Playing Chess with Machiavelli:
Improving Interactive Entertainment with Explicit Strategies

Andrew S. Gordon

IBM TJ Watson Research Center
30 Saw Mill River Road
Hawthorne, New York 10532
asgordon@us.ibm.com

Abstract

Current interactive entertainment is not strategy-aware, in that it does not allow users to execute a wide range of real-world strategies in the context of the artificial environment, it does not recognize when users are trying to execute these strategies, and it does not include software opponents and partners that are capable of executing a varied range of strategies in response to user actions. A solution to this problem is to explicitly identify a broad range of real-world strategies that people would use if the environment of the interactive entertainment were real rather than fiction, to represent these strategies to identify their component features, and to design the entertainment environment so that these component features are supported and recognized. This paper gives an example of how the first two steps of this solution can be done in the political domain by examining 60 explicit strategies outlined in Machiavelli's *The Prince*.

The Problem of Strategy

One could imagine that playing chess with Niccolo Machiavelli (1469-1527) would be a very entertaining experience. Florence-born political strategist and author of *The Prince* (Machiavelli, 1998), Machiavelli had interesting views on the role of rulers (the king and queen), religious leaders (the bishops), and the citizenry and soldiery (the knights and pawns). Among the possible tactics that he might use are:

1. *Hold awe inspiring military exercises* (chapter 10, *The Prince*). In reference to the cities of Germany, Machiavelli would periodically move his knights and pawns in elaborate, awe-inspiring synchronization with the intent of dissuading his opponent from launching an attack in his well-protected territory.

2. *Appear religious and humane* (chapter 18, *The Prince*). Machiavelli's king would periodically move next to the bishop, publicly praising the divine work and importance of this piece. When isolated from the other pieces, however, the king would ruthlessly slaughter his opponents and countrymen alike to advance his own interests.

3. *Use an arbiter to manage conflict among the subjects* (chapter 19, *The Prince*). Aptly noticing a territorial dispute between his own queen and rook, Machiavelli would have

the king summon a nearby knight, and entrust him to listen to both parties' claims before making a judgment on the king's behalf, insulating the king from either party's subsequent vengeance.

4. *Foster then crush animosity to improve renown* (chapter 20, *The Prince*). Noticing that an opponent's pawn had entered his territory, Machiavelli would have his king instruct his knights to ensure that the enemy pawn had some military success, as much as to appear that the king himself was in grave danger, before slaughtering the helpless pawn in an awe-inspiring attack.

To win or to lose a chess game against Machiavelli would be equally entertaining, as either scenario would unfold as a tale of interesting interpersonal conflict, camaraderie, treachery, and above all, rich strategic thinking. But if we were to imagine that somehow Machiavelli could play with a modern computer chess program, or any other type of interactive entertainment, we would guess that he would find the experience frustrating, at best. Machiavelli was primarily a strategist, and as such, three critical problems with modern interactive entertainment can be identified.

1. *The possible actions available to users do not permit strategic playing*. The strict rules of chess leave little room for negotiating agreements between opponents, rousing troops with propaganda or fear, gathering information with spies and double-agents, or disarming existing troops in favor of raising a more loyal military force. The most advanced modern computer games in the domain of political and military leadership do not fare much better. The challenge for game designers is to offer users a wide breadth of actions to enable them to execute sequences that express a strategic approach.

2. *The software does not recognize the strategies that users are executing*. Even when users manage to execute interesting strategies within the constraints of possible actions, current interactive entertainment software will make no acknowledgment of it. When Machiavelli allows opponent's pawn to threateningly approach his king before mercilessly destroying it, a strategy-aware version of computer chess should respond in a manner that reflects the renown that the king has gained through this act, perhaps by

pulling back some opposing pieces in retreat, or by selecting less risky plans of attack for future moves.

3. *Software opponents and partners do not respond with strategies of their own.* While designers of interactive entertainment spend an enormous amount of time programming opponents and other non-player characters to act in strategic ways, this behavior is rarely *in response* to the interesting strategies of the human user. When Machiavelli touts the religiousness of the king through his close association with the bishop, an entertaining computer chess program should launch a counter strategy that reveals the maliciousness of the king when he is out of proximity with his subjects, perhaps by forcing this user to decide between saving the bishop piece instead of a knight or pawn that is currently instrumental to some malicious goal of the king. The challenge for designers is to construct software that not only *plays with* the human user, but also *plays along* with the strategies that these users employ.

