 1981).

Of course, the creation of disjunctive and recursive structures has potential for overgeneralization, as demonstrated by research on the induction of context-free grammars (e.g., Langley & Stromsten, 2000). Our technique for determining the start conditions on new skill clauses is much simpler than standard techniques for analytical learning or rule induction. In fact, at first glance, the learned clauses in Tables 3 and 4 appear highly overgeneral, but this ignores the fact that ICARUS does not interpret skills in isolation. Recall that the architecture must find an entire path through the skill hierarchy before it can execute the primitive skill at its terminus. This means the system collects conditions dynamically, as it descends the hierarchy, guarding against overgeneralization by carrying out limited analysis at performance time rather than doing it all at learning time.

Unlike some approaches to incremental learning, ICARUS' methods require no additional mechanisms for skill refinement. Each skill clause is generalized when the architecture constructs it, and its start conditions are assumed to be accurate. The knowledge it acquires from solving a given problem may well be incomplete, but this will simply lead to further impasses that produce additional learning. Skill clauses acquired later complement, but do not compete with, those learned earlier because they cover different situations or the older clauses would have avoided the impasse. Thus, learning is purely monotonic, as in frameworks like Soar.

We should note that our current implementation restricts the use of learned skills in future problem solving. In particular, we have adopted Mooney's (1989) idea that one should not chain off the preconditions of learned skills. This does not restrict their use by the execution module, but it does mean that the problem solver considers a learned skill only when its start conditions are already satisfied. As a result, clauses acquired from chaining off skills always have a left-branching structure in which the second subskill is primitive. This assumption may seem restrictive, but, like Mooney, we believe it provides an effective guard against the utility problem (Minton, 1990), in which the creation and use of complex structures reduces search but actually slows performance.

6. Experimental studies of skill learning

Although these methods for learning hierarchical skills seem plausible, whether they improve an ICARUS agent's performance is an empirical question. In this section, we report the results of basic tests of these mechanisms on three distinct domains: urban driving, the Blocks World, and FreeCell solitaire. After this, we report more systematic experiments with the domains that examine the effects of learning in more detail. As one measure of performance, we used the number of recognize-act cycles required to solve the problem in the simulated environment, including both problem solving and execution steps. However, we also measured the CPU time required to solve each problem, to determine whether ICARUS suffers from the utility problem.

6.1. Domains and basic demonstrations

To ensure that our approach to learning hierarchical skills operated as intended, we developed ICARUS programs for the three domains. In each case, we provided a set of primitive skills sufficient for solving problems with means-ends analysis and a set of hierarchical concepts sufficient for recognizing situations that were relevant to executing those skills. For example, we devised some 41 concepts and 19 skills for the urban driving domain, 11 concepts and four skills for the Blocks World, and 24 concepts and 12 skills for FreeCell solitaire. The Appendices A, B, and C give the names of the primitive concepts, nonprimitive concepts, and primitive skills provided for each domain, which should also suggest their functions.

In addition, we also provided the architecture with a set of sensors and effectors for each simulated environment.

We have already discussed the Blocks World, but both urban driving and FreeCell merit some explanation. The first domain involves a dynamic simulation of a downtown driving environment. The city contains objects represented as rectangles of different sizes, including buildings and sidewalks organized into square blocks that are divided by street segments and intersections. Each segment includes a yellow center line and white dotted lane lines, and it has a marked street name and speed limit. Each building has a unique street address to help the agent navigate through the city and to support tasks like package delivery. The city configuration used in our experiments has nine blocks with four vertical streets and four horizontal streets. The ICARUS agent must operate under physical laws and follow the rules of driving, such as staying on the right side of the street and turning from the proper lane. We provide the agent specific with goals to achieve, such as getting onto another street segment or delivering a package to a certain address.

FreeCell solitaire involves eight stacks of cards, the first four of which contain seven cards and the last four of which contain six cards. All 52 cards are dealt face up, making them visible to the player. In addition, there are four free cells, which can serve as temporary holding spots for one card each during the game, and one foundation cell for each suit. The goal in FreeCell is to get all cards on the foundation cells in ascending order (where the ace is one and the king is thirteen) grouped by suit. Once on its foundation cell, a card cannot be removed. Only fully-exposed cards at the top of each stack and card in the free cells are in play. The agent can move one card at a time to an available free cell, to the appropriate foundation cell, to an empty stack, or to a stack in which the top card has a different color and value one higher than the moved card.

Sample runs with the urban driving domain, the Blocks World, and a reduced version of FreeCell indicated that the extended version of ICARUS was able to solve problems in their respective domains with some search and, from their respective traces, learn hierarchical skills in the manner described earlier. We found that, when given the same task to solve a second time, the system utilized this knowledge to handle it without problem solving. Moreover, because the system generalizes its learned structures beyond the specific instances on which they are based, they transfer fully to any tasks that are isomorphic to those it has already solved. The only constraint is that this isomorphism must involve the same goal and have the same concepts satisfied or unsatisfied in the initial environment.

