

DEVELOPMENTAL PROGRESS IN THE
USE OF WEAK METHODS:
FROM HILL-CLIMBING TO SUBGOALING

Sharon McCoy Carver and David Klahr

Society for Research in Child Development

April 1985 Biennial Meeting

Toronto, Canada

Developmental Progress in the use of Weak Methods: From Hill-Climbing to Subgoaling

[Slide 1] The title of this talk highlights our central issue: the development of general problem-solving skills, or weak methods. The subtitle foreshadows our conclusion: development progresses from hill-climbing to subgoaling [Slide 2] Our talk will be organized as follows: first, we will describe these two weak methods, and then we will describe a series of experiments that assess children's and adults' ability to use them.

[Slide 3] Research on the development of problem-solving abilities is facilitated by our understanding of how normal adults solve problems. We know that when faced with novel problems, adults call on a set of general strategies, called weak methods, that require little task-specific knowledge. However, little is known about what forms of these methods are available to young children and what kinds of difficulties children have in using them. We don't have time here to discuss the full repertoire of methods that adults use. Instead, we will contrast two very different kinds of weak methods: means-ends analysis and hill-climbing.

[Slide 4] Means-ends analysis requires that the problem solver characterize the problem in terms of differences between the current state and the goal state. Once the differences have been characterized, the problem solver must set a subgoal to eliminate one of the differences. This subgoal constitutes a subproblem to which the entire problem-solving procedure is applied recursively. When all of the subgoals have been achieved, the problem is solved. A crucial property of means-ends analysis is the ability to generate and order subgoals appropriately.

[Slide 5] For example, in the well-known Tower of Hanoi problem, the first subgoal is to get the largest disk to the goal peg. Once that is accomplished, a subgoal can be set to get the next sized disk to the goal peg, and so on. But what if the subgoal ordering is not clear? [Slide 6] In this version of the Tower of Hanoi, it is not clear that any one disk should reach its goal peg first. Here, in fact, the medium-sized disk would reach its position first in the minimum path solution. In situations where the subgoal ordering is ambiguous, means-ends analysis is not very effective.

[Slide 7] Hill-climbing is a different type of weak method. In hill-climbing, the problem solver generates a few alternative moves and computes an evaluation function for the outcome of each potential move. One of the simplest evaluation functions merely counts the number of pieces that are currently in their desired configuration. A hill-climber prefers the move that yields the largest increase in the local evaluation. Problems with an evaluation

function that is easy to compute would facilitate this weak method

[Slide 8] Several previous investigations (such as Spitz & Borys and Klahr & Robinson) have shown that pre-school children spontaneously use means-ends analysis which requires the use of subgoals. However, Klahr and Robinson also demonstrated that children's effective use of subgoaling was limited to problems having few subgoals and clear subgoal ordering.

[Slide 9] In our work, we investigate pre-schoolers' and adults' ability to solve problems in which means-ends analysis is ineffective. We have chosen to use several versions of the depth of search puzzle invented by Suzanne Borys because the choice of which piece should reach its destination first in such puzzles is ambiguous. By using a problem that precludes the use of subgoals, we can assess the extent to which children and adults use other weak methods. As you will see, the evaluation function for the Depth of Search puzzles can be calculated easily, so we expected it to be a good candidate for evoking hill-climbing.

[Slide 10] This slide summarizes our research. On a simple depth of search puzzle, the Dog-Cat-Mouse puzzle, where subgoals were ambiguous and means-ends analysis was not possible, children did not make moves at random but rather they fell back on a simple hill-climbing procedure. On a more difficult depth of search puzzle, the Five puzzle, children again used a hill-climbing strategy. Some adults also start with hill-climbing; but as they get experienced with the puzzle, they form an abstract representation which facilitates the use of a subgoaling strategy.

[Slide 11] We began by studying children solving a simple depth of search puzzle, isomorphic to the one described earlier by Suzanne Borys. The Dog-Cat-Mouse puzzle, as we call it, consists of three toy animals and three toy foods that "belong" to the animals (a bone for the dog, a fish for the cat, and cheese for the mouse). The goal of the problem is to move each animal to its corresponding food. Remember that this puzzle was chosen because it has ambiguous subgoal ordering: the order in which the animals will reach their foods is not at all obvious. In this case, it will be the mouse.

[Slide 12] In this diagram of the DCM state space, each of the 24 nodes represents one of the legal configurations of the three animals. All problems are defined in terms of their initial states (the arrangement of the animals) and their final states (the arrangement of the foods). For example, the problem shown in the last slide starts at node 11 and ends at node 20. Three properties of the state space are relevant to our discussion. First, problems can be characterized in terms of their path length: the minimum number of moves required to get from the initial state to the final state. In our example, 11 to 20

has a path length of 5. Second, problems can be either rotation or permutation problems. Rotation problems have both initial and final states on the same hexagon - either the inner or the outer. They have minimum paths that do not use the diagonal of the game board. Permutation problems have initial and final states on different hexagons and require use of the diagonal. These problems start and end with different permutations of the three animals. The permutation order can be changed only by using the diagonal. In our example, 11 to 20 is a permutation problem which begins with the animals in the order D-M-C-D and ends in the order D-C-M-D. Thirdly, problems can begin at nodes with either open or closed diagonals. Nodes with open diagonals have three adjacent states so there are three possible moves from an open node. Nodes with closed diagonals have two adjacent states so there are only two possible moves from a closed node. Our example problem begins at node 11, a closed node.

[Slide 13] Eight problems varying in path length, type of initial node, and problem type were used. In addition, four three-move training problems were used to familiarize children with the rules of the game. The data reported are from 39 children, ranging in age from 45 to 70 months old, who attended the Carnegie-Mellon University Children's School. Problems were presented in order of increasing path length. Children were given two chances to produce a minimum path solution to each problem.

For each problem, subjects were assigned a 1/0 score based on whether or not they found a minimum path solution by the second presentation of the problem. Each problem was assigned a score based on the number of subjects passing it. Subjects' performance varied widely: the highest performing subject solved all but one problem, while three subjects failed all but one.

[Slide 14] Problem difficulty also varied widely: from nearly all subjects passing the easiest problem to over 80 percent failing the hardest problem. This graph shows the proportion of subjects who found a minimum path solution to each problem; problems are numbered in order of increasing path length.

The differences in problem difficulty cannot be explained by the structural variables described earlier. Path length is a poor predictor of problem difficulty. The two easiest problems (1 and 2) are also the two shortest, but even though they both have a path length of 4, there is a 30% difference in the proportion of subjects passing them. The next easiest problems (7 and 8) are the two longest (7 moves). The four hardest problems are intermediate in path length; even within that set, there is a large difference between the pairs with the same length. Overall, the correlation between path length and problem difficulty is not significant. Neither of the two other independent variables (permutation vs

rotation and open vs. closed node) had a reliable effect on problem difficulty

Another possible index of problem difficulty is what Spitz and Borys call subgoal length or "depth of search". This index is defined as the number of steps on the minimum path before the first object reaches its final position. Thus defined, subgoal length explains only 43% of the variance in problem difficulty. This marginally significant correlation is further evidence that the DCM puzzle has ambiguous subgoals.

[Slide 15] Path length, problem type, node type, and subgoal length are all structural variables: features of the problem rather than of the problem-solving process. Even if they are good predictors of difficulty, they leave unstated the underlying processes that they affect. For example, the subgoal length calculation just described is based on a tenuous assumption: that children can determine which piece to focus on first in choosing subgoals. But, this is precisely what makes the DCM puzzle difficult.