Explicit Strategies as a Design Solution

A solution to these three problems of strategy in interactive entertainment is to design these environments from a strategy-centric perspective. The suggested approach is to facilitate strategic play by being explicit about the strategies that users may wish to employ. This approach involves three critical steps:

1. *Identify strategies within the domain.* In any domain of an interactive entertainment title, be it sports, war, politics, government, combat, business, or crime, there are a large number of real-world, culturally-shared strategies that can be reasonably employed to achieve goals within the domain. Machiavelli helped define the breadth of culturally-shared political strategy in writing *The Prince* for the ruling Medici family in Florence, a set that has evolved and grown as events in our political history unfold and are discussed. The first step in designing strategy-aware interactive entertainment is to examine the domain strategies in breadth, to name those that don't already have a name in natural language, and catalog them in a list.

2. *Represent strategies in planning terms.* Strategies are knowledge that assist people in the process of planning to achieve a goal. As such, they can be represented by the planning-theoretic elements with which they are composed, as demonstrated by previous on work the representation of strategies for use in AI planning systems (Collins, 1986; Jones 1992). In representing the set of culturally-shared strategies in a domain, a set of planning elements are revealed that must be supported in an interactive entertainment environment if it is to enable strategy use. In our own work on strategy representation, we have found that these elements typically include abstract actions that agents can perform (e.g. preventing and defending), characteristics and relationships between elements in the environment (e.g. adversarial relationships and territorial boundaries), and the

processes and mental states involved in human planning (e.g. auxiliary goals and planning constraints).

3. *Instantiate planning terms in game design.* In any given interactive entertainment environment, the set of planning elements that are revealed by strategy representations must be instantiated, or grounded, in the actual computational elements of the software. For example, individual or sets of user commands must be identified that correspond to the action elements in the strategy representations, including abstract notions of defending and threatening. Concepts such as resources and information must be referenced as variables or objects in the environment. New environmental variables must be created to monitor the various, changing relationships that hold among software characters and the user.

Having grounded the planning elements of the strategies to the processing of the interactive environment, these systems can monitor for users' execution of strategies by matching user actions and environmental variables to the representations themselves - and respond accordingly when various strategies are triggered.

Admittedly, none of these three steps are at all straightforward. Indeed, each of them entails an enormous amount of laborious effort, as well as the resolution of many open research issues. Perhaps the best approach to executing these three steps is divide the labor involved into three research and development camps. The task of identifying the set of culturally shared, real-world strategies that people use is probably best done as research in the cognitive sciences, specifically that of cognitive anthropology. The task of representing these strategies in planning-theoretic elements falls squarely on the shoulders of AI planning research. The task of instantiating planning elements into interactive entertainment environments is best done by the developers of these titles.

The remainder of this paper describes our research efforts in attempting to execute the first two steps of this approach in the domain of political strategy.

Step 1: Identify Strategies

In exploring this three-step approach to developing strategy-enabled interactive entertainment, we focused on the domain of Machiavelli's interest: political strategy. Computer games in this domain have been particularly popular, and include titles that place the user in diverse roles such as city mayor, conquering emperor, and civilization ruler. To investigate the breadth of real-world, culturally shared strategies that these titles should support, we treated Machiavelli's *The Prince* as a strategic catalog in this domain. That is, we analyzed this text to extract the set of strategies that it advances, with the assumption that this work was so influential to western political theory that the identified strategies would reference ideas that are commonly understood in western culture.

In retrospect, it is clear that political strategy has advanced significantly since the 1500's, and this catalog is far from comprehensive. Nevertheless, the 60 strategies that we identified from this text do seem to have a commonsensical quality to them, a claim that is weakly supported in that they can be easily referenced with very short descriptive titles. Figure 1 lists these 60 strategy titles, which are extracts or paraphrases from Machiavelli's text, roughly in the order in which they appear.

Step 2: Represent Strategies

The aim of strategy representation is to define any given strategy such that all situations that match the definition are positive examples of the strategy, and all cases that do not match the definition are not examples of the strategy. Recognizing that the same strategy may be applicable in a wide variety of situations - even those that cross domain boundaries - our efforts have focused on strategy representations that are of the highest possible level of abstraction while still meeting these definition requirements.

In defining the abstract structure of strategies we are, in effect, explicitly explaining 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 strategies like those in Figure 1. Instead, we have