However, we should note this ability does not mean that the system can complete a familiar problem in a single cycle. Recall that, unlike traditional work on cognitive architectures, ICARUS resorts to problem solving only to enable action, and it must still execute its acquired skills to achieve a goal. Thus, for a problem that requires four primitive steps, the system takes six cycles on the second

encounter, with one to retrieve the hierarchical skill and one to realize it has finished. However, the agent requires neither search or backward chaining over skills or concepts to complete any problem it has solved previously.

6.2. Experiment with urban driving

Although some sample runs were encouraging, we desired more than anecdotal demonstrations that the new mechanisms supported incremental learning of hierarchical skills. We also wanted evidence from systematic experiments that this learned knowledge produces more effective behavior. Our first study along these lines focused on urban driving, which is the most dynamic of the three settings and thus the one most appropriate for evaluating our methods for learning skills that support reactive execution.

As noted above, we provided ICARUS with 41 concepts and 19 primitive skills relevant to this environment. With the basic knowledge, the agent can characterize its situation at multiple levels of abstraction and perform actions for accelerating, decelerating, and steering left or right at realistic angles. Thus, it can operate a vehicle, but this is not sufficient to drive safely in a city environment. The agent must still learn skills for staying aligned and centered within lane lines, change lanes, increase or decrease speed for turns, and stop for parking.

To encourage such learning, we presented the agent with the goal of driving on a different street segment than its current one. To achieve this objective, it resorted to problem solving, which found a solution path that involved changing to the rightmost lane, staying aligned and centered until the intersection, steering right into the target segment, turning the corner, and finally aligning and centering in the new lane. We let the ICARUS agent practice this task for five trials to examine its improvement with experience. We repeated this procedure ten different times with slightly different starting positions, collected performance measures for each run, and averaged the results.

Fig. 3 shows the cumulative number of cycles used for problem solving as a function of the number of trials to achieve the goal. As the agent accumulates knowledge about this task, problem solving disappears almost entirely, which results in the flat curves for the learning case. However, it is notable that the total number of cycles required to complete the task, which is not shown here, reveals a reduced effect. This is due to the fact that this problem is dominated by execution time, as the agent must actually drive the vehicle to its destination in the city. Thus, the benefit of learning is diminished when one takes the total number of cycles into account.

Table 4 shows the four skill clauses acquired during one of these runs. The clause for *in-segment* specifies a decomposition the system found for achieving this top-level goal. It refers to the learned subskill for *in-intersection-for-right-turn* and the primitive skill *steer-for-right-turn*. The former refers to *in-rightmost-lane*, which invokes another learned skill for *driving-in-segment*, but it also calls on itself recur-
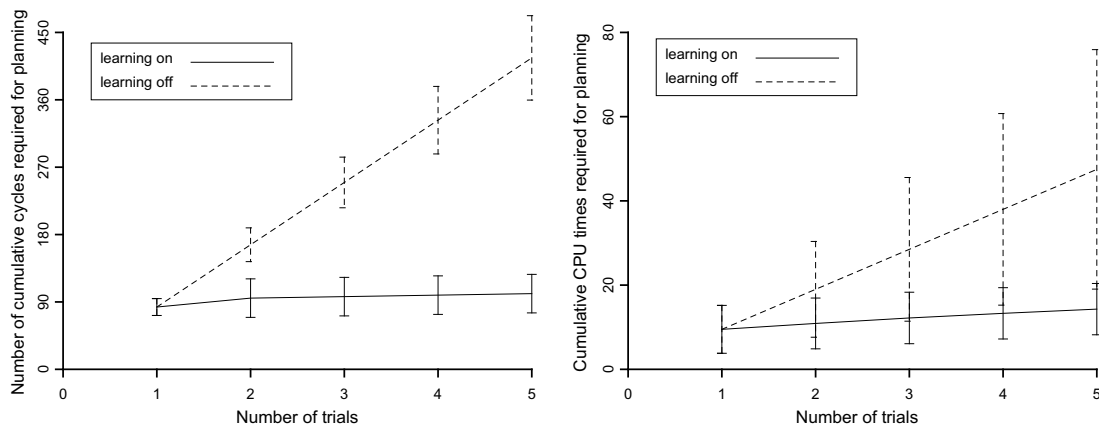
Fig. 3. Cumulative number of cycles and CPU times required to plan for a particular right-turn task in the driving domain, as a function of the number of trials. Each learning curve shows the mean over ten sets of trials and 95% confidence intervals.

sively, at which point a primitive skill clause with the same head is used as a base case. For clarification, the table also presents this primitive clause for *in-intersection-for-right-turn*, which the system was given as background knowledge.

