Strategies in Analogous Planning Cases

Andrew S. Gordon (asgordon@us.ibm.com)
IBM TJ Watson Research Center, 30 Saw Mill River Road
Hawthorne, NY 10532 USA

Abstract

Strategies are abstract patterns of planning behavior that are easily recognized, compared, and used by people in everyday planning situations. In an effort to better understand them as a type of mental knowledge structure, three hundred and seventy-two strategies were identified from ten different planning domains, and each was represented in a preformal manner intended to describe the common characteristics of their instances. In doing this large-scale representation work, two observations were made with significance to current theories of analogical reasoning. First, strategies are portions of the relational structure shared by analogous planning cases. Second, representations of strategies include propositions about the reasoning processes of the agent employing them. We propose that a theoretical understanding of analogical reasoning allows us to use strategies as an investigative lens to view the mental models that people have of others and of themselves.

Introduction: Strategy Representations

The term *strategy* periodically appears in cognitive science literature to refer to the abstract patterns that can be recognized in planning behavior. People appear to have a near-effortless ability to use and reason about strategies despite their complexities - as when considering the case of a corporation that underprices part of its product line to bankrupt their competitors, or a case where a parent bird pretends to be wounded to lure a predator away from a place where their nest would be discovered. From the cognitivist's perspective, strategies can be viewed as knowledge schemas, with the assumption that these schemas are mental representations that can be manipulated, compared, and used in intelligent planning and problem-solving behavior. In order to understand this view more deeply, we undertook a project to systematically analyze and represent strategies on a large scale.

We began this project by identifying three hundred and seventy-two strategies from ten diverse domains of planning and problem-solving. Three competitive planning domains were examined: business, warfare, and dictatorship. Three cooperative planning domains were examined: scientific research, education, and personal relationships. Two individual performance domains were included: artistic performance and object counting. Finally, two anthropomorphic domains, where non-people are viewed as planners, were studied: immunology and animal behavior. Strategies for each of these domains were collected using a variety of methods, which included formal interviews with subject-matter experts, introspection, and the analysis of texts such as Niccolo Machiavelli's *The Prince* and Sun Tzu's *The Art of War*, which are nearly encyclopedic of strategies in the domains of dictatorship and warfare, respectively.

For each of these strategies, we authored a definition in the form of a representation such that all situations that match the definition would be positive examples of the strategy, and all cases that do not match the definition would not be examples of the strategy. Recognizing that the same strategy could be applicable in a wide variety of situations - even those that cross domain boundaries - our efforts focused on strategy representations that were of the highest possible level of abstraction while still meeting these definition requirements.

While some descriptive and functional planning languages are beginning to emerge in the artificial intelligence planning community (Tate, 1998; Gil & Blythe, 2000; McDermott, 2000), we chose not to attempt to use them for this representation work, following the belief that these current efforts are not yet expressive enough to describe the subtle planning features found in strategies. Instead, we adopted a style that can be best viewed as *preformal*, somewhat similar to the strategy representations found in smaller-scale efforts (Collins, 1986; Jones, 1992), and where the content of these representations was loosely drawn from a wide range of content theories of planning, notably from Owens (1990). The motivation for using this preformal style was to enable the scaling-up of representation work by relaxing the syntactic formality of logic while preserving the unambiguity of representational terms.

Figure 1 gives three examples of the 372 preformal representations that were authored. Words and phrases in the representations meant to refer to planning concepts and abstractions were capitalized and italicized as they were authored, which allowed us to algorithmically extract them so that they could be analyzed outside of the context of specific

Education strategy: Pop quiz. Ensure that students do their homework with the threat of a unannounced quiz

Representation: The planner has the goal that an agent achieve a set of *Knowledge goals* and has a *Plan* with a *Plan duration* that includes the *Agency* of the agent to execute *Subplans* that are not *Observed executions* by the planner. The planner *Envisions a threat* to the plan in that the agent do not execute the *Subplan* with cause that it is not a *Subplan* of a goal of the agent, and the planner *Envisions the possibility* that the *Successful execution of the plan* would not achieve the goal. The planner *Modifies the plan* by *Scheduling* a set of *Subplans* at *Start times* that are *Randomly selected* from *Moments* in the *Duration*. In each subplan the planner *Requests the execution* of a plan by the agent such that the *Successful execution* of the plan *Requires* that the agent *Executed the subplan* in the goal that were *Subplans previously scheduled*, and where a *Failed execution* will cause a *Violation of a goal* of the agent. The planner adds a *Planning constraint* against plans that cause the agent to have *Knowledge* of the *Scheduled start time* of these subplans.

