A First-Order Theory of Stanislavskian Scene Analysis

Leora Morgenstern
IBM T. J. Watson Research Center
Hawthorne, NY 10532
leora@steam.stanford.edu; leora@us.ibm.com

Abstract

At the turn of the last century, Constantin Stanislavski developed a new system of acting, replacing the mannered gestures and forced emotion then popular with a more natural style. The core of his system lay in having actors perform a process of scene analysis, in which an actor would flesh out the circumstances of the play so that the character's motivations and actions would follow logically. This paper is an attempt to ground Stanislavski's method of scene analysis in a formal theory of action. We discuss the relations between Stanislavskian and formal AI theories of action and planning, give a formal definition of the end product of a scene analysis, and characterize the conditions under which a scene analysis is coherent.

1. Background, Scope, and Motivation

Background: At the turn of the last century, Constantin Stanislavski, the founder of the Moscow Art Theatre, developed an innovative system of acting that broke with centuries of tradition in the theatre. Prior to Stanislavski, acting often relied heavily on stock mannerisms, such as putting one's hand to one's brow to indicate despair. A small minority of actors could express genuine emotion on the stage. But it was unclear, even to these actors themselves, how they achieved this display of emotion. An actor could work himself up into some emotional state during one performance, but might subsequently be unable to reproduce it. Still less could he teach others how to perform.

Stanislavski sought to develop a technique that could be taught and replicated. He was opposed to what he believed were the pitfalls of conventional acting: playing to the audience, conventional gestures, working oneself up into an emotional state. Instead he proposed that actors immerse themselves in the circumstances of the play. If an actor would sufficiently flesh out the circumstances of a play, he argued, he would be able to act in a realistic manner. The trick was the "magic if": hypothesizing enough facts, consistent with the play, to make it real to the actor, and enable him to feel, rather than pretend or artificially work-up, the appropriate emotions.

The cornerstone of Stanislavski's (1936) system is the process of *scene analysis*(SA). The actor constructs a backstory for his character, which includes a detailed portrait and history of the character before the play's start and then chooses actions that further his objectives. This enables the actor to imagine in detail the circumstances of the play.

The aim of this paper is to explore the formalization of scene analysis, and characterize *coherent* scene analyses — the ones that seem to "work" for an actor.

Scope of this paper: This paper is in the field of formal commonsense reasoning (Davis 1998; Miller & Morgenstern 1997). The aim is to develop a formal object-level theory that can justify inferences of the form: Scene analysis SA is coherent with respect to Scene SC, where SC consists of a set of characters and a script, and SA consists of a backstory, a set of objectives, a dramatic history, and a mapping between the dramatic history and the script.

What is out of scope: The paper is not concerned with how to automate the construction of a scene analysis. It is not concerned with the NLP issues; it is assumed that the script has already been translated into the knowledge representation language of choice. It is not concerned with the creation of a story line, as in the field of interactive drama (Mateas 2002).

Motivation: We list two motivations for this work:

First, there are striking similarities between the concerns of Stanislavski and of the formal AI/KR community; these are evident in the vocabulary and ontology used by Stanislavski and his followers. Common concepts include characters/agents, actions, objectives/goals, intention, causation, plans, and obstacles. Indeed, Stanislavski wrote about the need for a character's actions to follow *logically* from the circumstances that the actor has imagined; we are exploring to what extent we can develop a theory in which this happens.

Second, this is a promising domain for commonsense formalization. Much research in formal CSR has focussed on artificial problems (McCarthy 1998; Miller & Morgenstern 1997; Sandewall 1999). In contrast, the scene analysis problem is real. Actors and directors frequently use Stanislavski's methods, and have a good sense for when a scene analysis works. The process relies more on commonsense reasoning than acting technique: even novice actors can do scene analysis.

Indeed, examining the SA problem offers a new perspective on traditional CSR domains. For example, traditional planning focusses on plans in which success is guaranteed or at least likely. But characters in plays (as in real life) often construct and execute plans that don't have a strong chance of success. (Consider a hero's many attempts to woo the heroine.) More important is a character's ability to recognize and recover from failure. These concepts are central to the development of coherence in this paper.