1. Extinguish the family of the former lord
2. Change no laws or taxes
3. Take residence in new land
4. Send colonials to new lands
5. Avoid maintaining armed forces
6. Become defender of neighbors
7. Weaken powerful neighbors
8. Prevent foreign footholds
9. Support weak neighbors
10. Conquer by using discontented members
11. Dismantle lands with tradition of self rule
12. Back ideological revolutions with force
13. Establish stable power base in favor of popularity
14. Abolish old armies for new ones
15. Abolish old alliances for new ones
16. Partner with forces that can block your enemies
17. Schedule all blows to citizens at once
18. Avoid oppressing the people
19. Secure popularity after being established by nobles
20. Make yourself needed by your people
21. Fortify your towns but not the countryside
22. Build protecting walls and ditches
23. Establish sufficient artillery
24. Stockpile food and war supplies
25. Give useful work to the community
26. Hold awe inspiring military exercises
27. Give hope that sieges will be brief
28. Cause fear of the cruelty of a sieging enemy
29. Use your own armies rather than mercenaries
30. Study only the art of war
31. Keep armies organized
32. Keep armies drilled
33. Learn the geography of your country
34. Learn how to learn the geography of other countries
35. Study the history of past rulers
36. Permit yourself to be mean once you are established
37. Favor being liberal than mean before you are established
38. Be conservative with the wealth of your people
39. Be liberal with the wealth obtained from other lands
40. Support armies by pillage, sack, and extortion
41. Favor cruelty to mercifulness
42. Favor being feared to being loved
43. Uphold the appearance of good qualities while acting otherwise
44. Appear religious and humane
45. Use an arbiter to manage conflict between people and nobility
46. Arm conquered people that are unarmed to build loyal armies
47. Disarm a new territory when adding to an existing land
48. Foster then crush animosity to improve renown
49. Look for loyalty in groups distrusted at the beginning of your rule
50. Build fortresses when you fear your people more than foreign powers
51. Acknowledge extraordinary things done by citizens
52. Declare favor to one party when two are warring
53. Remove barriers to the financial success of your citizens
54. Entertain the people with season festivals
55. Support commercial guilds and societies
56. Maintain the majesty of the rank
57. Choose servants that think only of their master
58. Allow free speech and advice only from selected wise people
59. Ignore the unsolicited advice and stick to your plans
60. Be adventurous rather than cautious to control fortune

Figure 1. Sixty Machiavellian Strategies from *The Prince*

Strategy 8. Prevent foreign footholds: "Again, the prince who holds a country differing in the above respects ought to [...] taking care that no foreigner as powerful as himself shall, by any accident, get a footing [amongst neighbors]; for it will always happen that such a one will be introduced by those who are discontented, either through excess of ambition or through fear, as one has seen already."

Representation: The planner has a *Role* in a *Power relationship* over a *Community*, and has the goal that this role be *Permanent*. *Adjacent* communities exists as well that are less *Powerful* than the planner's community, that is, the *Envisioned likelihood* that the community could *Achieve* an *Unspecified goal* is less than for the planner's community. Other communities that are not *Adjacent* exists as well, which are equal to or more *Powerful* than the planner's community. The planner *Assigns as a threat* the *State* in which the *Adjacent* communities are in a *Cooperative relationship* with these powerful communities, and plans to *Prevent* the *Adjacent* community from having the goal of having a *Cooperative relationship* of this sort.

Strategy 38. Be conservative with the wealth of your people: "Either a prince spends that which is his own or his subjects' or else that of others. In the first case he ought to be sparing, [...]"

Representation: The planner has a *Role* in a *Power relationship* over a *Community* and has the goal that this role is *Permanent*. The planner has *Acquired resources* as a result of the *Agency* of the agents in the *Community*, or by a *Transfer of possession* between these agents and the planner. The planner is doing *Planning* and has a *Partial plan* that includes these *Resources*. The planner adds a *Planning preference* to *Avoid the action* of *Expending these resources*, and *In the case that* this *Preference is negated* the planner makes a *Planning decision* to *Select a range value* which is the *Minimum value* for any *Range* in the plan of the *Amount of resources* that are *Expended*.

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."

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.

Figure 2. Three examples of pre-formal representations of Machiavellian strategies

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 (Collins, 1987), and where the content of these representations is loosely drawn from a wide range of content theories of planning (notably Owens, 1990).

Figure 2 gives three examples from the 60 representations that we constructed for the strategies of Machiavelli. Terms in the representations that refer to planning elements are capitalized and italicized.

This representation work has been done in the context of a larger effort of our group to catalog a wide range of strategies from many different domains, and to identify a representational vocabulary that is rich enough to adequately define these strategies as well as other types of planning knowledge. Accordingly, we expect that the representations of these 60 strategies will need to be changed in the future to reflect insights gained from the analysis of strategies in other domains, ultimately leading toward formal representations of strategies that can be employed in cognitive simulations of human planning.

For the purpose of enabling strategy use in interactive entertainment, this style of preformal representation is adequate to reveal the design components that are required in these environments. By extracting the planning elements of the 60 representations of the Machiavellian strategies in Figure 1 (indicated by italics and capitalization in Figure 2), we can identify environmental variables and user actions that should be implemented to support and recognize strategy use. In all, 1197 instances of planning elements were used in the 60 representations that we authored. By grouping duplicate and synonymous terms, 262 unique types of planning elements were revealed. These terms were then categorized into one of three classes: 46 terms (246 instances) referred to abstract actions that agents perform, 66 terms (512 instances) referred to features of the environment, and 150 terms (439 instances) referred to mental states and processes involved in creating and executing plans.

