The Representational Requirements of Strategic Planning

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

Abstract

An analysis of strategies, recognizable abstract patterns of planned behavior, reveals an enormous divide between the kind of planning that artificial intelligence planning systems do and the kind of strategic planning that people do. This paper describes a project to collect and represent strategies on a large scale to identify the representational requirements of strategic planning in order to inform the design of future artificial intelligence planning systems. Three hundred and seventy-two strategies were collected from ten different planning domains. Each was represented at an abstract level and in a preformal manner designed to reveal the planning concepts that each strategy contains. The contents of these representations, consisting of nearly one thousand unique planning concepts, were then collected and organized into forty-eight groups that outline the representational and functional components of strategic planning systems.

1 Strategic Planning

Throughout its history, research on artificial intelligence planning systems has weathered criticism concerning the viability of these systems to solve real world problems. Persistence in this area has yielded algorithms that are both more scalable and more flexible, and there is growing consensus among researchers about what constitutes a real planning problem and how they should be solved [McDermott, 2000]. While this consensus has afforded new metrics for comparison between approaches, it has also had the unfortunate side-effect of underscoring the enormous divide between the kind of planning that computers do and the kind of planning that people do.

To illustrate this divide, consider the planning that was done by a budding concert pianist with whom we recently spoke. During one of his early competitive performances his execution of a solo piano piece was derailed by an audience full of distractions, which included audible conversation, chair noises, and even the ringing of a portable telephone. Determined to avoid this problem in the future, he would purposefully cause distractions to occur during his practice sessions. By setting alarm clocks for odd times, turning on kitchen oven timers, and asking some friends to telephone him during certain hours, he became accustomed to surprise events during his performances, and learned how to ignore them.

It would be difficult to argue that the reasoning done by this pianist was not planning - it is indeed an example of the use of a more abstract strategy for learning to ignore distractions during execution. This same strategy might also be useful for surgeons and soldiers, for whom distractions can be deadly. What is difficult to imagine, however, is how this instance could be contorted to somehow fit into the current models of artificial intelligence planning systems. While some initial conditions, operators, and effects are a part of this planning problem, at its core are more subtle planning concepts, including *execution failures*, *distracting events*, the *ability to ignore*, and *learning through practice*.

This example warns against models of intelligent planning that are too narrowly construed, but also suggests that strategies, the recognizable abstract patterns of planned behavior, may be useful in revealing the breadth of planning concepts that have a role in the kind of planning that people do.

This paper describes a project to exploit the conceptual richness of strategies to identify the representational and functional components of intelligent planning. Our approach was to represent strategies on a large scale - hundreds of strategies from many different planning domains - and in so doing, progressively generate a vocabulary of planning that is broad enough to describe the richness and subtleties that strategies contain. Our goal was to describe the representational requirements of *strategic planning*, a term that we use to refer both to the use of strategies in planning systems, as well as other modalities of planning that use the same representational and functional components.

This project was composed of three parts, each described in the following three sections of this paper. In the first part of this project (section 2), we collected a set of 372 strategies from 10 different planning domains. In the second part (section 3), we authored an abstract representation for each of the strategies in the set, defining the features that are common in all instances of the strategy. In the third part

(section 4), we extracted all of the defining features out of the set of strategy representations, and attempted to organize and summarize them to learn something about the nature of strategic planning systems.

2 Collecting Strategy Examples

With the intuition that the analysis of strategies could reveal the representational requirements of the sort of planning that people do, our first challenge was to collect a broad range of strategies from many different planning domains to feed the analysis process. Although the concept of a strategy has been discussed at least since classical times, we were unable to find an existing encyclopedic collection of strategies across various domains. In particular planning domains, such as warfare and governance, works such as Sun Tzu's *The Art of War* and Nicollo Machiavelli's *The Prince* are almost encyclopedic in nature, and the strategies in each have become widespread in the world's literate cultures. In other planning domains, such as business planning or musical performance, the common strategies seem to be more loosely shared among practitioners, and may vary more widely between cultures. Rather than attempting to perform an exhaustive search, or even to collect a representative random sample, we chose instead to collect strategies from the encyclopedic texts where they were available, from interviews with colleagues that had particular domain knowledge, and introspectively in our own areas of expertise.