2. The Scene Analysis Process

2.1 Example of scene analysis

There are many variations of the scene analysis process (SAP). We use here a modification of (Gewirtz 2007), itself based on (Lewis 1991) and (Hagen 1991). The examples here refer to Stanislavski's production plan for *Othello* (Stanislavski 1983). We focus on the first scene, which begins with Iago's (I)urging Roderigo (R), who has unsuccessfully pursued the Venetian lady Desdemona (D), to tell her father, Brabantio (B), of her elopement with Othello (O).

The SAP includes the following steps:
1. Constructing a *backstory* for one's character, including the personality traits of the character, and the actions that have happened prior to the start of the play
2. Determining a character's *scene objectives*
3. Determining which *strategies* the character uses to achieve his objectives, and the *actions* that each strategy comprises
4. Identifying the *obstacles* that stand in the way of the character executing his strategies
5. Choosing the *strategies* and *actions* that a character uses to overcome the obstacles.

For example, an actor playing Iago could construct an (informal) scene analysis containing the following:
Iago's objective throughout the play: Avenging himself on Othello.
Iago's scene objective in the first scene: Breaking up Othello and Desdemona's marriage.
Strategy: Get R to inform B of the elopement so that B will ask officials to annul the marriage.
Obstacles: R is angry at I, doesn't trust I's judgement, and doesn't want to help him.
Strategy to get around obstacles: Appease R; convince R of I's trustworthiness; remind R of how I has also been hurt by O; persuade R that informing B will indeed result in the desired outcome.

The actor could construct a backstory explaining why R is so angry at I and doesn't trust his judgement.

2.2 Scene analysis vs. classical AI planning

We highlight some important differences between SA and classical AI planning.

First, the term *action* is used differently in the two contexts.

In the acting domain, actions can be divided into 3 groups:
1.**locutionary actions**: the utterances (from the script) that the character speaks;
2. **blocking actions**: the movements that the character makes while on stage;
3. **dramatic actions**: the essential actions that move a play forward and move each character toward his objectives. Example: Iago's convincing Roderigo to ally himself with Iago. Other dramatic actions include tempting, pleading, and stealing.

Stanislavskian Scene Analysis (SSA) is primarily concerned with dramatic actions. Much of the work in the SAP consists of mapping the locutionary actions to dramatic actions. For example, R's utterance to Iago in Act 1,"By heaven, I rather would have been his hangman," can be interpreted interpreted either as his *rejecting* I's explanation or as *sympathizing* with I. Both rejecting and sympathizing are dramatic actions; merely uttering a sentence is not. Dramatic actions are frequently not explicit in the script; the SAP makes them explicit.

Indeed, instances of locutionary and blocking actions may be co-extensional with instances of dramatic actions: i.e., one may perform a dramatic action *by performing* (as in (Goldman 1970)) a locutionary and/or blocking action. E.g., A might perform the dramatic action of *consoling* B by placing his arms around B (a blocking action) and uttering "there, there" (a locutionary action). An analysis of the relationship between locutionary and/or blocking actions, on the one hand, and dramatic actions on the other — i.e., determining what really happens when a character utters a line and moves on stage — is at least as difficult as the general story understanding problem and beyond this paper's scope.

Second, the focus on planning is different. AI planning focusses on constructing a plan which is guaranteed or at least is likely to achieve the desired result. But plays would be of limited interest if characters' plans usually succeeded. Rather, a character does his best to choose actions which he believes will help him pursue his objectives, and replans as (is often the case) necessary.

Consider, e.g., I's objective to break up O and D's elopement. His strategy is to convince Roderigo to inform B, so that B will appeal to Venetian authorities to annul the marriage. I knows that a precondition of convincing R is having R favorably disposed toward him. He knows that if B would turn out to be unmoved by the elopement, then his strategy would fail. In fact, he knows when his strategy fails (for another reason, when the duke refuses to annul the marriage.) And he replans, choosing another strategy (making O insanely jealous of D) that will achieve his larger scene objective, namely, avenging himself on O.

3. A Formal Theory of Scene Analysis

3.1 Theory of Action

Our language is based on (Davis 2004). We use a situation-based branching theory of time. Intervals are defined by their starting and ending situations. Fluents are properties that change over time. *Holds(s,f)* (*Holds([s1,s2],f)*) indicates that fluent *f* holds over situation *s* (interval *[s1,s2]*).

Actions take place over intervals of time. *Occurs(ac,s1,s2)* denotes the action *ac* occurring between *s1* and *s2*. We can also say *Occurs(Do(a, act), s1, s2)*, which denotes the action of agent (or character) *a* performing *act* between *s1* and *s2.act* denotes an *actional*, an action which is not anchored to a particular agent.