Step 3: Instantiate Strategies

Having identified and represented the strategies in a domain, we are left with the difficult task of mapping these

Abstract actions: Acquire resources, Acquire possession, Acquire wealth, Actions, Activities, Agency, Attack, Avoid, Block, Block successfully, Change normal activities, Communicate, Compensated work, Decode, Defend, Destroy, Do normal activities, Enable, Execute a plan, Execute a skill, Execute a partial plan, Execute successfully, Execution ability, Execution problem, Expend resources, Failed execution, Give possession, Inform, Instrumental use, Iterate, Locate agent, Make a work product, Monitor, Observe, Perform a service, Persuade, Prevent, Prohibit, Reorganize, Request, Retaliate, Separate, Take possession, Threaten, Transfer possession, Work service

Environment features: Acquired resources, Activity role, Adjacent objects, Adversaries, Agent having a role, Assistants, Blockades, Boundary, Boundary definition, Causal agent, Cause, Change, Class, Class instance, Community, Component, Composed object, Consecutive events, Controller, Cooperator, Employee, Employer, Encoding, Event, Evil agent, Expected adversary, Expended resources, Giving agent, Highest role, Instrumental object, Located agent, Location, Observed event, Obstacle, Occurrence, Organization, Organizational process, Organizational structure, Past events, Permanent state, Persistent state, Physical characteristics, Physical environment, Physical object, Physical state, Place, Possessed object, Power, Protector, Quantity of resources, Quantity, Randomness, Rate, Region, Relationship, Resource, Role, Similar objects, Simultaneous events, Starting time, State, Subregion, Successional role, Successor, Wealth, Worker

Figure 3. Abstract actions and environment features in Machiavellian strategies

representations onto the elements of a particular interactive entertainment environment. The first challenge is to offer users and other agents in the environment actions or sets of actions that can be mapped onto the abstract actions in these representations. This facilitates the first goal of strategy-aware software: enabling users to engage in strategic play. The second challenge is to identify the correspondence between objects and variables in the environment and the environmental features in the strategy representations. By pairing the mappings from these two challenges with a reasonably sophisticated matching algorithm, designers can achieve the second goal of strategy-aware software: algorithmically recognizing the strategies that users are executing.

Figure 3 lists the 46 abstract actions and 66 environment features that must be instantiated to allow for the execution and identification of the 60 Machiavellian strategies. Some of the elements, such as the action of Attack and the feature of Boundary, may have straightforward correlates in an interactive entertainment environment. Others, such as the actions of Prevent and Enable, may require substantially more work, but successful research implementations of these elements in microworlds have demonstrated its feasibility (see Ortiz, 1999).

The third challenge that designers face is enabling software agents to respond appropriately to user actions with interesting strategies of their own. In the near term, the best approach may be to pair each recognizable user strategy with a static, but appropriate and entertaining response. In the long term, we can imagine that software agents will be based on cognitive models of human planning that incorporate strategic elements. Toward this goal, our future work in this area will be directed at further explicating the mental states and process that people use in strategic thinking - starting with the 150 terms that refer to these elements in Machiavellian strategies - in an effort to broaden the strict search-based approaches of traditional AI planning systems.

Conclusions

Developing strategy-aware interactive entertainment will require a knowledge-rich approach, encompassing a range of research and development disciplines. The identification and analysis of real-world strategies reveal planning elements that are shared across strategic situations. Aligning the components of interactive entertainment titles with these elements will facilitate strategic play, recognition, appropriate response, and ultimately lead to software systems that entertain by exhibiting rich strategic thinking.

References

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

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

Gentner, D. 1983. Structure-Mapping: A Theoretical Framework for Analogy. Cognitive Science, 7, 1555-170.

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.

Jones, E. 1992. The Flexible Use of Abstract Knowledge in Planning. Ph.D. dissertation, Yale University.

Machiavelli, N. 1998. The Prince. Translated by W. Marriott. Urbana, IL: Project Gutenberg.

Ortiz, C. 1999. A Commonsense Language for Reasoning About Causation and Rational Action." Artificial Intelligence, 111, pp 73-130.

Owens, C. 1990. Indexing and Retrieving Abstract Planning Knowledge. Ph.D. dissertation, Yale University.

Tate, A. 1998. Roots of SPAR - Shared Planning and Activity Representation. Knowledge Engineering Review, 13, 121-128.