In all, we collected 372 different strategies from ten different planning domains, which are outlined in Figure 1. These ten planning domains include several that were chosen specifically because of some interesting characteristic of the strategies in the domain. The strategies of business planning, for example, often reflect the view that organizations of people are agents in their own right, with plans, goals, and even values. The strategies that people recognize in animal behavior are anthropomorphic, as modern zoologists prefer to explain animal behavior as the result of natural selection rather than intentional planning. The extreme anthropomorphic view is reflected in the strategies of immunology, where immune system cells, viruses, and bacteria are seen as adversaries at war. The performance strategies that musicians use reveal much about the way that people manage skills that are largely internalized. The counting domain is unique in this list in that people are typically unaware that they use strategies to count objects; these strategies had to be derived through observation of human counting behavior.

3 Representing Abstract Strategies

After collecting a wide range of strategies, we began the work of analyzing the strategies to uncover the representational requirements that each of them contained.

1. **Animal behavior** (37 strategies). Anthropomorphic strategies of wild animals. Example: *Cowbirds abandon their eggs in the nests of other bird species who will raise the hatchlings along with their own offsprings*.
2. **Business** (62 strategies). The strategies of corporations and their executives. Example: *Companies that compete for repeat business from consumers gain customer loyalty by issuing redeemable points (e.g. frequent flyer miles) with each purchase*.
3. **Counting** (20 strategies). The internalized strategies that people use to count things. Example: *When counting identical things organized in a circle, people will mark their starting location to avoid over-counting* .
4. **Education** (32 strategies). The strategies of students and educators. Example: *Occasionally delivering an unscheduled "pop quiz" encourages students to study materials in the duration between larger exams* .
5. **Governance** (60 strategies). The strategies of Machiavelli's *The Prince*. Example: *If established as lord by nobility, a new prince should work to become popular to establish a power base among the citizenry* .
6. **Immunology** (21 strategies). Anthropomorphic strategies of immune system cells and the foreign organisms that they protect against. Example: *One way that cells guard against the spread of viruses is to destroy themselves through programmed cell death when an infection is detected*.
7. **Military** (45 strategies). The strategies of Sun Tzu's *The Art of War*. Example: *By leaving the enemy a means of retreat, the attacking army prevents the threat of extraordinary defenses enabled by heroic desperation* .
8. **Performance** (39 strategies). The strategies musicians use to manage their performance abilities. Example: *To avoid being derailed by distracting events during a performance, musicians will purposefully cause surprising distractions to occur during their practice sessions in order to become accustomed to them*.
9. **Relationships** (22 strategies). Strategies that people use to find, keep, and leave a loving partner. Example: *On a first date, a skeptical person may ask a friend to telephone halfway through to give them a chance to fabricate an emergency that requires them to leave*.
10. **Scientific research** (34 strategies). The strategies that scientists use to advance their field and their careers. Example: *To moderate the difficulty of obtaining research funds, many researchers tie their work to some application domain that is backed by wealthy interests, such as the military or pharmaceutical companies* .

Figure 1. Ten strategic planning domains with example strategies (372 strategies total)

The aim of this strategy representation effort was to define a given strategy 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.

In defining the abstract structure of strategies we have, in effect, explicitly described the structural similarity between analogous situations of strategy use. Structural similarity of case representations has long been regarded as the basis of analogical reasoning [Gentner, 1983]. However, computer simulations of analogical reasoning have placed a heavy burden on the development and use of rich representational ontologies [Falkenhainer *et al*, 1989]. As a result, the most compelling examples of analogical reasoning in computer systems are in domains where rich semantic theories exist, notably physics.

In the field of AI planning, some formal representational ontologies exist [Tate, 1998; Gil & Blythe, 2000], but are not, in our experience, expressive enough to adequately represent the strategies identified in this project. Instead, we have taken the approach of authoring representations that might best be described as preformal, somewhat similar in format to the work of previous researchers in this area

Governance Strategy 48. Foster then crush animosity to improve renown: "For this reason many consider that a wise prince, when he has the opportunity, ought with craft to foster some animosity against himself, so that, having crushed it, his renown may rise higher." (from Machiavelli's *The Prince*)

Representation: The planner has a *Role* in a *Power relationship* over a *Community* and has the goal that a set of agents *Believe* that the planner is the *Best candidate* for this role. There exists an agent that has goals that are in *Goal conflict* with the planner, and that has had a *Planning failure* to achieve their goals with a *Cause* of *Lack of resources* or *Lack of agency*. The planner plans and executes for the goal that the agent *Believes* that there exists an *Opportunity* for this agent to achieve their goals, where no *Opportunity* exists. The planner then *Monitors* the agent for the *Start of execution* of plans to achieve the goals. When this occurs, the planner executes a plan to *Destroy* the agent, with the goal that the first set of agents have *Event knowledge* of the planner's execution.