The logic presented is monotonic. The frame problem is handled using explanation closure axioms.

The theory of knowledge and belief is based on the possible-worlds theory of knowledge introduced by (Kripke 1963) and extended by (Moore 1980). *B(a,s1,s2)* denotes that situation *s2* is indistinguishable to *a* from *s1*, given *a*'s beliefs.

Expected Effects, Success, and Failure: Because characters must frequently deal with their plans not succeeding, it is important to explicitly incorporate a concept of failure into the theory. The domain theory therefore includes:

Precondition axioms, of the form *Occurs(ac, s1, s2)* $\Rightarrow$ *Holds(s1, f)*. *f* is a **precondition** of *ac*,

Success-Effect axioms, of the form *(Occurs(ac, s1, s2)* $\wedge$ *Holds(s1,f1))* $\Rightarrow$ *Holds(s2, f2)*. *f1* is a **success condition** of *ac*, and *f2* is the **effect**, and

Failure axioms, of the form *(Occurs(ac, s1, s2)* $\wedge$ *Holds(s2, f2))* $\Rightarrow \neg$ *Holds(s1,f1)*. *f1* is a **failure condition** for *ac*.

The predicates *Precond(ac,f)* (*f* is a precondition of performing *ac*) and *FailCond(ac,f1,f2)* (*f1* is a sufficient condition for the failure of *ac* to achieve *f2*) will be useful in stating the definition of coherence in Sec. 3.3.

Objectives and Strategies:

Objectives: An objective (goal) is represented as a fluent, a state that an agent wants to achieve.

Strategies: Our development of strategies is influenced by (Gordon 2004). A strategy is a relatively loose structure of actions, much like a partial plan, with the following characteristics:

- It comprises a set of actions, or other strategies, that may be used to accomplish a particular strategy.
- It may, but does not necessarily, mandate the order in which these actions/strategies must be performed/executed.
- It does not necessarily contain the complete set of actions necessary for success; in fact its execution does not guarantee success in or even likelihood of achieving one's objective.
- It will typically have gaps; there is not an action prescribed for every time period in which the strategy is executed. (An agent may execute a second strategy during a gap in the first strategy.)
- It may include actions performed by other agents. E.g., I's strategy for breaking up O and D's marriage includes R's action of informing B of O and D's elopement. (The strategy will, of course, fail if the other agent does not do his action in a timely manner.)
- It may include *reactive actions* — actions that respond to a particular situation, or action of another agent. E.g., (Section 4), a strategy to keep a secret may include the reactive action of refusing to answer an agent's question if doing so would entail the secret becoming known.

Strategy Execution: Analogous to the concept of an action occurring during an interval is the notion of a strategy being *executed*,completely or incompletely, during that interval. We distinguish between complete and incomplete executions. The predicate *Executes(a,strat,f,s1,s2)* denotes that *a* executes strategy *strat* in pursuit of objective *f* between *s1* and *s2* and that the execution is completed.*StartExecute(a,strat,f,s1,s2)* denotes that *a* begins to execute the strategy *strat*, and that the execution takes place between *s1* and *s2*. The execution may not be complete at *s2*.

Constructing Strategies:

(We first introduce some syntactic sugar: *Occurs*$(ac(\vec{x}) \mid P(\vec{x})$*, s1,s2)* denotes the occurrence of action $ac(\vec{x})$ with the range restricted to *P* in the obvious way. We extend this syntactic sugaring convention to the predicates *Executes* and *StartExecute*).

$\Phi_i(\vec{x_i})$ is a *simple plan segment* if it is of the form $[\neg]Occurs(ac(\vec{x_i} \mid P_i(\vec{x_i})), ss_i, se_i)$ or $[\neg]Executes(a,strat(\vec{x_i} \mid P_i(\vec{x_i})), f, ss_i, se_i)$ or $[\neg]StartExecute(a,strat(\vec{x_i} \mid P_i(\vec{x_i})), f, ss_i, se_i)$.

$\Psi_i(\vec{x_i})$ is a *plan condition* if it is of the form $\bigwedge_i Holds([sa, sb], f_i(\vec{x_i})) \wedge \bigwedge_i \Phi_i(\vec{x_i})$ where each $\Phi_i(\vec{x_i})$ is a simple plan segment.