Fig. 4 shows a trace of the agent's behavior on the task during learning, in a situation that involves a street with two lanes, and afterwards, in a setting that instead involves three lanes. The trace of the vehicle's movement demonstrates that the learned skills generalize to cases that involve more lanes than were present during training. This ability follows directly from the recursive structure of the learned *in-intersection-for-right-turn* clause. Behavior after learning is also smoother, presumably because the agent need not engage in problem solving when it overshoots slightly after getting into the target lane in preparation for the right turn.

Fig. 4. A trace of the ICARUS driving agent's behavior, during and after learning, on a task that required changing to the rightmost lane and turning at the intersection. The trace demonstrates generalization to a new setting with a different number of lanes.

6.3. Experiment with the Blocks World

Although the Blocks World is far less dynamic than urban driving, it lends itself to scaling studies that involve generalization to tasks with varying numbers of objects. For this domain, we provided ICARUS with the four primitive skills in Table 2 and 11 concepts that were sufficient, in principle, to solve any problem. We then presented the agent with the problems in sequence, using each task as a training problem but also recording the number of cycles and CPU time required to complete it. Because misguided search combined with execution can lead the problem solver into undesirable physical states, we told it to halt if it had not finished a run within 100 cycles and to start over from the initial state. However, the agent could attempt a given problem only five times, and thus spend at most 500 cycles before giving up entirely. We also limited the stack depth to ten goal elements. We enforced these constraints for reasons of practicality and because we think they reflect the manner in which humans tackle novel problems.

We generated randomly a set of Blocks World tasks that involved settings with 5, 10, 15, 20, 25, and 30 blocks. Each complexity class had 67–69 distinct problems, and we ordered the classes by difficulty (five-block tasks first and 30-block tasks last). The intuition was that the system would learn more effectively if we presented it first with simpler problems, which it could then use in solving more difficult ones. To this end, ICARUS retained skills acquired on successful runs for use in later tasks. We provided the system some 400 randomly generated problem orders and recorded the number of cycles and CPU times needed for each task. As a control, we also ran the system with its learning mechanisms off for another 400 problem sets that were ordered randomly within difficulty classes. Because the problems require different amounts of effort, traditional learning curves are not very informative. Instead, following Minton (1990), we report *cumulative* cycles and CPU times as a function of the number of training problems.
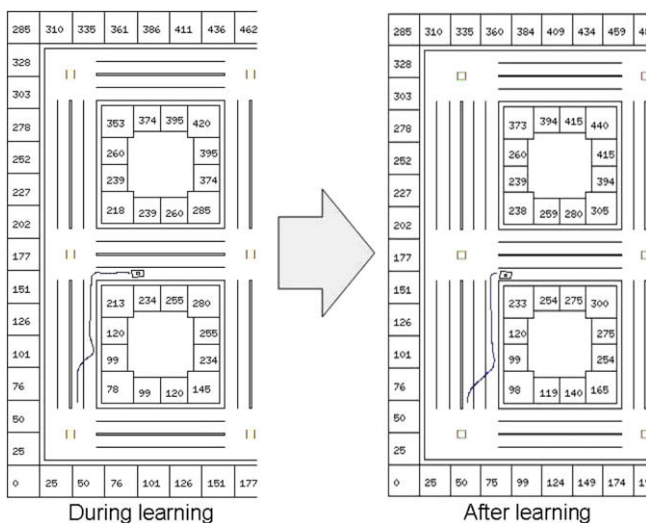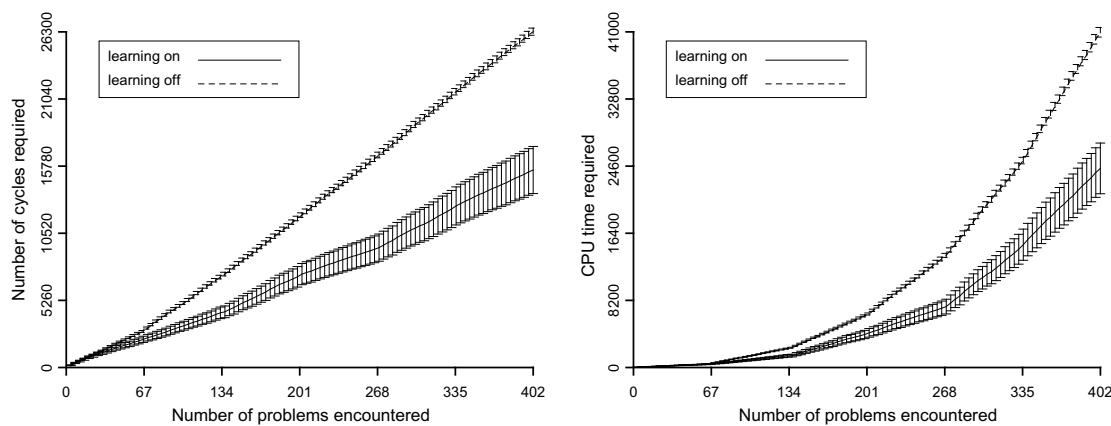
Fig. 5. Cumulative number of cycles and CPU times required by Icarus to solve a Blocks World task as a function of the number of problems encountered, averaged over 400 runs and with problems ordered by difficulty, with the goal stack of size ten. Tick marks on the horizontal axis indicate shifts in problem complexity.