[Slide 16] The general point is that structural variables alone do not cause behavior directly: they are mediated by underlying processes. In some cases, the process model is so obvious that it need not be made explicit. But in situations of even modest complexity, such as the DCM puzzle, there are several plausible processes - or components of weak methods - and their interactions can only be understood by formulation of an explicit process model.

[Slide 17] We will present evidence for children's sensitivity to three components of weak methods: a constraint against back-up, a 2 or 3 move search for goal states, and a partial evaluation of progress toward the goal.

One of the most rudimentary forms of efficient problem solving is avoidance of unnecessary moves. In the DCM puzzle, moving the same piece twice in succession always results in a two-move sequence having no effect: the puzzle returns to the state occupied at the start of the sequence. If moves are made at random, without regard to this "no-backup" constraint, then we would expect 33% of moves at open nodes and 50% of those at closed nodes, or 42% of all moves, to be double moves. In fact, double moves were rare: out of 3,350 moves, only 10% were doubles.

Another rudimentary ability inherent in several of the weak methods is evaluation of the quality of a proposed move. The simplest evaluation is binary: a state either matches the goal or it does not. A solver could search as far forward as necessary, testing each state for whether or not it was a goal state. Once having detected the goal, the problem solver would simply follow the path that led to it. If a subject had the ability to search n moves ahead for the goal, then we should see perfect performance from n steps away. Two-thirds of the subjects could produce perfect solutions from at least 2 moves away, and one-third

could even do it from 3 moves away. None could reliably find minimum path solutions more than 4 moves distant. We grant subjects a 2 to 3 move capacity to search for the goal state.

Much more useful than a binary evaluation, however, is a partial evaluation that gives some measure of how well the current state matches the goal state. If the children were using such an evaluation function, then we should see two types of biases in their move patterns. One bias would show up as a tendency to favor moves that increase the number of pieces in their goal positions or move pieces closer to their goal positions. [Slide 18] For example, in Problem 2, a first move of the cat increases the evaluation function, while moving the dog does not. The dog is also off the minimum path. Over all trials and all subjects on this problem, the cat was moved 81% of the time. Even more revealing are the "garden path" problems. [Slide 19] In problem 4, the minimum path move is the mouse, which does not increase the evaluation function. Only the cat increases the partial evaluation function, and it is preferred on 66% of the trials even though it is off the minimum path.

The other bias would be a reluctance to remove pieces from their goal locations - to reduce the value of the partial evaluation. [Slide 20] This can be assessed on problem 3 where the minimum path sequence requires that the dog be temporarily removed from its goal position. On 65% of all trials with Problem 3, subjects preferred to move the cat rather than the dog, even though this took them off the minimum path. As you can see, this reliance on a partial evaluation - a hill-climbing strategy - can be dysfunctional since it can lead to local maxima that are isolated from the goal.

[Slide 21] Let us return to the measure of problem difficulty. This slide shows the solution rates for the eight problems listed in order of increasing difficulty. These problems are divisible into four types. On type I problems, all of the moves increase the evaluation function, so the solution is easily reached using a hill-climbing strategy. Type II problems can be solved by moves which either maintain or increase the evaluation function; the hill-climber is forced to make one or more moves at random so performance is not as good as on type I problems. Type III problems require the subject to make a move which decreases the evaluation function; at this point hill-climbing is reduced to random search. Finally, type IV problems require subjects to avoid moving a piece which would increase the evaluation function; a hill-climber will rarely do this as you can see by the poor performance on those three problems.

We have used a computer simulation model to test whether our characterization of the children's strategy is sufficient to predict their performance. The model makes moves

according to the following rules:

- If there is a two-move sequence that can reach the goal state, then make it, otherwise:
- Generate all candidate moves. On all but 10% of the trials, delete the piece just moved from the candidate set. (This percentage reflects the finding that 10% of the children's moves were backup.)
- If there is more than one candidate, then compute the evaluation function between each candidate node and the goal node. Choose the move with the maximum evaluation on 69% of the trials (since this is how often children did), or if all evaluations are equal choose randomly.
- Go to step 1.

This model represents an imperfect hill-climber who occasionally backs up and who does not always follow the evaluation function. Each problem was presented to the program 400 times; the solutions generated by the program were scored using the same 1/0 rating as we used for the children. The simulation model accounted for 70% of the variance in the children's performance.

Let me summarize the findings from the first study. We presented preschoolers with problems having ambiguous subgoal ordering to discover what weak methods they could invoke when means-ends analysis was not useful. In this situation, they did not resort to random trial and error. Their avoidance of double moves suggests a rudimentary knowledge about thoroughly useless actions that is not conveyed by the "trial and error" view of young children. The second important finding is that solutions are not simply arrived at by chance, since there is a lookahead to the goal state, and little deviation from the minimum path once it is in sight. Third, children use partial evaluations of nearly correct states to guide their choice of moves. This sensitivity to incremental progress may actually degrade children's performance (as in the garden path problems). Nevertheless, it is reasonable for children to attempt to use such information.

For the second phase of our research, we chose a more difficult version of the DCM puzzle so we could observe adults' solution strategies and then compare children and adults on the same puzzle. [Slide 22] The 5-puzzle has the same simple rules as the DCM puzzle, but problems can vary in difficulty enough to challenge adults. In contrast to the DCM's 4 spaces with 3 pieces, the 5-puzzle has 6 spaces with 5 pieces, 3 of which are identical. As with the DCM puzzle, pieces can be moved into the blank space from any of the adjacent spaces. The goal is to get from an initial configuration to a final configuration in the minimum number of moves. The state space for the 5-puzzle contains 120 possible

configurations of the pieces: this extensive space allows for problems up to 18 moves in length. As with the DCM puzzle, two types of moves are possible: rotation moves maintain the permutation of the pieces and permutation moves change the permutation of the pieces. In the 5-puzzle, there are 4 possible permutations of the pieces as opposed to only 2 for the DCM puzzle.

Seven Carnegie-Mellon University Undergraduates participated in a preliminary protocol study using the 5-puzzle. Each subject solved one sample and 10 test problems on each of two versions of the puzzle. The problems ranged in minimum path length from 5 to 18. Problems also varied in the number of necessary permutation moves (from 0 to 2). The order of the test problems was randomized.

Two performance measures were used: the percentage of solutions and the percentage of efficient solutions (solutions close to the minimum path).

[Slide 23] Solution rates for both versions of the puzzle did not differ so the results were collapsed across version. The total solution rate was near 100% on most problems. Problems did differ, however, in the percentage of efficient solutions. On this puzzle, efficient solutions are rare on long problems because there are so many solution paths.

Two distinct problem representations were evident from the subjects' protocol statements. [Slide 24] The 5-puzzle and the DCM puzzle can be represented as a set of individual pieces in positions. In addition, the unique pieces are represented separately from the common pieces. Goal statements such as "just get the red closer to where it's supposed to be" and "I'm going to try to get the yellow closer to where it's supposed to be" characterize this representation. The second representation is an abstract representation which focuses primarily on the permutation of the pieces and only secondarily on their exact positions. Subjects using this abstract representation typically make goal statements such as "I wanna switch the red and the yellow and no blues in between them now we have the right order. all's we have to do is rotate them around to where we want them, and presto magic"

Subjects using the piece-position representation have five subgoals to accomplish, one for getting each piece to its desired location. However, as in the DCM, the order of accomplishing these subgoals is ambiguous. Since subjects represent the unique pieces as special cases, they try to reposition them first. Choice of which unique piece to try first and which order to try the common pieces remains ambiguous. Subjects' goal statements reveal that they use a hill-climbing strategy to resolve the ambiguity. Their move choice is based on an evaluation of how much each move would bring them nearer to the goal. Subjects using the hill-climbing strategy can successfully solve problems 1, 2, 3, 4, 5, and

8. Problems 6, 7, 9, and 10 are garden path problems: they require temporary reductions of the partial evaluation. Hill-climbing often fails on these problems because it is reduced to random search when moves which increase the evaluation are not available

Subjects using the abstract permutation-position representation must accomplish two subgoals in a fixed order. First they establish the correct permutation of the pieces, and then they rotate the correctly ordered pieces into their desired positions. Choice of individual moves to correct the permutation is determined by further subgoaling which varies from subject to subject. The general procedure involves getting the unique pieces in the right order with respect to each other and then getting the right number of common pieces between them. Subjects using this subgoaling strategy can successfully solve all ten of the problems. However, they cannot necessarily solve them efficiently because each subgoal may be accomplished in a variety of ways. It could be that greater familiarity with the puzzle would lead to greater efficiency. However, the importance of being able to use the subgoaling strategy is not its efficiency but its ability to solve problems on which hill-climbing would fail.