Counting strategy 8. Transfer into a container: Determine the quantity of a set of things by moving them one at a time into an empty container.

Representation: The planner has the *Knowledge goal* of the *Quantity* of members in a set of *Physical objects*. The *Locations* of the *Physical objects* are *Contained within a region*. The planner has possession of a *Physical container* where the *Contents* of the container is a set of *Physical objects* that does not include members of the set of unknown quantity. The planner has a subplan to *Transfer the location* of the physical objects of unknown quantity from the *Region* to a *Location* such that they are members of the set of *Contents* of the container. The planner *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*.

Animal behavior strategy 15. Stay close to the fort: Prairie dogs stay close to their burrows to guard against a quick attack by a bird-of-prey.

Representation: The planner has an *Adversarial relationship* with a set of agents that have an *Adversarial plan* that includes *Transferring locations* to the *Location* of the planner to *Achieve a precondition* of having the *Location* of the planner. The planner has a *Counterplan* that has the *Precondition* that the planner has a *Location* that is in a *Region*, where the planner has a *Planning failure* in *Other agent planning* for an *Arbitrary member of the set* of adversaries to *Achieve the precondition* that the agent has a *Location* in the *Region*. The planner has *Envisioned future plans* that include having a *Location* that is not in the *Region* for a *Duration*, and *Envisions a threat* that an agent in the set of adversaries will *Attempt the execution* of the *Adversarial plan* in this *Duration*, and that this would be a *Successful execution*, and where there is a *Duration* between the *Moment* that the planner *Observes the execution* of the *Adversarial plan* and the *End time* of the *Transferring locations* by the agent. The planner executes a plan to *Imagine a region* where *Locations* in the region have a *Condition* that the planner has an *Envisioned likelihood of success* of a *Transfer of locations* between the *Location* in the *Region* an *Arbitrary location in the region* of the *Counterplan* with an *Execution constraint* that the *Execution duration* is less than the *Duration* in the *Threat*. The planner then adds the *Planning constraint* against *Future plans* that include *Locations* that are not in the *Imagined region*.

Figure 2. Three examples of preformal representations of abstract strategies

[Collins, 1987], and where the content of these representations is loosely drawn from a wide range of content theories of planning (notably Owens [1990]). The use of this preformal style was intended to enable the scaling-up of representation work by relaxing the syntactic formality of logic while preserving the unambiguity of representational terms.

Figure 2 gives three examples of the 372 preformal representations that were authored. These ordered examples were authored early, in the middle, and late in the process, and reflect a progressing style where the strategy representations are increasingly more verbose. Words and phrases in the representations that referred to planning elements were capitalized and italicized as they were authored to aid in extracting them for later analysis.

4 Organizing Representation Terms

After representing the 372 strategies, we turned our attention to the set of planning concepts that were used to define these strategies. The capitalized and italicized words and phrases were extracted from the representations, resulting in a list of 8,844 instances. A controlled list of terms was then created by removing duplicate instances, selecting a representative lexical form for sets of instances that differed only in their inflection, and combining the forms that were determined to be synonymous, i.e. referring to the same planning concept. The result of this vocabulary-creation effort was a list of 974 terms with enormously broad scope.

An analysis of the use of these terms over the set of strategy representations showed a wide variance in the frequency that a term appeared, where the ordered frequency appears as a standard Zipf distribution. The rate of growth of new terms decreased as new strategy representation were authored, but remained positive at the end of our representation work.

During our initial review of the terms, it became clear that they could be easily grouped into sets of related terms, where some sets consisted of terms that were instances of a common abstraction while others were directly related to a particular functionality of the planning process. Fortunately, but not surprisingly, these sets of terms appear to describe a coherent model, consisting of classes of planning knowledge coupled with component subsystems that have a role in the kind of planning that people do. Accordingly, these sets of terms outline the representational requirements of strategic planning and provide a design model for future artificial intelligence planning systems that have a more human character. In all, 48 sets of related terms were identified, each of which is summarized here.

1. **World states** (19 terms): The world is described in states that change, are partially observable, and where portions of these descriptions delineate environments.

2. **Time** (24): Moments of time, including the current moment, compose durations that have relationships to other moments and durations, and which may be only partially defined.

3. **Space** (42): Space is composed of locations in regions with boundary definitions, and can define paths through space with directionality and endpoints.