Fig. 5 shows the resulting curves, including 95 percent confidence intervals around each mean. As expected, the curves mainly take the form of superlinear functions whose slopes increase with problem difficulty. Although the large scale of plots made the learning and non-learning curves look similar for early parts of the curves, there was some benefit for learning even from the beginning, but the difference grows substantially as the systems encounter harder problems. Clearly, prior experience reduces search substantially when it reaches problems with many blocks, and there is no evidence that learning produces a utility problem. Remember that we have made the transfer of learned knowledge challenging in that none of the problems are isomorphic, although they may involve isomorphic subtasks. The results indicate that Icarus can take advantage of this similar substructure to reduce its effort on later problems.

6.4. Experiment with FreeCell solitaire

To ensure that our conclusions held for more than the Blocks World, we carried out a similar experiment with FreeCell solitaire, which we described earlier in this section. We gave the Icarus agent only the 12 basic skills needed to move cards and the top-level goal of getting all cards into foundation cells, along with 24 concepts for describing situations. Unlike the Blocks World, this domain has only one goal condition, but it still has many possible starting states.

For this study, we randomly generated 20 problems each that involved 8, 12, 16, 20, and 24 cards.[8] We ran the system on 300 different sequences of tasks, with simpler problems being presented earlier but ordered randomly within each of the five difficulty classes. As before, we expected

that the agent would learn skills from the easier problems that would assist on the harder ones, thus reducing problem-solving effort. For comparison, we presented another 300 random sequences to a non-learning system with the same initial skills and concepts.

Fig. 6 presents the cumulative results for this experiment, with error bars that indicate the 95% confidence intervals. As in the Blocks World, the difference between the learning and non-learning conditions is substantial. However, problems with 20 cards or more require a different class of skills that involve column-to-column moves, which caused the lessened gap between the two conditions around the 80th problem. However, once they have been acquired, these skills provide some advantage, as evidenced by the downturn in the curve for the learning system on the far right of the graphs. Again, we detected no sign of a utility problem as the agent accumulates knowledge in this domain.

7. Related research

Research on learning cognitive skills from problem solving has a long history within both AI and cognitive science. For example, work on explanation-based learning often aimed to improve efficiency on problem-solving tasks and combined experience with a domain theory to create new cognitive structures. Some techniques focused on the acquisition of search-control rules to guide problem solving, but other efforts dealt instead with the construction of macro-operators from primitive operators (e.g., Iba, 1989; Mooney, 1989; Shavlik, 1989). Our approach to skill learning comes closer to the second paradigm, since both involve composing knowledge elements into larger structures. However, Icarus adapts this idea to the creation of disjunctive and even recursive skill hierarchies, whereas traditional methods emphasized the creation of 'fixed-sequence' macro-operators that were far less flexible.

Icarus also bears some similarity to other cognitive architectures that incorporate varieties of analytical or

---

[8] Icarus' problem solver has difficulty with FreeCell tasks that involve 30 or more cards, apparently because they involve goal interactions that basic means-ends analysis cannot handle.
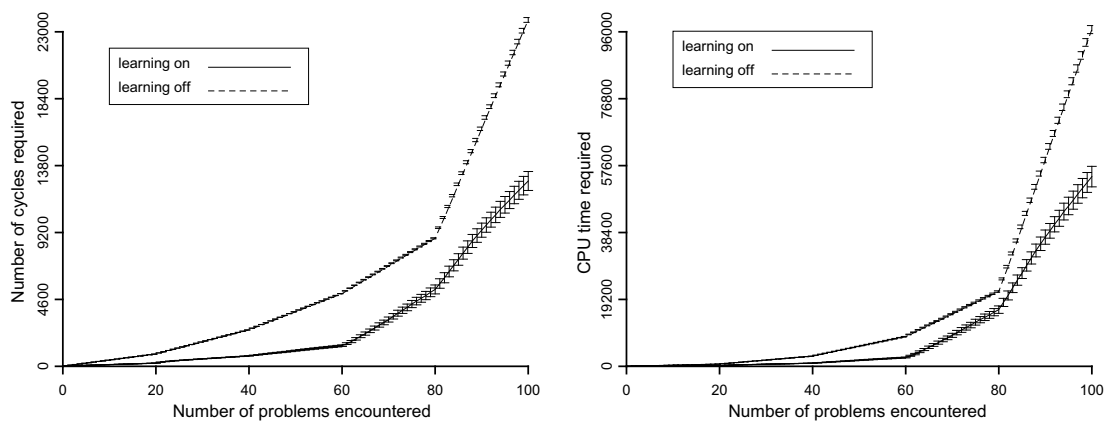
Fig. 6. Cumulative number of cycles and CPU times required by ICARUS to solve a FreeCell task as a function of the number of problems encountered, averaged over 300 runs and with problems ordered by difficulty.