Most of the adults represented the 5-puzzle as a permutation of pieces rotated into position. Those who did not begin with this representation often developed it when conflict arose, i.e. when hill-climbing failed.

The protocol study gave us some evidence that adults can shift from local evaluation and hill-climbing to a more abstract representation that facilitates subgoaling. In order to study this further with young children, we decided to present similar problems to children and adults in the form of a computer game. 17 Carnegie-Mellon Undergraduates and 4 Children's School Kindergarteners participated in the study.

[Slide 25] The findings for the adults replicate those from the preliminary protocol study. Once again the adults solved the problems at a high rate. The order of problem difficulty was also quite similar. The common levels of performance in the two studies suggests that the strategies adults used to solve the 5-puzzle were not affected by the presentation format.

Children, unlike adults, simply could not solve the majority of the problems presented to them. These children reached a total of nine solutions, all of which were on problems 1, 2, 3, and 5, problems on which a hill-climbing strategy could be effective. Informal observation of the children's performance also suggested that they used hill climbing; they avoided moving pieces once they were correctly positioned.

To further investigate the children's strategies, we designed an easier set of 5-puzzle problems. The set consisted of two 4-move, two 5-move, and four 6-move problems. Two

of the 6-move problems (called 6-star) required moving a unique piece out of its goal position on the first move of the minimum path. The set was designed to contrast three distinct models of the children's strategy: constrained random search (random search without backup), hill climbing, and subgoaling [Slide 26] If children use a constrained random search strategy, then predictions about their performance on specific problems can be generated by multiplying the probabilities of making each move on the minimum path. If children use the subgoaling strategy described earlier, they should be able to solve all 8 problems. Our hypothesis was that children do not use the abstract permutation-position representation so they cannot use subgoaling, but that they can use the piece-position representation which facilitates hill-climbing. If this is the case, children should be able to solve all but the 6* problems. One of the 6* problems has two moves on the minimum path which decrease the evaluation function and the other has only one. On the first, the hill-climber must make two random choices so has a 25% chance of reaching the solution in the minimum path; on the second, there is only one random choice so the chance is 50%.

11 Children's School kindergarteners participated in the study and all of them tried all of the problems. [Slide 27] Both the total solution rates and the minimum path solution rates are high for all but Problems 7 and 8, the 6* problems. So, the children are not using random search. But the poor performance on problems which require decreasing the evaluation function indicates that children are not able to use a subgoaling strategy but are, rather, using a hill-climbing strategy. The difference in performance on Problems 7 and 8 confirms our hypothesis that in detour situations, when no uphill moves are possible, children resort to random search.

Again, we wrote a computer simulation model to incorporate and evaluate all of our assumptions about how children solve this problem. It is a hill-climber that uses an evaluation function to rate each possible move. It has a 15% probability of backup. When none of the possible moves increase the evaluation function, the choice is made randomly. This model has a more complex evaluation function, which reflects the difference in the structure of the task. The evaluation function incorporates two extra factors: the difference between the unique and the common pieces and the distance from the goal. The unique pieces are given a weight of 2 and the common pieces a weight of 1; this reflects the subjects' focus on the unique pieces first and then the common pieces. The distance from the goal is measured by the number of moves necessary to get a piece from the current position to the goal position without any other pieces on the board. As with the other simulation, the model was run repeatedly to determine the model's prediction for the

percentage of minimum path solutions this hill-climbing strategy should produce. This model accounted for 90% of the variance in problem difficulty on the 8 problems.

Let me summarize the major findings of our research. We began with the goal of investigating children's repertoire of weak methods. We deliberately chose puzzles on which subgoaling would not be particularly effective so that we could determine what other weak methods children are able to use. [Slide 28] We have found that children have some rudimentary knowledge of what moves would be useless, that they have a 2 to 3 move look-ahead to the goal state, and that they are able to use a hill-climbing strategy. However, when the hill-climbing strategy fails, children search randomly. [Slide 29] In contrast, when adults find hill-climbing to be ineffective, they are able to re-represent the problem by abstraction. This abstraction allows them to use means-ends subgoaling.

DEVELOPMENTAL PROGRESS IN THE
USE OF WEAK METHODS:
FROM
HILL-CLIMBING TO SUBGOALING

Sharon McCoy Carver
David Klahr

Carnegie-Mellon University
Pittsburgh, PA
#1 city in Rand McNally's
Places Rated Almanac

Issue: the development of general problem-solving skills

Focus on two "weak methods"

Means-ends analysis

Hill-climbing

Describe empirical studies designed to assess use of methods

Conclusion: development progresses from hill-climbing to subgoaling

Adult Repertoire of Weak Methods

Generate and Test

Heuristic Search

Depth First

Breadth First

Best First

MEANS-ENDS ANALYSIS

HILL-CLIMBING

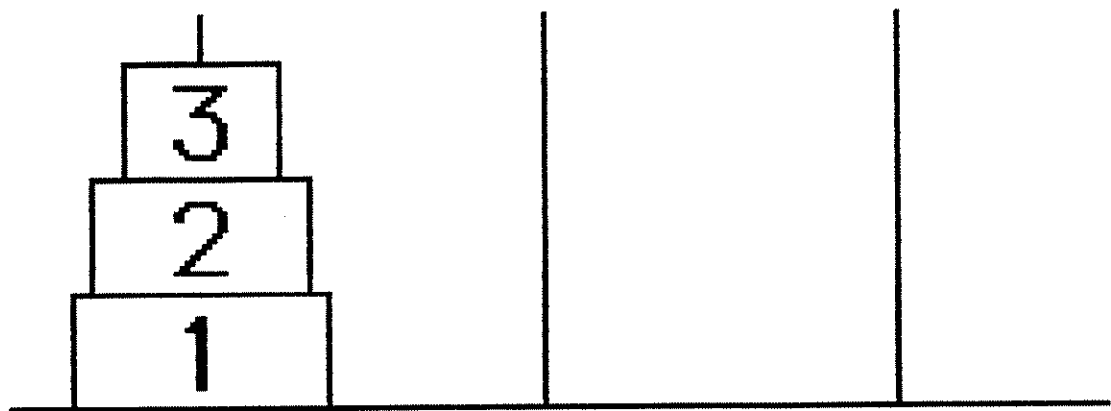
Trial and Error

MEANS-ENDS ANALYSIS

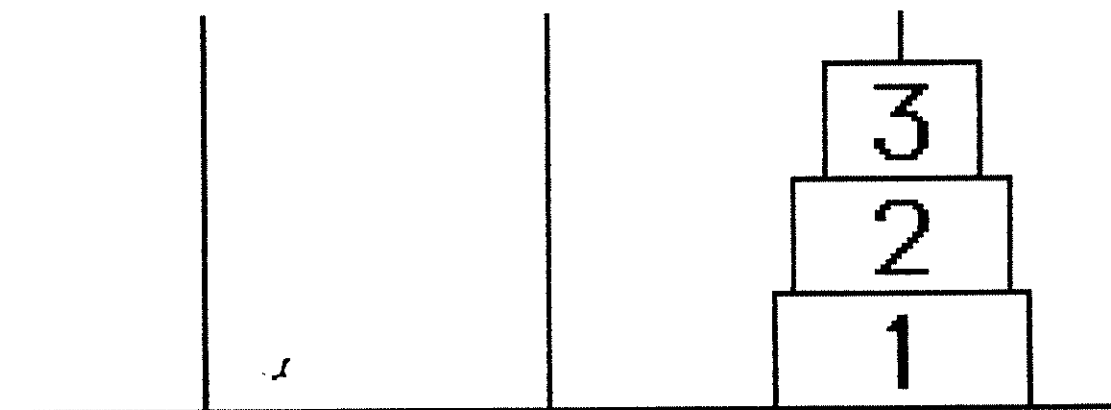
Goal: Transform C into G

Procedure:

- a) determine differences between C and G
- b) establish ordered sequence of subgoals
(g_1, g_2, g_3, \dots)
- c) apply procedure to each subgoal



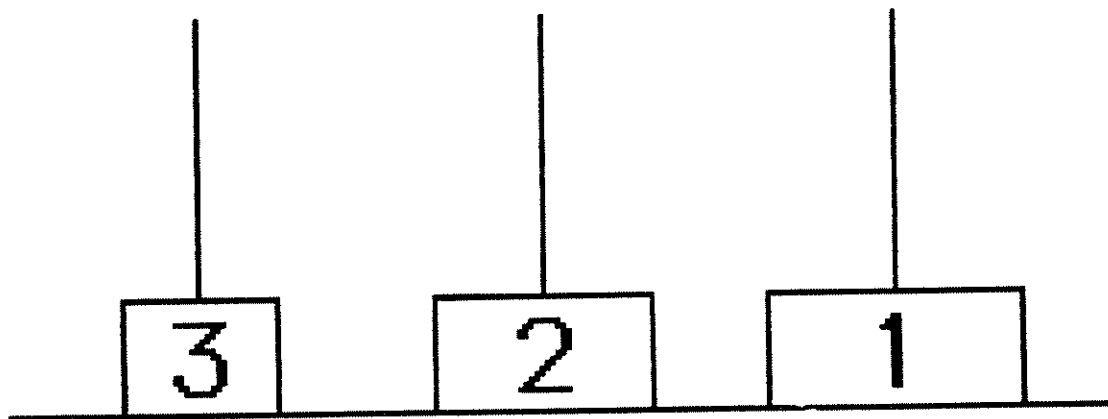
Initial State



Goal State



Initial State



Goal State

HILL-CLIMBING

Goal: Transform C into G

Procedure:

- a) generate alternative moves
- b) evaluate "goodness of fit"
- c) choose move that maximizes evaluation function

Child Repertoire of Weak Methods

Means-Ends Analysis

if few subgoals

if clear subgoal ordering

Trial and Error

Depth of search puzzles

- * ambiguous subgoal ordering precludes use of means-ends analysis

- * easy computation of evaluation function facilitates use of hill-climbing

DCM Puzzle

4-5 year olds

hill-climbing

2 move search for goal

Five Puzzle

5 year olds and adults

children use hill-climbing

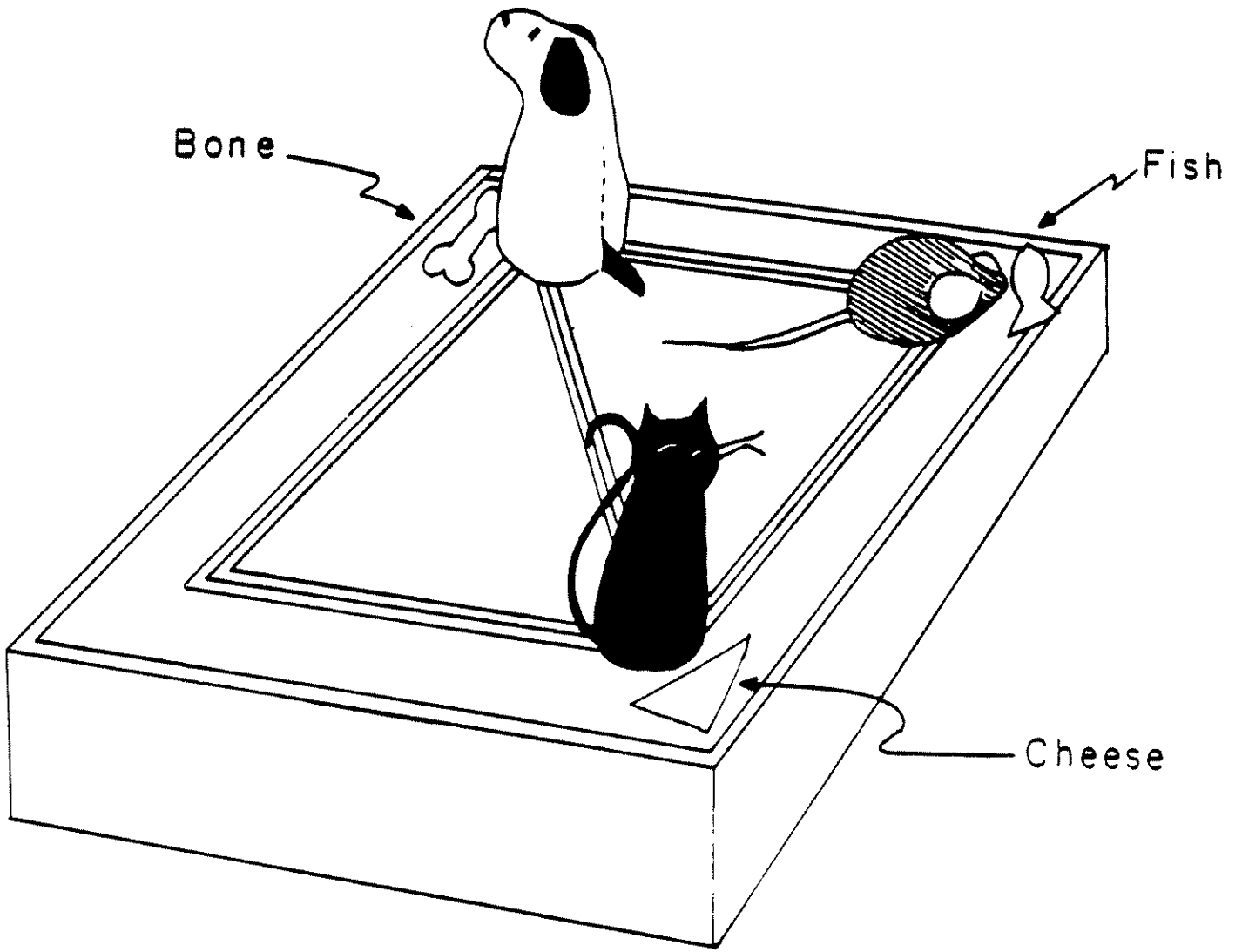
(as in DCM)

adults:

start with hill-climbing

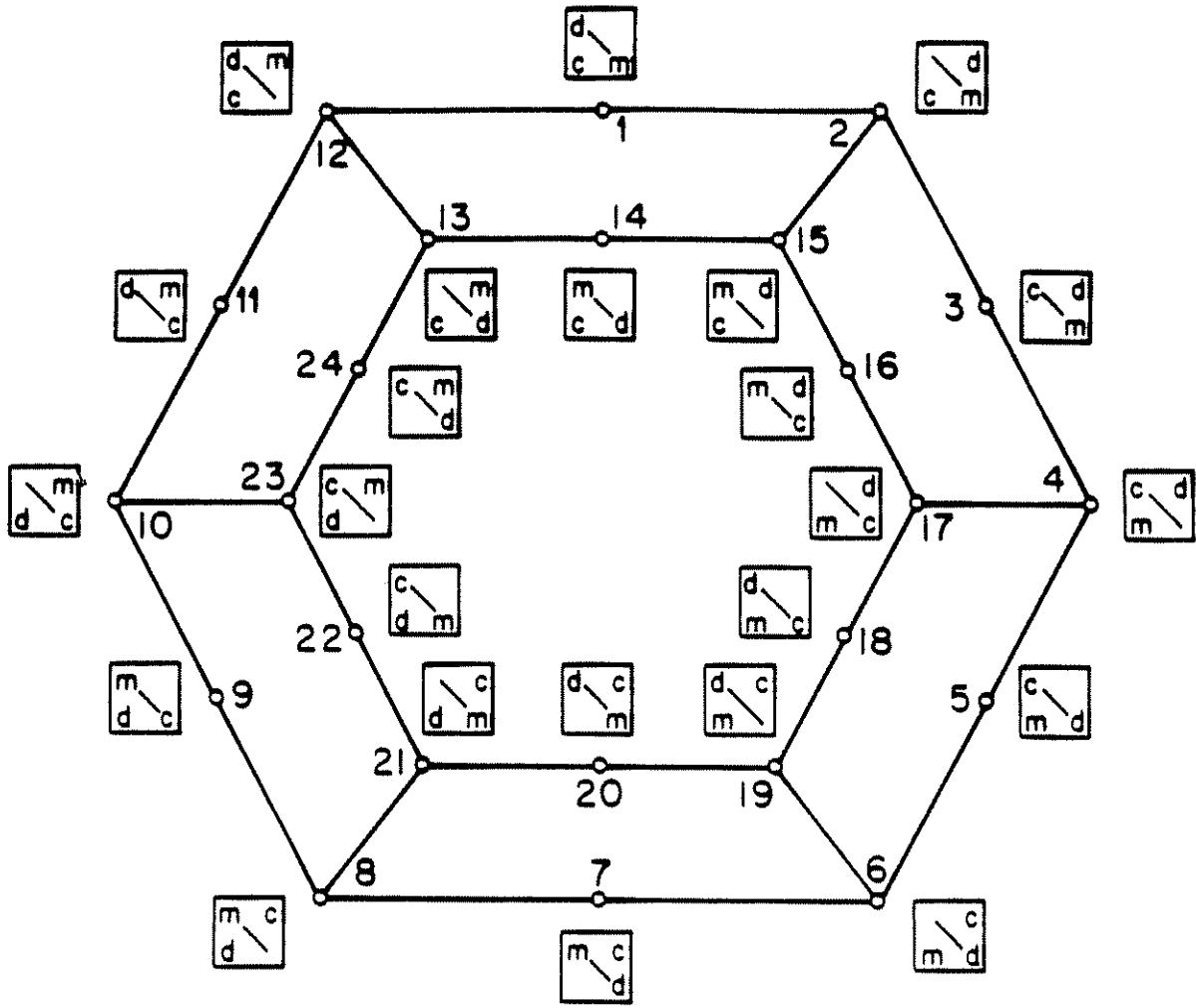
form abstract representation

then use subgoaling



DCM Puzzle

DCM State Space



.

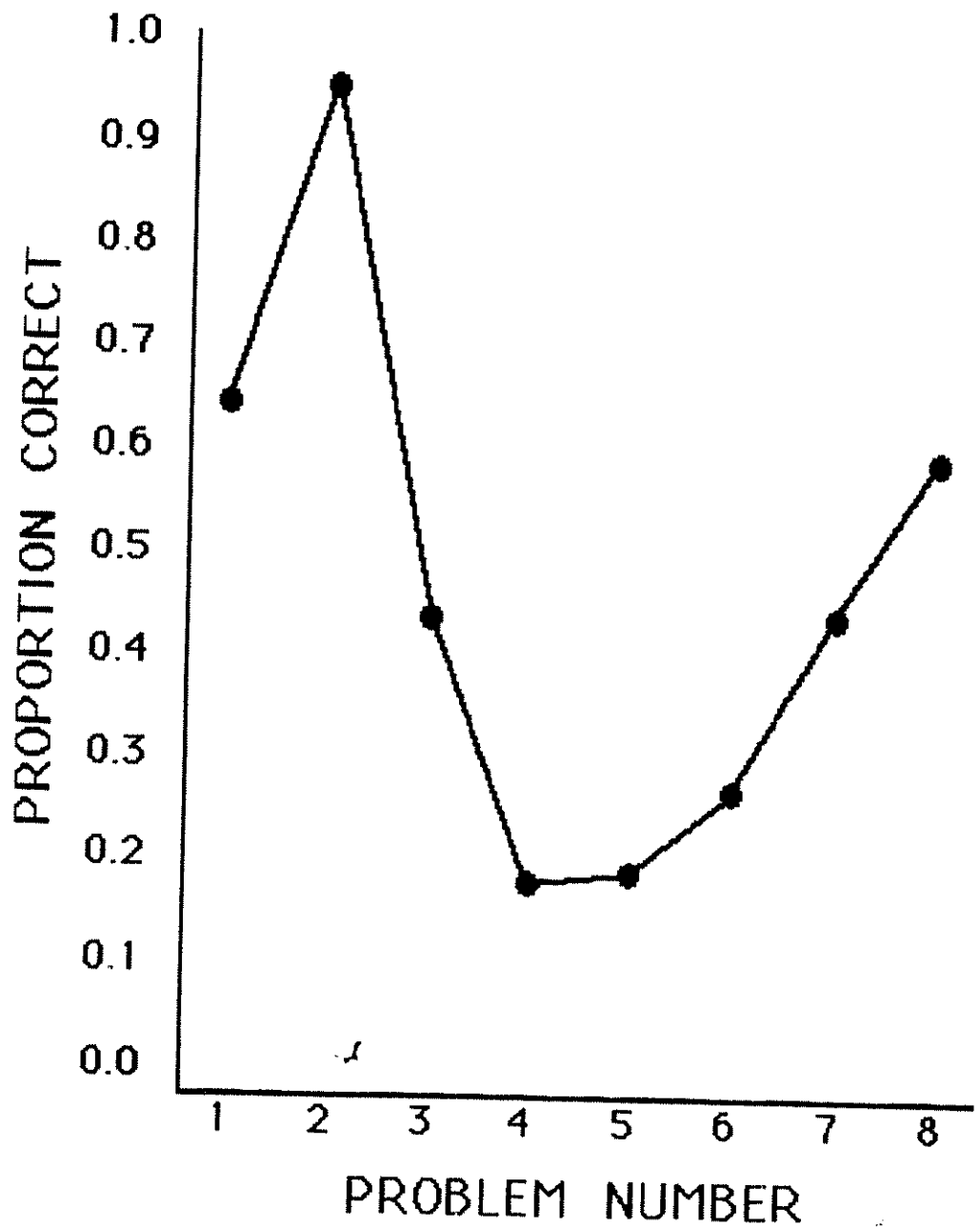
DCM Puzzle

3 Independent Variables:

Path Length (4 – 7)

Problem Type
(Rotation or Permutation)

Node Type
(Open or Closed)



Path Length

Problem Type

Node Type

Subgoal Length

Structural
Variables

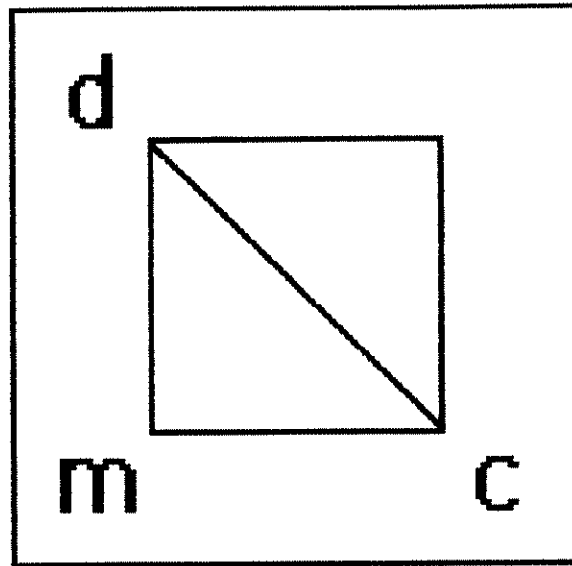
Structural variables do not
cause behavior;
they are mediated by
underlying processes.