A *conditional plan statement* is of the form $\Psi_i(\vec{x_i}) \Rightarrow \bigwedge_j \Phi_j(\vec{x_j})$, where the earliest time point of each $\Phi_j(\vec{x_j})$ is no earlier than the latest time point of $\Psi_i(\vec{x_i})$.

A *strategy* is of the form $\{(ss,se) \mid \bigwedge_i \Upsilon_i(\vec{x_i}) \wedge \kappa \}$ where each $\Upsilon_i(\vec{x_i})$ is either a simple plan segment or a conditional plan segment and κ is a conjunction of constraints of the form $s_i < s_j$ or $s_i \leq s_j$.

Example 1: { *(s1,s6)* | *Occurs(Do(a1,Act1(x1)* | *P1(x1)), s1,s2)* $\wedge$ *((Holds(s3,P2(x2))* $\wedge$ *Occurs(ac2(x3),s4,s5))* $\Rightarrow$ *Occurs(ac3,s5,s6))* $\wedge s2 < s4 < s3 < s5$ } is an example of a general strategy.

Example 2: {*(s1,s4)* |
Occurs(Do(a1,Give(a2,x1))|*Book(x1)* $\wedge$ *Rare(x1)), s1,s2)*
$\wedge$ *(Holds(s3,Likes(a2,x1))*
$\Rightarrow$ *Occurs(Do(a2,Give(a1,x2))* | *Baseball(x2)* $\wedge$
$\wedge$ *Autographed(x2,BabeRuth),s3,s4))* $\wedge s2 < s3$ }
is an example of a specific strategy, e.g., that of trading treasured objects.

Interaction between strategies and objectives:

At any situation in time, an agent has at least one primary objective, his scene objective. Strategies are used to achieve objectives; the pursuit of a strategy may generate other objectives. As an agent uses a strategy to achieve his objectives, he may form objectives to achieve preconditions or avoid failure conditions for the actions in his strategy. In order to formalize the interaction between strategies and objectives, we need to express how an agent proceeds through his strategy. The following functions and predicates will facilitate this discussion:

- *ActionOf(strat, ac)* denotes that *ac* is one of the actions in strategy *strat*.
- *Precursor(ac1,ac2,strat)* denotes that *ac1* must be performed prior to *ac2* when *strat* is executed. (This relation is entailed by the temporal constraints in a strategy.)
- *StrategyFor(f,strat)* denotes that strategy *strat* is a strategy for pursuing objective *f*.
- *StratPart(strat1, strat2)* denotes that strategy *strat2* is a part of strategy *strat1*.
- *Holds(s,SObj(a,f))* denotes that *f* is the scene objective of *a* in *s*.
- *Holds(s,CObj(a,f))* means that *f* is a current objective of *a* in *s*. (There may be multiple objectives.)
- *Holds(s,CStrat(a,f,strat))* denotes that *strat* is *a*'s current strategy in *s* in pursuit of his objective *f*.
- *Holds(s, CAction(strat, a1, f, do(a2,act)))* denotes that the action *do(a2,act)* is a current action for *a1*'s strategy *strat* to achieve objective *f*. Note that *a1* may be distinct from *a2*.

An action *ac* is said to be *done* in *s* relative to some agent *a* and strategy *strat* if there was some interval, ending in *s*, in which *strat* was the current strategy of *a* for achieving objective *f*, and *ac* occurred at some point during that interval.

Definition of done:
Holds(s, Done(ac,a, strat)) $\Leftrightarrow$
ActionOf(ac,strat) $\wedge$
$\exists$ *f, ss', ss, sa, sb Holds([ss,s], CStrat(a,f,strat))* $\wedge$
$\forall$ *s' ss'* $\leq s' < s \Rightarrow \neg$ *Holds(s',CStrat(a,f,strat))* $\wedge$
sa $\geq$ *ss* $\wedge$ *sb* $\leq$ *s* $\wedge$ *Occurs(ac, sa, sb)*

An actional *act* is a *potential action* for an agent *a* pursuing some strategy *strat* if all the precursors of *ac* in *strat* have already been done:

Definition of potential action:
Holds(s,PotAct(a,act,strat)) $\Leftrightarrow$
$\forall$ *ac Precursor(ac,Do(a,act),strat)*
$\Rightarrow$ *Holds(s,Done(ac,a,strat))*

Strategy Failure: The notion of strategy failure is central to the development of a formal theory of scene analysis. An agent needs to realize when his strategies are not working out. *Holds(s, StrategyFailed(a,f,strat))* says that at situation *s*, the particular strategy *strat* that agent *a* has chosen in his pursuit of objective *f* has failed.