4. **Values and quantities** (23): Qualitative values and their quantifications may be defined by their ranges or in relationship to others, and are used to describe amounts in the world.

5. **Events** (10): Events are interpretations of states changing over time, may occur multiple times with some frequency, and have relations to other events.

6. **Classes and instances** (15): Things in the world are conceptualized as instances of class categories at various levels of abstraction, with characteristics that sometimes must be guessed.

7. **Sets** (34): Things can be conceptualized as being members of sets that have membership rules, and which may be ordered using an ordering principle.

8. **Physical entities** (29): Physical objects and substances, including the physical bodies of agents, are composed of components, can be configured into a state, and can contain or be attached to other things.

9. **Agents** (19): Agents are intentional things having characteristics and roles, and are sometimes unknown, unacquainted, or unspecified to another agent.

10. **Goals** (24): Goals of agents describe world states and events, and include the goals in agent relationships, auxiliary goals, preservation goals, knowledge goals, envisionment goals, planning goals, and scheduling goals.

11. **Goal themes** (6): The existence of goals in agents is explained by goal themes, which reference the various roles that an agent holds or to characterizations of being good or evil.

12. **Agent relationships** (26): Relationships between agents are determined by their comparative goals and the plans that they are executing, and include adversarial, competitive, and assistive relationships.

13. **Communities and organizations** (10): Sets of agents may be described as organizations and communities, where agents have roles in a structure, and where processes and goals of the set agents can be referenced.

14. **Emotions** (7): Agents understand a wide variety of types of emotions, such as pride and sympathy, the effects of which are important in plans and strategies.

15. **Plans** (33): The plans of agents are descriptions of behaviors envisioned to achieve certain goals, and can be defined by the types of goals that they achieve or by the degree to which they are reused by one or many agents, and may be composed and partial.

16. **Plan elements** (28): Plans are composed from subplans, may branch if certain conditions are met, may have iterative or repetitive components, and may include

preconditions, triggering conditions, and elements that specified at the time of execution.

17. **Resources** (16): Physical entities that are widely reusable across multiple plans can be characterized as resources of a certain amount, and as such may be acquired, generated, expended, transferred, and maintained.

18. **Ability** (9): Ability is a characteristic of an agent that is predictive of the successfulness of their execution of a shared plan, and may be qualified at various degrees of skill.

19. **Activities** (12): Activities are shared plans where agents have roles with scripted expectations, and some are normally part of an agent's execution.

20. **Communication** (22): Some abstract plans deal with transferring information to other agents, and include making and accepting offers, asking and answering questions, requesting permission, persuasion, negotiation and threatening another agent.

21. **Information** (12): Some types of knowledge can be externalized as information, sometimes encoded and decoded in some external physical entity, which can be passed in signals and that may be either true or false.

22. **Agent interaction** (23): Several persistent abstract patterns of events involve the interaction of agents, such as the assignment of ownership, a defense against an attack, and the execution of work.

23. **Physical interaction** (11): Some persistent patterns of events describe the interaction that agents have with physical entities in the world, such as configuring a thing or using an instrument for its design purpose.

24. **Managing knowledge** (28): Agents refer to knowledge that they have in the form of beliefs that may be shared, may be true or false, that may be justified, and used for prediction, and where knowledge can be actively assumed, affirmed, or even disregarded.

25. **Similarity comparisons** (16): Agents can compare the similarity or analogous nature of things that they reason about, from conceptualizations of things to representations of events, and where various similarity metrics can be used.

26. **Memory retrieval** (3): Agents have a memory that they use to store information through memorization and retrieve from using memory cues.

27. **Explanations** (16): Agents generate candidate explanations for unknown causes of all sorts through a process that is modified with the addition and removal of explanation preferences.

28. **World envisionments** (46): Agents follow preferences and constraints to construct causal envisionments to make prediction about the past, present, and future, where these envisionments include likelihoods with dependencies, possible branches, and hypothetical entities.

29. **Execution envisionment** (23): One mode of envisionment is that of imagining the execution of a plan for the purpose of predicting possible conflicts, execution failures, side effects, and the likelihood of successful execution.

30. **Causes of failure** (31): Agents attempt to explain failures as examples of explanation patterns, such as scheduling failures caused by a lack of time or planning failures caused by a lack of resources.

31. **Managing expectations** (8): Agents manage a set of expectations that are the result of envisionment and that are validated or violated over time, where violation adds the goal of explanation.

32. **Other agent reasoning** (7): To predict the actions of other agents, agents envision what they would be thinking if they were them.