1

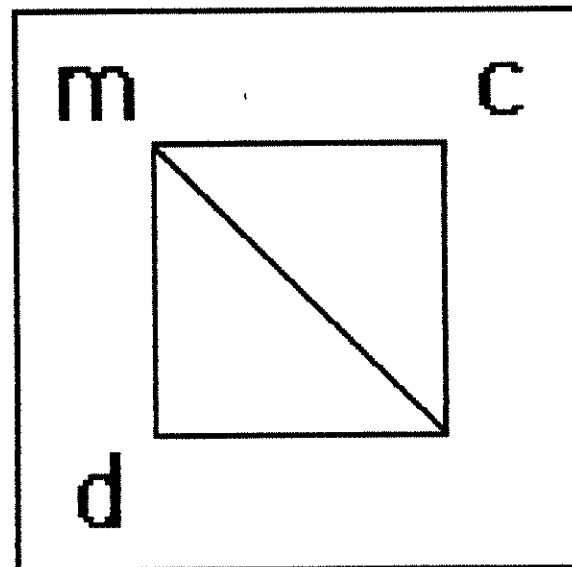
Children's strategies for DCM

- 1) constraint against backup
 - 2) 2 move search for goal state
 - 3) partial evaluation of progress toward goal
- = HILL-CLIMBING