Animal behavior strategy: Mark your territory. Leave scent marking to avoid unneccessary defensive conflicts

Representation: The planner has a *Competitive relationship* with a set of agents, and has a *Competitive plan* that includes the *Monitoring of execution* of plans of the other agents for *Locations of execution* that are *Locations* in a *Region*, and *In this case* the execution of an *Attack* on the agent. A *Threat* exists that the execution of the *Attack* will be a *Failed execution* with a cause of a *Successfully executed counterplan*. The planner envisions that *If it were the case* that agents in the set had *Knowledge* of the *Competitive plan* of the planner, then a subset of the agents would add the *Planning preference* against plans that had a *Location of execution* in the *Region*. The planner executes a *Repetitive plan* to *Encode information* that is the *Competitive plan* of the planner and *Transfer locations* of the *Encoding of information* to a *Location* that is in the *Region*, where the *Location* is *Selected from the set* with a *Selection criteria* of *Random choice*.

Counting strategy: Transfer between spaces. Count objects as they are moved into an empty location

Representation: The planner has the *Knowledge goal* of the *Quantity* of *Physical objects* in a set. There exists a set of two *Disjoint regions*, where every object has a *Location* that is *Contained within a region* that is a member of the set. The planner has a *Subplan* to *Transfer the location* of a specific object that is *Contained within a region* of the set to a different *Location* that is *Contained within the other region*. The planner executes a plan to achieve the goal that all of the objects in the set have a *Location* that is *Contained within the region* that is the *Start location* in the *Transfer of location* subplan. The planner then *Repetitively executes* a subplan where the planner executes the *Transfer of location* subplan and *Imagines a number*. In the *First repetition*, the number is 1, and in *Subsequent repetitions* the number is the addition of 1 to the *Imagined number* in the *Previous iteration*. The *Termination of repetition condition* is that the planner has an *Execution failure* of the subplan with a cause of *Unfound object in start location*. The planner then *Achieves* the *Knowledge goal* that is the *Imagined number* in the *Last repetition*.

Figure 1. Three preformal representations of strategies from different planning domains

representations. In all, 8,844 italicized words and phrases were extracted from the representations, which was reduced to a set of 974 terms by removing duplicate instances, selecting a representative lexical form for sets of instances that differed only in their inflection, and combining the forms that we determined to be synonymous, i.e. referring to the same planning concept.

The driving motivation behind this large-scale representation work was twofold. First, we aimed to identify the broad representational requirements of strategic planning and outline the mental models that people have of intentional agents. The findings in reference to this first motivation are reported in a separate publication (Gordon, 2001). Our second motivation, which is the subject of this current report, was to further our understanding of the peculiar role that strategies play in the way that people reason about analogous planning cases.

During and after the completion of this large-scale representation work, we made several observations that contribute to our theoretical understanding of strategies. In particular, two points are presented here that are specifically targeted at the cognitive science research area of analogical reasoning. First, we argue that strategies are themselves portions of the relational structure that serves as the basis for analogical reasoning about planning cases. Second, mental representations of strategies include references to the reasoning processes of intentional agents, providing us with a means of describing the models that people have of their own reasoning processes and those of others. Both of these arguments are developed in the following two sections.

Strategies are Relational Structures Shared by Analogous Planning Cases

In June of 1941, Germany invaded Ukraine with three million troops, threatening to advance eastward into Russia. Soon after, Soviet leader Joseph Stalin announced a scorched-earth policy for Ukraine, ordering that retreating Ukrainians destroy everything that could be of value to the advancing German army, including grains, fuel, and engines that could not be transported east to Russia. In an analogous case, Iraq invaded their oil-rich neighbor Kuwait in August of 1990, leading to the Persian Gulf war. The following January, the United States launches an attack against Iraq, and Saddam Hussein responded by blowing up Kuwaiti oil wells and dumping millions of gallons of oil into the Persian Gulf.

Analogies of exactly this sort have been the subject of a number of experimental studies of analogical reasoning (especially Spellman & Holyoak, 1992), and competing theories have been proposed as cognitive models for this sort of mental processing. The two theories that have received the most attention are Structure-mapping theory (Gentner, 1983, 1989) and Multiconstraint theory (Holyoak & Thagard, 1989, 1995). Although they have their differences, the two theories agree that analogical reasoning is based on structural similarity, the similarity of the systems of relationships that exist between the represented entities in two different cases. Both agree on the constraint of *one-to-one correspondence* between represented entities in analogous cases, e.g. Kuwait could potentially correspond to Germany, Ukraine, or Russia, but not more than one of these in any given analogical mapping. Both also agree on the constraint of the *systematicity*, requiring that sets of relationship correspondences have consistent argument correspondences, e.g. because Iraq is an argument in both the relationships of invading and destroying, both of these relationships cannot be a part of the same system of analogical mappings to the WWII case, where Germany did the invading and the Ukraine did the destroying.