33. **Managing threats** (13): Agents envision threats that are events or states that violate goals or cause plans to fail, and which may be merely possible or already realized.

34. **Managing goals** (27): Agents manage their goals through preferences and by assessing their justifications, which may cause them to suspend, add, modify, specify, abandon, or pursue a goal from their current set.

35. **Planning modalities** (17): Agents plan to achieve their goals constructing new plans, by adapting old plans, or by considering the plans of others, and make reference to past, future, and normal planning.

36. **Planning goals** (28): Agents direct their planning using planning goals, such as the goals of blocking a threat, delaying an event, enabling an action, preserving a precondition, or satisfying the goals of others.

37. **Plan construction** (30): Agents construct new plans through the specification of partial plans, adding and ordering subplans, and resolving the various planning problems that arise.

38. **Plan adaptation** (18): Existing plans can be adapted and modified by substituting values or agency, and by adding or removing subplans, all in a process that can fail and that has an adaptation cost.

39. **Design** (8): Agents generate designs for things, processes, and information that does not yet exist, and that can satisfy certain design constraints, and where the actual things have a degree of adherence to the design ideal.

40. **Decisions** (38): Agents make decisions across planning, envisionment, scheduling, and execution, where there is a choice set, a choice criteria with decision factors, consequences, preferences, random selections, and sometimes the determination that decisions are insignificant.

41. **Scheduling** (23): Unscheduled or pending plans are scheduled and unscheduled for execution in a process managed by modifiable scheduling constraints, preferences, and deadlines for starting or completing plans.

42. **Monitoring** (18): Agents can purposefully monitor both states and events in the world as well as their own internal planning processes for certain trigger conditions, causing the execution of some triggered plans.

43. **Execution** (21): Plans can be executed concurrently, consecutively, repetitively, periodically, in a manner coordinated with other events, or in a manner constrained by execution rules.

44. **Plan following** (20): Agents track of the progress of a plan during its execution, noting the achievement or violation of preconditions, missed deadlines, and the success or failure of plans and their component parts.

45. **Execution control** (27): Plans can be attempted, suspended, resumed, abandoned, delayed, failed, or completed successfully.

46. **Repetitive execution** (14): Repetitive or iterative plans or subplans have previous, current, and remaining executions of certain quantities, and continue until they are terminated.

47. **Body interaction** (14): Agents have a physical body that informs them of expended energy and effort, that has some movements more natural than others, that executes unintentional actions, that perceives the world through constrained modalities, and that offers a sensation of execution.

48. **Observation of execution** (26): Agents observe the execution of other agents and make judgments as to its adherence to some performance specification.

5 Conclusions

It is common in cognitive science research to draw a distinction between a *mental model* and a *cognitive model*. The former is meant to connote the various ideas and theories that people have and use to explain the world around them, e.g. one might have a mental model of how an engine carburetor works. The latter refers to the theories that cognitive scientists (as well as artificial intelligence researchers) devise to explain some component of human intelligence, e.g. one might have a cognitive model of inference by analogy. The model of strategic planning that is revealed by strategy representation is certainly an example of one of these two types, but which one?

On one hand, we'd like to argue that this is a mental model of human planning. We believe both that each of the strategies represented in this work describe a family of analogous strategy instances, and that the mechanisms of analogical reasoning require that these instances be represented as mental entities with deep structural alignment.

On the other hand, this model is far from intuitive, and we traditionally like to think of mental models as things that can be easily described by the people that have them. Strategies, by their very nature, are uncomfortably abstract for most people. That these representations required the skills of AI knowledge representation experts (cognitive scientists, no less!) suggests that the model of planning that is presented here is a cognitive model, and should be treated as a disprovable theory of human planning.

Perhaps the more interesting possibility is that this model is an example of both types, i.e. that the mental model that people have of themselves and others (albeit inaccessible) has some sort of correspondence to the cognitive processing that they engage in. Presumably our ability to effortlessly recognize, evaluate, and utilize strategies as a part of our daily life requires some alignment between the way we think we work and how we actually do.

References

[Collins, 1986] G. Collins. Plan Creation: Using strategies as blueprints. Ph.D. dissertation, Yale University.

[Falkenhainer *et al*., 1989] B. Falkenhainer, K. Forbus, D. Gentner. The Structure-Mapping Engine: Algorithm and examples. *Artificial Intelligence*, 41, pp 1-63.

[Gentner, 1983] D. Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 1555-170.

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

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

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

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