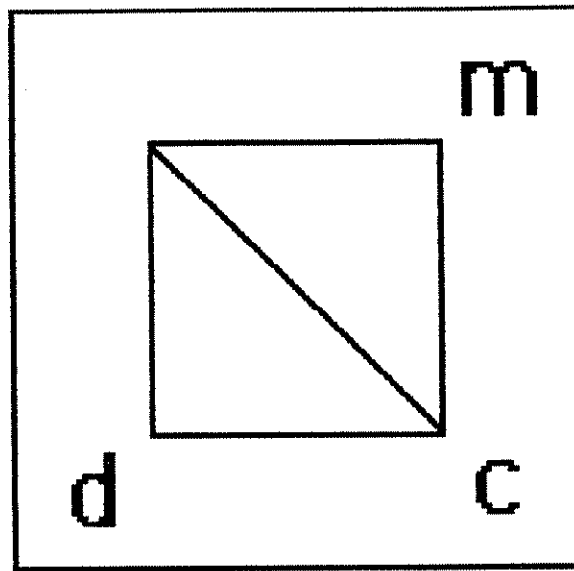
Initial
State



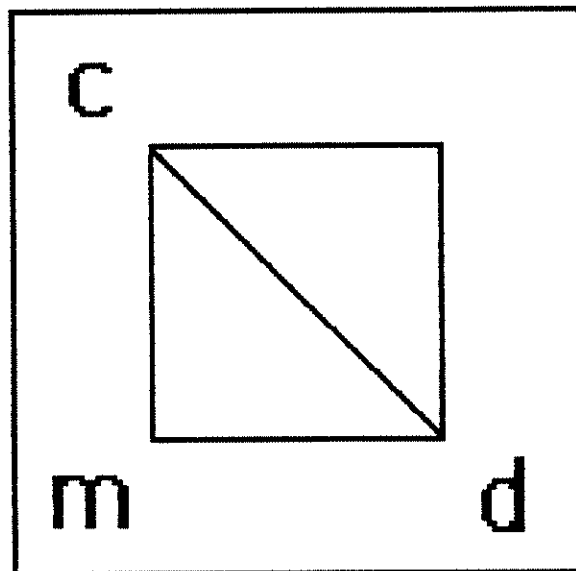
Goal
State



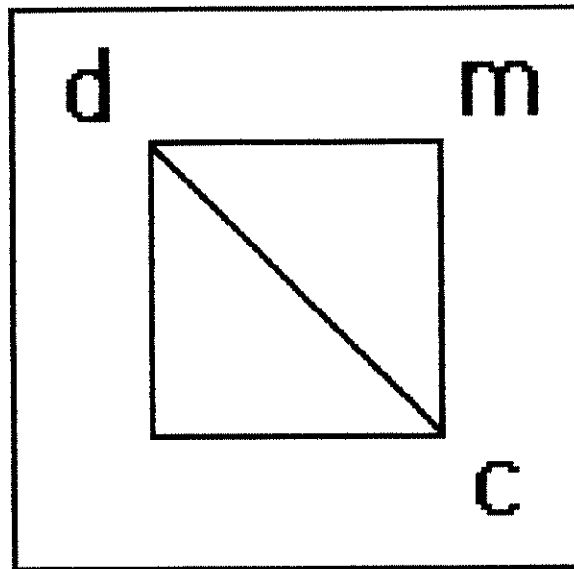
Initial
State



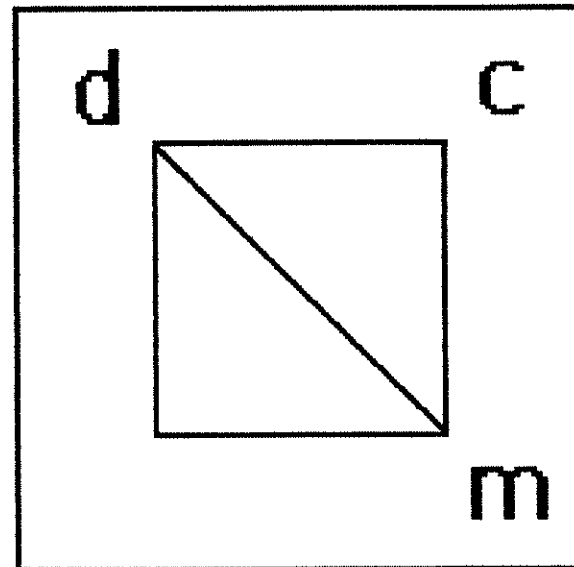
Goal
State

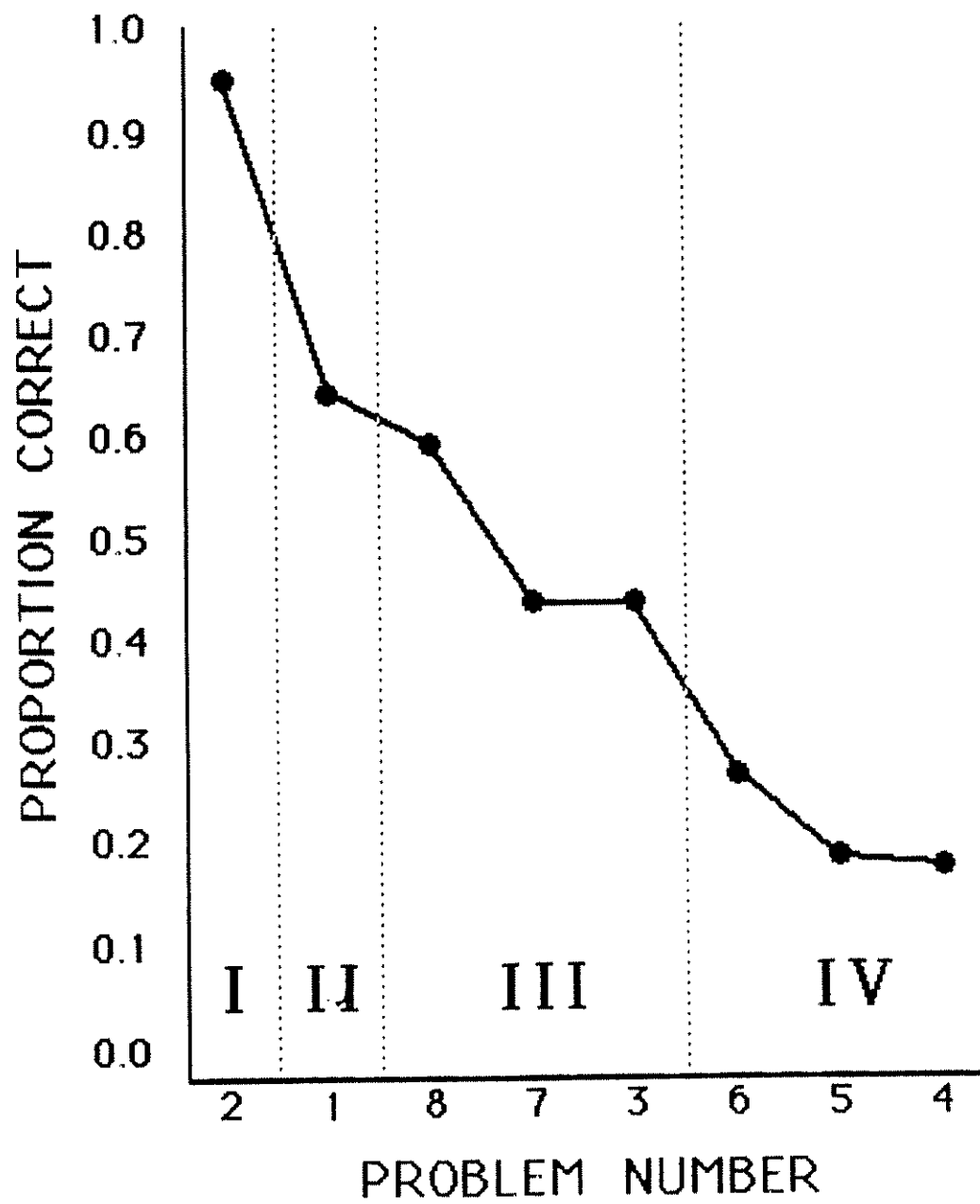


Initial
State



Goal
State



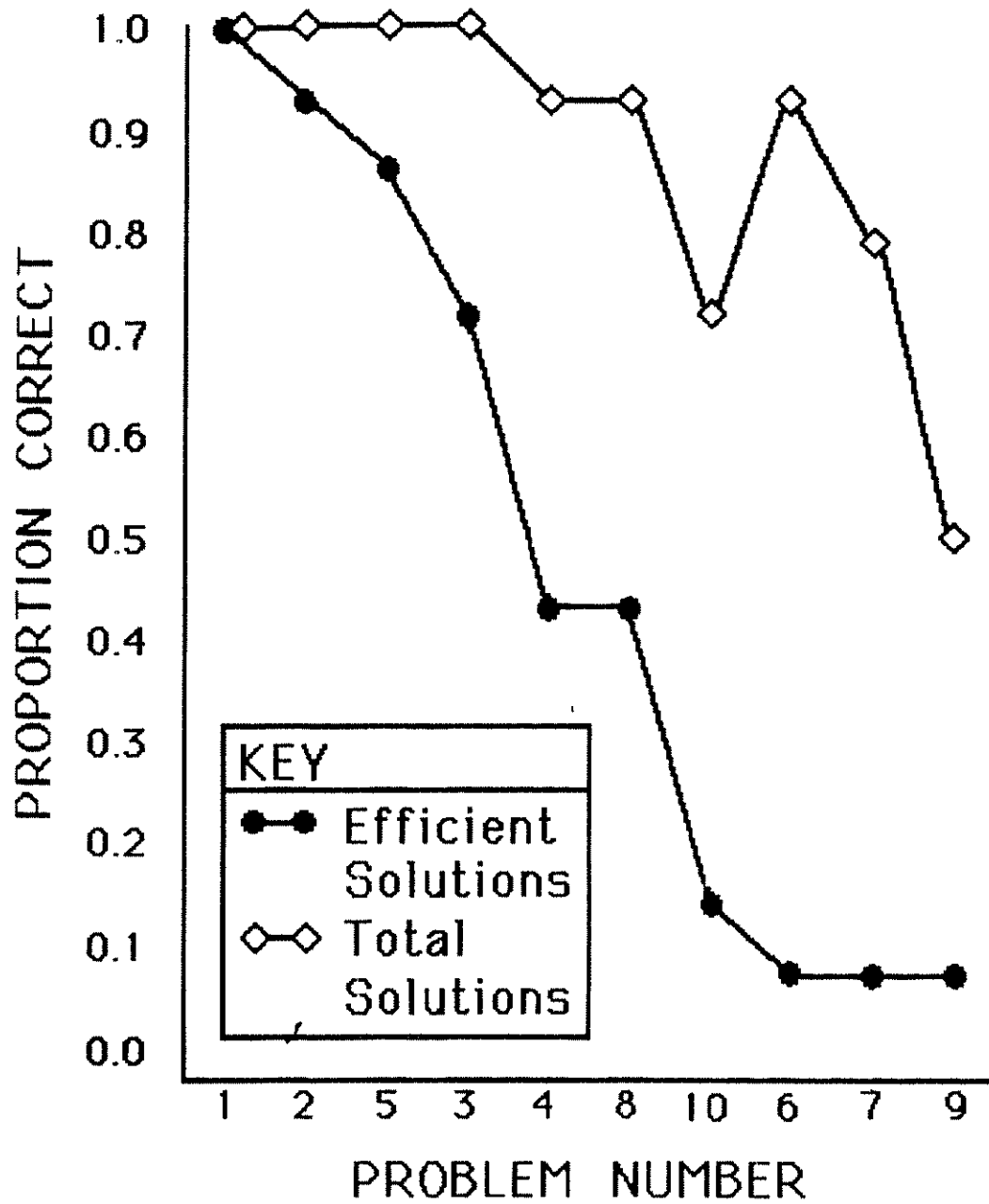


Initial State

A	A	B
A		C

Goal State

A	A	C
A		B

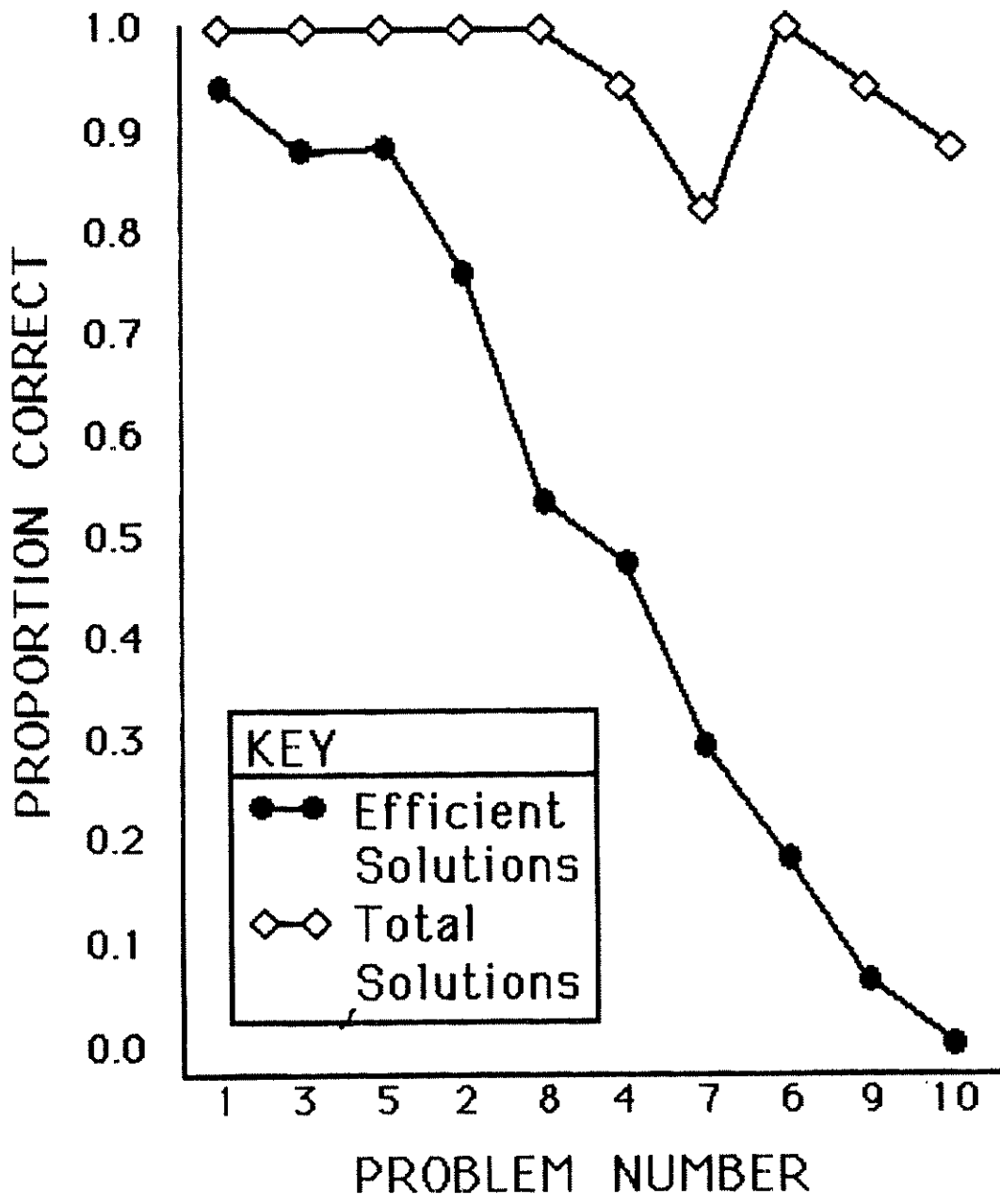


Piece-Position

- 1) local focus on individual pieces
- 2) ambiguous subgoal order
- 3) hill-climbing

Permutation-Position