Given the constraints of one-to-one correspondence and systematicity, these theories predict that the strength of any given analogy is strongly dependent on the way that the cases are represented. If we assume representations that are too sparse, we risk predicting that the analogy between a Persian Gulf war case and a WWII case would be a relatively weak one. If Germany is mapped to Iraq, then we have correspondence between the relationship of invading (Germany/Ukraine and Iraq/Kuwait) and the relationship of contained-within (Ukraine/Grains and Kuwait/oil wells). If the Soviet Union is instead mapped to Iraq, then we have a correspondence between the relationship of destroying (Soviet Union/grains and Iraq/oil wells) and that of possession or ownership (Soviet Union/Grains and Iraq/oil wells).

However, this analogy is intuitively very strong and unambiguous. The obvious mapping is that Iraq is like the Soviet Union, as their decision to destroy the Kuwaiti oil wells was analogous to when Stalin ordered the destruction of resources in the Ukraine. These cases are two examples of the *exact same strategy* - instances of an abstract pattern of planning behavior that is so prevalent in our culture that we've given it a name, *scorched earth policy,* so that we could refer to it again and again in analogous cases, whether they appear in warfare or in completely different domains such as politics or business. To account for the comparative strength over this interpretation of the analogy over others, we must assume that the representations of these cases are much richer.

When we consider the abstract similarities that are found in the planning that is done in these two cases, the structural alignment becomes clear. The agent that is doing the planning in these cases (Stalin/Hussein) has some adversarial relationship with some other agent (Hitler/Bush). This planning agent imagines a likely future where the adversary acquires possession of some resources (grain/oil) that are currently possessed by the planner. They imagine that the adversary will make use of these resources to further the pursuit of their adversarial plan (march on to Russia/control the middle east). They decide that the best plan is to do something (destroy grain/blow up oil wells) that will cause these resources to be destroyed, or to make it impossible that the adversary could make use of them, and to do so before the adversary gains possession.

While the rich relational alignment between these two military examples is described using natural language in the preceding paragraph, a corresponding mental representation language would necessarily include structures to refer to the adversarial relationship, the imagination of a likely future, the acquisition of resources, the expenditure of resources in an adversarial plan, the goal of disabling a resource, and the execution deadline. It is this collection of relationships that constitutes the representation of the strategy, and which also makes a significant contribution to judgments of analogical similarity to *every other case* that describes an instance of this sort of strategic behavior.

To clarify, we would like to point out that not all of the relational structure shared between analogous planning cases can be thought of as part of a strategy. Certainly there are analogous planning cases that are so not because of any similarity in the strategies of the participating agents. For example, a case where a person marries someone just before the other person

wins the lottery may be analogous to a corporation that acquires another business just before it has an unexpected licensing windfall, but the commonalties in these cases have more to do with unforeseen benefits than strategic thinking. In contrast, if it turns out that the person and the parent corporation both had selected their candidate acquisitions based on a perception of how lucky the candidates were, we would say that they shared the same strategy - in reference to the portion of the shared relational structure that concerned the planning processes of these agents.

The research opportunity that is evident here concerns the apparent ease that people have in making casual references to large portions of shared relational structure, removing these portions from their context to be considered independently, and assigning to them names like *scorched earth policy* when they are particularly interesting for one reason or another. This ease with strategies enables researchers to collect whole catalogs of naturally occurring analogical mappings, and to argue about how they could be represented on a much larger scale than was possible in previous knowledge representation debates.

Strategy Representations Include the Reasoning Processes of Agents

The formal case representations that have appeared in computer models of analogical reason consist almost entirely of propositions about the external world (Falkenhainer *et al*., 1989; Forbus *et al*., 1994; Holyoak & Thagard, 1989; Holyoak *et al*., 1994). The most compelling examples that support their corresponding theories of analogical reasoning are often in the domain of physical systems, where existing representational theories support the way that these example cases were represented, notably Qualitative Process Theory (Forbus, 1984). Likewise, the case representations in the computer models of analogical reasoning that seem the most contrived are those that are more story-like in nature, involving the intentional behavior of intelligent agents.