explanation-based learning. For example, Laird, Rosenbloom, and Newell's (1986) Soar revolves around a problem solver that proceeds until the system encounters an impasse, in which case it creates a subgoal to resolve it. This resolution may require search and take some time to produce the information necessary. Once the impasse has been handled, Soar creates a *chunk* that encodes a generalized explanation of the result in terms of the original goal context. Intermediate steps from the solution are lost, but the acquired chunk lets the system sidestep similar impasses in the future.

Anderson's (1993) ACT-R employs a related mechanism, called *compilation*, which creates new production rules from ones that are involved in the same reasoning chain. This scheme produces very specific rules that replace variables with the declarative elements against which they matched, rather than forming generalized structures, as do ICARUS and most other systems that learn macro-operators or search-control rules. In fact, our approach is more akin to the composition process that played a role in earlier versions of ACT (Neves & Anderson, 1981), though this mechanism produced fixed behavioral sequences rather than flexible skill hierarchies.

ICARUS' closest architectural relative is PRODIGY (Minton, 1990), which invokes means-ends analysis to solve problems and uses an analytical method to learn either search-control roles or macro-operators from problem-solving traces. Veloso and Carbonell (1993) also describe an extension that records these traces in memory and solves new problems by derivational analogy with earlier ones. None of these mechanisms generates explicit hierarchical structures, but Veloso and Carbonell's approach provides flexibility similar to that found in ICARUS, and the two systems record and utilize very similar information in their goal stacks.

Some other systems support learning in problem-solving domains without making strong architectural commitments. Ruby and Kibler's (1991) SteppingStone learns generalized rules for decomposing complex problems into simpler ones, which it obtains through mixed use of existing problem-reduction rules and forward-chaining exhaustive search when it reaches an impasse. Marsella and Schmidt's (1993) PRL system also acquires task-decomposition rules that incorporate partial orderings among components. Their system combines forward and backward search to identify candidate state pairs, which in turn produce hypothesized problem-reduction rules that are revised based on further experience.[9]

Perhaps the closest relative to our approach is Reddy and Tadepalli's (1997) X-Learn, which acquires goal-decomposition rules from a sequence of training problems. Their system does not include an execution engine, but it generates recursive hierarchical plans in a manner that also identifies declarative goals with the heads of learned clauses. However, because it invokes forward-chaining rather than backward-chaining search to solve new problems, it relies on the trainer to determine hierarchical structure. X-Learn also uses a quite sophisticated mixture of analytical and inductive techniques to determine conditions on skills, rather than the much simpler method that ICARUS incorporates.

Another key difference from X-Learn, PRL, and Steppingstone is that ICARUS learns skills for use in reactive execution rather than for use in planning. There has been other work on this topic, but it has emphasized the acquisition of flat controllers rather than hierarchical structures. For instance, Benson's (1995) TRAIL learns teleoreactive controllers for physical agents, but it invokes inductive logic programming to determine rules for individual actions. Fern, Yoon, and Givan (2004) report an approach to learning reactive controllers that trains itself on increasingly complex problems, but that also acquires decision lists for action selection. Khardon (1999) considers the related task of learning hierarchical controllers, but his

---

[9] Ilghami, Nau, Muñoz-Avila, and Aha (2002) present another system that organizes plan knowledge in a hierarchical manner, but it learns conditions for clause selection rather than the structure of the hierarchy itself.

formal analysis assumes the agent is provided with annotated sample solutions rather than being generated through problem solving.