We make some observations about the circumstances in which a strategy fails:

1. *a* is pursuing a strategy and has performed an action in this strategy, but the expected effect does not hold.
2. *a* is pursuing a strategy, which calls for him to perform an action. The action has a precondition which does not hold, and he either doesn't know how to or can't perform an action to establish that precondition.
3. *a* is pursuing a strategy, which calls for another agent *a'* to perform an action. *a'* performs the action, but the expected effect does not hold. Or *a'* does not perform the action, or performs the contrary of the action.

An example of the last type of failure can be seen in I's strategy to break up D's and O's marriage. His strategy consists of his convincing R to inform B of the elopement; for B to alert the duke; and for the duke to annul the marriage. However, the duke does not annul the marriage; in fact, he confirms that it is valid.

When a strategy fails, an agent may repeat the strategy or choose another strategy to achieve his objective. E.g., when I's initial strategy to break up D and O fails, he chooses another strategy: making O jealous of D. It is not trivial to characterize in what circumstances an agent will switch strategies or repeat a strategy/action. Certainly, plays — and life — are rife with examples of agents who persist in a strategy and prevail. I, for example, must repeatedly entreat R before the latter agrees to inform B of the elopement. Yet unrestrained persistence is, at best, the stuff of slapstick comedy.

An analysis of strategy persistence vs. strategy switching might formalize the following: that agents may persist in a strategy for a certain amount of time, or repeat an action several times until they reach some threshold of tolerance; that an agent's threshold for repetition depends on a variety of factors, including the ease of performing an action, expected payoff, availability of other strategies, or difficulty of executing such other strategies. This is left for future work.

Motivation: One wishes not merely to posit an agent's scene objective, but to ground this objective. The backstory can provide this grounding. For example, Stanislavski's backstory for *Othello* explains the past connection between O and I, detailing occasions where I saved O's life; I's low-born background; O's decision to choose as lieutenant the high-born but unworthy Cassio, because he needs to appear polished in elegant Venetian society. This *motivates* I's resentment and explains why I wants to avenge himself on O.

The notion of motivation used here is significantly weaker than that, say, of (Stein & Morgenstern 1994) (where an action was motivated if its occurrence was entailed). This theory retains the concept of free will: No matter what has happened, a person is never forced to choose an objective. Rather, we introduce the predicate *Motivated(a, f)*, provide axioms for this predicate, and then show that particular backstories entail particular instantiations of the *Motivated* predicate. Even if an objective is motivated for a particular character, however, it is not necessarily the character's scene (or current) objective.

3.2 Scenes and Scene Analysis

We define a scene *SC* as a tuple $< Char, \Sigma >$, where *Char* is the set of agents/characters in the scene and Σ is a sequence of (mostly) locutionary actions. (Σ

may include dramatic actions that are entailed by the script, e.g., the script of *Othello* entails that Othello kills Desdemona. However, in general, most dramatic actions are introduced during the SAP.)

We define a scene analysis *SA(SC, A')* as a tuple $<$ *Char*, Σ, *[SS,SE]*, *BStory(A',SS)*, *Obj*, Δ*(A',SS,SE)*, $\Pi >$, where

- *Char* and Σ are as above.
- *SS* and *SE* are the starting and ending situations of this instantiation of the scene.
- *BStory(A',SS)* is the backstory of character *A'* up to situation *SS*, defined as a set of sentences, each of which is of the form *Holds(s, f)* where $s \leq SS$. or is an occurrence sentence whose latest time point is earlier than *SS*.
- *Obj* is a set of fluents, the objectives of *A'*.
- Δ is the dramatic history of the scene, defined as a set of sentences each of which is of the form *Holds(s,f)* or is a simple strategy segment whose active agent (the agent executing the strategy) is in *Char* and whose active interval (the interval in which the strategy is executed) is contained in the interval *[SS,SE]*.
- Π relates subsets of Δ to subsets of Σ. That is, Π associates dramatic actions with lines in the script.

Γ*(SA(SC,A'))* is defined as the union of the sentences in the backstory and the dramatic history.