A somewhat notorious example of this problem can be seen in the "Karla the hawk" stories that were used in support of the Structure-Mapping Engine (Falkenhainer *et al*., 1989). The main story concerned the actions of Karla, an old hawk, and her encounter with a hunter who had attempted to shoot Karla with an arrow, but had missed because the hunter's crude arrow lacked the feathers needed to fly straight. Karla knew that the hunter wanted her feathers so she offered some of hers to him, and he gratefully pledged never to shoot at a hawk again. In the dozen propositions that were authored to represent this story, emphasis was placed on describing the actions that where done (e.g. seeing, attacking, offering, obtaining, promising) with only a single, highly ambiguous predicate of *realizing* to refer to the reasoning processes that Karla undertook.

What is lacking from these representations is the richness of planning and reasoning that is ascribed to these characters when we read stories like this - the conceptual glue that allows us to make sense of the story in the first place. Much of this knowledge can be packaged into the form of a single strategy, the one that we guess that Karla had in mind when she offered the feathers. Figure 2 offers a preformal representation of this strategy, where the capitalized and italicized words and phrases come from the set used in our large-scale strategy representation project. Of course, it is possible to imagine that Karla wasn't thinking strategically at all - often characters in this genre of fable-like stories seem to stumble across some course of action by mere chance, without thinking things through. However, it is difficult to imagine that a reader could be as ignorant. Indeed, it is the planning lessons that can be learned from stories of this genre that make them compelling and valuable.

What should be noticed in Figure 2 is that the strategy

Karla's Strategy: Turn enemies into friends by improving their capabilities
The planner has an *Adversarial Relationship* with another agent that has had an *Execution failure* of an *Adversarial plan*, with a *Cause of failure* of *Instrument failure*. The planner *Envisions* that this agent will *Attempt the execution* of this *Adversarial plan* in *Future states* against the planner and *Envisions a possibility* that this plan will be *Successfully executed.* The planner has a *Partial plan* to *Reduce the probability* of *Instrument failure* in *Future executions* of the *Adversarial plan* by this other agent that includes *Modifying the instrument* and includes the *Instrumental use* of *Resources* of the planner. The planner executes a plan to *Make an offer* to the agent where the *Offered action* is that the planner *Transfers possession* of *Resources* that *Enable* the agent to execute the *Partial plan*. The planner then *Adds the expectation* that *If it is the case* that the agent *Accepts the offer*, that the agent will *Terminate their role* in the *Adversarial relationship*, which will cause them to *Abandon plans* that are *Adversarial plans* against the planner. The planner then *Adds a threat* that the *Expectation* is a *False expectation*, and that the agent will execute the *Partial plan* followed by the execution of the *Adversarial plan*.

Figure 2. The strategy of Karla the hawk as a preformal representation

contains a significant amount of references to mental states of both the planner and the adversary. There is the imagining of future states, of a partial plan not entirely worked out, an expectation of the consequences of an action, and an explicit threat of what might happen if this expectation is wrong. Each of these mental states is critical to the understanding of the story, and we argue that they should be included in the representation of this case to explain analogies to cases where only the strategy is shared.

Just as analogies in physical systems are based on mental models of processes in physical domains, analogies in intentional domains include assertions that only make sense with respect to a mental model of agents and intentional behavior. In arguing for the inclusion of these sorts of assertions in case representations in intentional domains, we are also making the argument that people have a rich model of agents and their reasoning processes.

A great deal of attention has been directed toward developing models of agents and intentional behavior among decision theorists and artificial intelligence logicians, often centered around the notion of Belief, Desire and Intention (BDI) agents. Formalizations of these models (e.g. Cohen and Levesque, 1990) typically strive to maximize the number of simplifying assumptions in order to retain the ability to prove related theorems, but to do so without sacrificing the expressiveness required to compute useful functions. The engineering value of this approach is demonstrated when theories lead to practical applications (see Rao & Georgeff, 1995), but we caution against this approach for the purpose of cognitive modeling. Certainly there are times where the mental representations that people have may appear to exhibit the qualities of elegance and simplicity, but our aim should be to describe the mental models of people as they are - without simplifying assumptions - if we are to understand and predict how they are manipulated by cognitive processes.

In our own investigation of mental representation through the lens of strategies, we have found that the models that people have of other agents and their reasoning processes is enormously complex in comparison to previous formalizations. From the most generalized perspective, the model appears to be comprised of many of the components that are commonly proposed. These components include representations of the current state, agents and their goals, the plans that are held by these agents, the envisionments that these agents construct of past and future states, of the plan construction process, the making of decisions, the scheduling of behaviors, the monitoring of events, and the process of executing an intended plan in the world. Given a closer look, we find that each of these model components is extremely complex. For example, the representations that people have of the process of executing an intended plan are rich enough to refer to the sensations that agents experience during an execution, and to remark on whether it had a natural quality to it or not - as in reference to a concert pianist that finds a particular passage in a piece cumbersome due to an awkward fingering that they selected. Indeed, a strategy that concert pianists employ is to reduce the risk of performance blunders by explicitly identify the sections of a musical piece that give rise to these sensations of awkwardness, and to rework their plan of execution for these sections so that they have a more natural quality.