Other researchers have built systems that support cumulative learning outside the context of problem-solving tasks. One early example was Sammut and Banerji's (1986) Marvin, which learns increasingly complex logical concepts that are composed of ones it has mastered previously. Stone and Veloso (2000) take a similar approach to learning concepts and controllers for playing robotic soccer, although their system acquires quite different types of structure at each level of description. Utgoff and Stracuzzi's (2002) STL algorithm receives training cases about many concepts in parallel, but it learns complex ones only when it has acquired simpler structures that let it master them with little effort. Pfleger (2004) describes another system that acquires hierarchical patterns in an on-line setting, in this case from unsupervised data. Like Marvin and STL, it learns conceptual structures from the bottom up, so that more complex patterns are apparent after simpler ones have been acquired.

8. Discussion

In the preceding pages, we presented ICARUS, a cognitive architecture for physical agents that uses stored concepts and skills, both organized in hierarchies, to recognize familiar situations and control behavior. We described a new module that supports means-ends problem solving on novel tasks, along with a learning mechanism that produces new skills from traces of problem solutions. This method operates in an incremental manner, creating hierarchical structures that refer to others learned earlier. In addition, we reported experiments with urban driving, the Blocks World, and FreeCell that showed such learning enables more effective behavior on unfamiliar problems than solving them with only basic knowledge about the domain.

We have focused on skill learning, which in humans occurs continuously throughout life, but which plays an especially important role early in development, when children are first learning to interact with their environments. There is a long tradition of using learning mechanisms to explain developmental phenomena that occur over extended periods. For example, Sage and Langley (1983) model development on the Piagetian balance scale task in terms of discrimination learning, whereas Jones and Van-Lehn (1994) model long-term changes in addition strategies using a form of probabilistic learning. There seems no reason, in principle, why our learning mechanisms cannot explain key aspects of cognitive development, but we cannot claim to account for them all.

For example, our framework relies on accurate models of the actions' effects, cast as primitive skills, while children clearly acquire such models from experience. However,

Benson (1995) reports one approach to learning action models that can be used to acquire teleoreactive controllers, so extensions seem possible. We have also assumed the ability to carry out means-ends problem solving, which appears fairly late in childhood. But elsewhere (Nejati, Langley, & Konik, 2006), we have reported an alternative mechanism that learns hierarchical skills from traces of successful exploration in the environment. This carries out a form of goal regression over the solution trace, much as in means-ends analysis, but without relying on conscious problem solving, providing a more plausible mechanism for early skill acquisition.

Nevertheless, our work on learning and development in ICARUS is still in its early stages. We should demonstrate its ability to acquire hierarchical structures on additional domains that include both cognitive tasks like multi-column subtraction and on dynamic domains that, like urban driving, require the integration of cognition with reactive control. One promising class of cognitive domains involves games like chess, which seem certain to introduce new challenges because of their extended duration. Future work on driving should show that our methods are sufficient to acquire more complicated skills that involve extended tasks like package delivery and complex settings that include other vehicles.

We are also interested in connecting the architecture to humanoid agents. We have taken initial steps in this direction by creating an ICARUS agent for Urban Combat, a first-person shooter game in which the player must move around in a three-dimensional setting, overcome obstacles, and capture a flag. The system uses the same representation, inference, and control methods we have described earlier, but it does not utilize means-ends problem solving and it learns hierarchical skills from execution traces, as outlined above. In the longer term, we hope our experience in this domain will let us connect the architecture to physical robots.

In addition, ICARUS' methods for problem solving and hierarchical learning would benefit from new capabilities. We noted earlier that the current system does not chain backward from the start conditions of learned skill clauses. Extending the problem solver to support this ability would mean defining new concepts that characterize the situations in which learned skills are applicable. This addition would also remedy another limitation of the current system, namely its inability to account for the origin of concept hierarchies, which it assumes are given. Such an extension would be straightforward for some tasks, but others will require the ability to acquire recursive concepts. Augmenting the system in this manner may also lead to a utility problem, not during execution of learned skills but during the problem solving used for their acquisition, which we would then need to overcome.

Another drawback is the architecture's reliance on purely deductive inference, which differs markedly from

the probabilistic approach taken by its earliest ancestor (Langley et al., 1991). Future versions of the framework should extend the representation of concepts and skills to incorporate probabilities, replace deductive processes with abductive methods that make plausible default inferences, and augment problem solving to operate over skills with uncertain outcomes. We hypothesize that the current mechanisms for learning the structure of skills can be adapted easily to this setting, but we should also introduce methods for estimating the probabilities that annotate the symbolic structures.

We should also note that, although our approach learns skills that generalize to situations with different numbers of objects, its treatment of goals is less flexible. For example, ICARUS can acquire a general procedure for clearing a block that does not depend on the number of blocks above it, but it cannot learn a procedure for constructing a tower with arbitrarily specified components. Extending the method's ability to learn about such recursive goal structures is another important direction for future research that will bring the architecture into closer alignment with the abilities observed in human cognition.