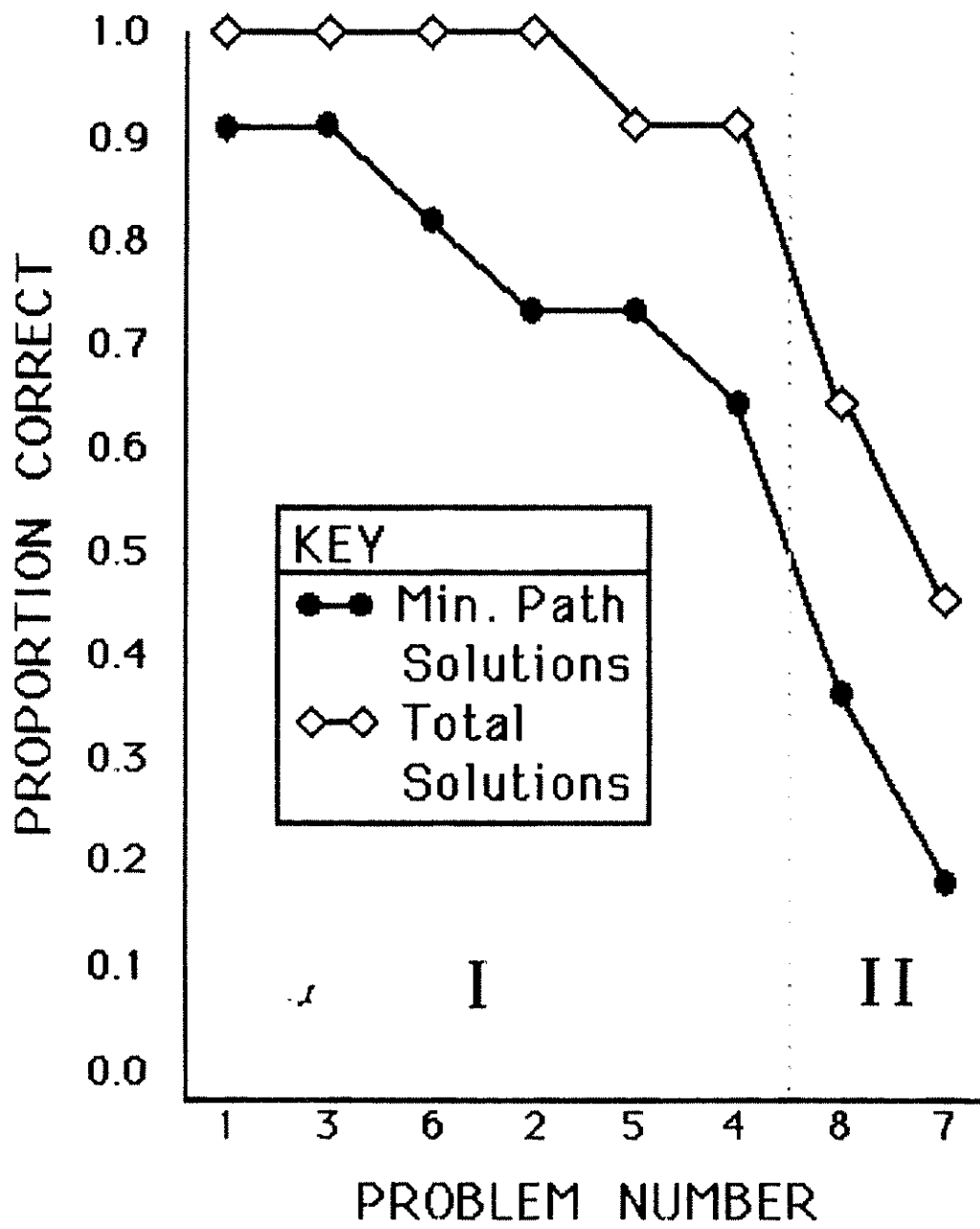
- 1) global focus on order of pieces
- 2) abstract subgoals with clear order
- 3) subgoaling



Increasing Difficulty



Random Search	Hill Climbing	Subgoaling
4 5 (25%)	4 4 5 5 6 6 (100%)	4 4 5 5 6 6 6* 6* (100%)
4 5 6 6 (12%)	6* (50%)	
6* 6* (6%)	6* (25%)	



Child Repertoire of Weak Methods

Means-Ends Analysis

if few subgoals

if clear subgoal ordering

Constraint Against Backup

2 Move Look-ahead for Goal

HILL-CLIMBING

Trial and Error

Adult Weak Methods

Flexible Representation

Local: use hill-climbing

Global: use subgoaling