Conclusions: Analogy as a Tool

"The association of solved problems with those unsolved may throw new light on our difficulties by suggesting new ideas. It is easy to find a superficial analogy which really expresses nothing. But to discover some essential common features, hidden beneath a surface of external differences, to form, on this basis, a new successful theory, is important creative work." (Einstein & Infeld, 1938)

Einstein and Infeld presented this idea to justify making an analogy between a violin string and the wavelike properties of elementary particles. The essential common features that Einstein and Infeld discovered were the concepts of a *standing wave* and a *moving wave*, and the new successful theory that they were forming was that of probability waves in quantum physics.

Einstein and Infeld's quote reveals something about the way that they approached scientific problem solving, but from a cognitive science perspective, we might also view it as a proposed methodology for forming theories of people's mental models. That is, by analyzing analogous cases, discovering the essential common features that are hidden beneath a surface of external differences, we can form theories of the mental models that people have that cause these cases to be analogous.

In this paper we have argued that planning strategies are particularly appropriate as the subject of analysis using the tools of analogy. Strategic analogies reveal features of our mental models of human planning, and in doing so, challenge our cognitive theories and models of intelligent planning and problem solving behavior. It is this hope that led us to author three hundred and seventy two representations of strategies in ten different planning domains, where each representation attempted to define the features of the planning situation that were common among all analogous instances of the strategy.

Einstein and Infeld made reference to two important activities in the quoted text. The first, to discover

essential common features of analogous cases, has been accomplished on a large scale for instances of strategies. Achieving the second, to form a new successful theory, will require a substantial amount of additional work. This paper has argued for the inclusion of two claims in the future successful theories that are developed. First, we argued that strategies are themselves portions of the relational structure that is the basis of similarity between analogous planning cases. Second, the representations of these strategies include propositions about the reasoning processes of the agents employing them, giving researchers a investigative lens to examine the rich mental models that people have of others and of themselves.

References

Cohen, P. & Levesque, H. (1990) Intention is choice with commitment. *Artificial Intelligence*, 42(3).

Collins, G. (1986) Plan creation: Using strategies as blueprints. Ph.D. dissertation, Yale University.

Einstein, A. & Infeld, L. (1938) The evolution of physics from early concepts to relativity and quanta. New York: Simon & Schuster, p. 273.

Falkenhainer, B., Forbus, K., & Gentner, D. (1989). The Structure-Mapping Engine: Algorithm and Examples. *Artificial Intelligence*, 41, pp 1-63.

Forbus, K. D., Ferguson, R. W., & Gentner, D. (1994). Incremental structure-mapping. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 313-318. Hillsdale, NJ: Lawrence Erlbaum Associates.

Forbus, K. (1984). Qualitative process theory. *Artificial Intelligence*, 24, pp 85-168.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*, pp. 199-241. London: Cambridge University Press.

Gil, Y. & Blythe, J. (2000) PLANET: A sharable and reusable ontology for representing plans. *AAAI 2000 workshop on Representational issues for real-world planning systems*.

Gordon, A. (2001) The representational requirements of strategic planning. *Proceedings of the Fifth Symposium on Logical Formalizations of Commonsense Reasoning*.

Holyoak, K. & Thagard, P. (1995) *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Holyoak, K. & Thagard, P. (1989) Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295-355.

Holyoak, K., Novick, L, & Melz, E. (1994) Component processes in analogical transfer: Mapping, pattern completion, and adaptation. In K. Holyoak & J. Barnden (Eds.), *Advances in connectionist and neural computation theory, Vol. 2: Analogical connections*. (pp. 113-180). Norwood, NJ: Ablex.

Jones, E. (1992) The flexible use of abstract knowledge in planning. Ph.D. dissertation. Yale University.

McDermott, D. (2000) The 1998 AI Planning Systems Competition. *AI Magazine*, 21(2), 35-55.

Owens, C. (1990) Indexing and retrieving abstract planning knowledge. Ph.D. dissertation. Yale University.

Roa, A. & Georgeff, M. (1995) BDI agents: From theory to practice. *Proceedings of the First International Conference on Multiagent Systems*.

Spellman, B. & Holyoak, K. (1992) If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology*, 62, 913-933.

Tate, A. (1998) Roots of SPAR - Shared planning and activity representation. *Knowledge Engineering Review*, 13, 121-128.