Acknowledgements

Appendix: Concepts and skills provided in experiments

Appendix A

Concepts and skills provided to ICARUS for the urban driving domain, with parentheses indicating the number of clauses for disjunctive skills.

Primitive concepts (15)	Nonprimitive concepts (26)	Primitive skills (19)
stopped	parked	in-intersection-for-right-turn
moving	aligned-with-lane-in-segment	aligned-with-lane-in-segment
in-segment	centered-in-lane	steering-wheel-straight
in-intersection-for-right-turn	steering-wheel-not-straight	centered-in-lane (2)
in-intersection	driving-in-segment	in-lane (2)
intersection-ahead	at-speed-for-right-turn	stopped
segment-to-right	ready-for-right-turn	moving
on-right-side-of-road-in-segment	in-leftmost-lane	adjust-speed-for-cruise
in-lane	lane-to-right	adjust-speed-for-right-turn (2)
steering-wheel-straight	lane-to-left	get-on-right-side-of-road
at-speed-for-cruise	in-rightmost-lane	cruise-within-segment
slow-for-right-turn	in-right-turn-lane	steer-for-right-turn
fast-for-right-turn	off-centered-to-right-in-segment	change-lane-to-right
first-lane	off-centered-to-left-in-segment	change-lane-to-left
last-lane	building-on-right	cruise-into-intersection
	building-on-left	cruise
	current-building	
	start-aligned-with-lane-in-segment	
	start-centered-in-lane (2)	
	start-adjust-speed-for-cruise	
	start-cruise-within-segment	
	start-change-lane-to-right	
	start-change-lane-to-left	
	start-in-lane (2)	

Appendix B

Concepts and skills provided to ICARUS for the Blocks World, with goal concepts in italics.

Primitive concepts (4)	Nonprimitive concepts (7)	Primitive skills (4)
on	*clear*	unstack
ontable	*three-tower*	pickup
holding	*two-tower-one-on-table*	stack
hand-empty	unstack-able	putdown
	pickup-able	
	stack-able	
	putdown-able	

Appendix C

Concepts and skills provided to ICARUS for FreeCell solitaire, with goal concept in italics.

Primitive concepts (10)	Nonprimitive concepts (14)	Primitive skills (12)
starthome	highest	column-to-home
successor	*game-won*	column-to-newhome
colcolpair	column-to-home-able	column-to-freecell
available-cell	column-to-newhome-able	lastcolumn-to-home
available-column	column-to-freecell-able	lastcolumn-to-newhome
clear	lastcolumn-to-home-able	lastcolumn-to-freecell
on	lastcolumn-to-newhome-able	freecell-to-home
bottom	lastcolumn-to-freecell-able	freecell-to-newhome
incell	freecell-to-home-able	freecell-to-column
home	freecell-to-newhome-able	column-to-column
	freecell-to-column-able	freecell-to-new-column
	column-to-column-able	column-to-new-column
	freecell-to-new-column-able	
	column-to-new-column-able	

References

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum.

Asgharbeygi, N., Nejati, N., Langley, P., & Arai, S. (2005). Guiding inference through relational reinforcement learning. In *Proceedings of the fifteenth international conference on inductive logic programming.* Bonn, Germany: Springer.

Benson, S. (1995). Inductive learning of reactive action models. In *Proceedings of the twelfth international conference on machine learning* (pp. 47–54). San Francisco: Morgan Kaufman.

Choi, D., Kaufman, M., Langley, P., Nejati, N., & Shapiro, D. (2004). An architecture for persistent reactive behavior. In *Proceedings of the third international joint conference on autonomous agents and multi agent systems* (pp. 988–995). New York: ACM Press.

Choi, D., & Langley, P. (2005). Learning teleoreactive logic programs from problem solving. In *Proceedings of the fifteenth international conference on inductive logic programming.* Bonn, Germany: Springer.

Doyle, J. (1979). A truth maintenance system. *Artificial Intelligence, 12,* 231–272.

Fern, A., Yoon, S. W., & Givan, R. (2004). Learning domain-specific control knowledge from random walks. In *Proceedings of the fourteenth international conference on automated planning and scheduling* (pp. 191–199). Whistler, BC: AAAI Press.

Forgy, C. L. (1982). Rete: A fast algorithm for the many pattern/many object pattern match problem. *Artificial Intelligence, 19,* 17–37.

Iba, G. A. (1989). A heuristic approach to the discovery of macro-operators. *Machine Learning, 3,* 285–317.

Ilghami, O., Nau, D. S., Muñoz-Avila, H., & Aha, D. W. (2002). CaMeL: Learning method preconditions for HTN planning. In *Proceedings of the sixth international conference on AI planning and scheduling* (pp. 131–142). Toulouse, France.

Jones, R. M., & VanLehn, K. (1994). Acquisition of children's addition strategies: A model of Impasse-free, knowledge-level learning. *Machine Learning, 16,* 11–36.

Jones, R. M., & Langley, P. (2005). A constrained architecture for learning and problem solving. *Computational Intelligence, 21,* 480–502.

Khardon, R. (1999). Learning to take actions. *Machine Learning, 35,* 57–90.

Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning, 1,* 11–46.

Langley, P., & Choi, D. (2006). Learning recursive control programs from problem solving. *Journal of Machine Learning Research, 7,* 493–518.

Langley, P., Cummings, K., & Shapiro, D. (2004). Hierarchical skills and cognitive architectures. In *Proceedings of the 26th annual conference of the cognitive science society* (pp. 779–784). Chicago, IL.

Langley, P., McKusick, K. B., Allen, J. A., Iba, W. F., & Thompson, K. (1991). A design for the ICARUS architecture. *SIGART Bulletin, 2,* 104–109.

Langley, P., & Rogers, S. (2005). An extended theory of human problem solving. *Proceedings of the twenty-seventh annual meeting of the cognitive science society.* Stresa, Italy.

Langley, P., & Stromsten, S. (2000). Learning context-free grammars with a simplicity bias. In *Proceedings of the eleventh European conference on machine learning* (pp. 220–228). Barcelona: Springer.

Marsella, S., & Schmidt, C. F. (1993). A method for biasing the learning of nonterminal reduction rules. In S. Minton (Ed.), *Machine learning methods for planning.* San Mateo, CA: Morgan Kaufmann.

Minton, S. N. (1990). Quantitative results concerning the utility of explanation-based learning. *Artificial Intelligence, 42,* 363–391.

Mooney, R. J. (1989). The effect of rule use on the utility of explanation-based learning. In *Proceedings of the eleventh international joint conference on artificial intelligence* (pp. 725–730). Detroit: Morgan Kaufmann.

Nardi, D., & Brachman, R. J. (2002). An introduction to description logics. In F. Baader et al. (Eds.), *Description logic handbook.* Cambridge: Cambridge University Press.

Nejati, N., Langley, P., & Konik, T. (2006). Learning hierarchical task networks by observation. In *Proceedings of the 23rd international conference on machine learning.* Pittsburgh, PA.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R.

Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

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

Newell, A., Shaw, J. C., & Simon, H. A. (1960). Report on a general problem-solving program for a computer. In *Information processing: Proceedings of the international conference on information processing* (pp. 256–264). Paris: UNESCO House.

Nilsson, N. (1994). Teleoreactive programs for agent control. *Journal of Artificial Intelligence Research, 1*, 139–158.

Pfleger, K. (2004). On-line cumulative learning of hierarchical sparse n-grams. In *Proceedings of the third international conference on development and learning*. San Diego, CA: IEEE Press.

Reddy, C., & Tadepalli, P. (1997). Learning goal-decomposition rules using exercises. In *Proceedings of the fourteenth international conference on machine learning*. San Francisco: Morgan Kaufmann.

Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory using EPAM IV. *Psychological Review, 102*, 305–330.

Ruby, D., & Kibler, D. (1991). SteppingStone: An empirical and analytical evaluation. In *Proceedings of the tenth national conference on artificial intelligence* (pp. 527–532). Menlo Park, CA: AAAI Press.

Sage, S., & Langley, P. (1983). Modeling development on the balance scale task. In *Proceedings of the eighth international joint conference on artificial intelligence*. Karlsruhe, West Germany: Morgan Kaufmann.

Sammut, C., & Banerji, R. B. (1986). Learning concepts by asking questions. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.). *Machine learning: An artificial intelligence approach* (Vol. 2). Los Altos, CA: Morgan Kaufmann.

Shapiro, D., Langley, P., & Shachter, R. (2001). Using background knowledge to speed reinforcement learning in physical agents. In *Proceedings of the fifth international conference on autonomous agents* (pp. 254–261). Montreal: ACM Press.

Shavlik, J. W. (1989). Acquiring recursive concepts with explanation-based learning. In *Proceedings of the eleventh international joint conference on artificial intelligence* (pp. 688–693). Detroit, MI: Morgan Kaufmann.

Stone, P., & Veloso, M. M. (2000). Layered learning. In *Proceedings of the eleventh european conference on machine learning*. Barcelona: Springer-Verlag.

Utgoff, P., & Stracuzzi, D. (2002). Many-layered learning. In *Proceedings of the second international conference on development and learning* (pp. 141–146).

Veloso, M. M., & Carbonell, J. G. (1993). Derivational analogy in PRODIGY: Automating case acquisition, storage, and utilization. *Machine Learning, 10*, 249–278.