3.3 Coherence

Our goal is to characterize those scene analyses that make sense, that "work" for an actor. Informally, we would like to say that a scene analysis is *coherent* if the following conditions hold:

[a] The scene objectives are motivated with respect to the backstory

[b] Any other objectives arise from the original scene objectives, the strategies taken to pursue objectives, and the facts that are true during the scene

[c] An agent will pursue a strategy only for an objective

[d] An agent's actions during the scene follow from his objectives and chosen strategies

[e] An agent will not continue a strategy that he believes has failed.

This is formalized in the definition below.

Definition: Let SC be a scene and SA(SC,A') a scene analysis for character A', as defined above. Let Γ*(SA(SC,A'))* be the set of wffs associated with the scene analysis, as defined above. Let $\Gamma(CSK)$ be a set of sentences representing a body of commonsense knowledge. (E.g., for *Othello*, this might include commonsense domain theories about wooing spouses.)

Then SA is coherent iff Γ*(SA(SC,A'))* $\cup\ \Gamma(CSK) \models$

1. (motivation of scene objectives **[a]**)

($\forall s \in [ss, se]$ *Holds(s,SObj(A',f))*
$\Rightarrow$ *Holds(s, Motivated(A',f)))* $\wedge$

2. (subgeneration of other objectives **[b]**)

(Holds(s, CObj(A',f)) $\Rightarrow$ *Holds(s, SObj(a,f))* $\vee$
$\exists$ *strat, ac,f' (Holds(s,CStrat(A',f, strat))* $\wedge$
ActionOf(ac,strat) $\wedge\neg$ *Holds(s, Done(ac,A', strat))*
$\wedge$ *(Precond(ac,f)* $\vee$ *FailCond(ac,*$\neg$*f,f')))* $\wedge$

3. (strategy pursuit only for objectives, and only if not failed **[c]**)

(Holds(s,CStrat(A',f, strat)) $\Rightarrow$
((Holds(s,CObj(A',f)) $\wedge$ *StrategyFor(strat,f))* $\vee$
(Holds(s,CObj(A',f')) $\wedge$*StrategyFor(strat',f')*
$\wedge$*StratPart(strat,strat')))* $\wedge$
$\neg$ *(B(A',s,s')*
$\Rightarrow$ *Holds(s',StrategyFailed(A',f,strat))))* $\wedge$

4. (actions performed by A' only if done as part of A's strategy **[d]** and only if it is not believed that they'll fail **[e]**)

(occurs(s,s',do(A',act)) $\Rightarrow$
Holds(s,CStrat(A', f,strat)) $\wedge$
Holds(s,PotAct(A',act,strat)) $\wedge$
$\neg\exists$ *f' (FailCond(Do(A',act),f,f')* $\wedge$
$\forall$*s(B(A',s,s')* $\Rightarrow$ *Holds(s',f))))*

4. Example and Proof Sketch

To demonstrate how one can perform a scene analysis and show that it is coherent, we use a small sample script, used in teaching principles of scene analysis, adapted from (Kahan & Rugg 1998).

A: Give me that.
B: No.
A: Give it to me.
B: I don't think so.
A: Come on: I really want it.
B: No!
(A grabs it from B.)
B: Well?
A: Well what?
B: Well, say something.
A: What do you want me to say?
B: You might have something to say.
A: I'm not going to say anything.

This scene is clearly ambiguous (are the characters two children arguing over a toy? a parent and child?); the point of scene analysis is fleshing it out.

What follows is an overview. (The full analysis and proof are at http://www-formal.stanford.edu/leora/stanislavski.) We do the scene analysis from B's point of view. First we present the elements of the scene analysis: the scene, a backstory, B's objectives, the dramatic history, and the mapping between locutionary actions and the dramatic history. Then we prove that this scene analysis is coherent according to the definition in Section 3.3.

We highlight the main points of the scene analysis:

The backstory: We posit a backstory where A and B are in a relationship. B wants to break up with A. B has already purchased a one-way ticket to the Bahamas. A has just seen that B is holding something, but doesn't know what it is. B is a non-confrontational person.

B's objective: breaking up with A without having to say to A's face that she wishes to break up with him.

B's strategy is the Runaway strategy. It consists of hiding her desire to break up with A until she begins her trip, taking her trip, and then writing a letter to inform A of the breakup. Hiding something is itself a strategy. It consists of a reactive action: if Y's objective is to hide something, and X asks Y to do something which would entail X's finding out, then Y must refuse.

The dramatic history: A asks B to hand him the ticket. B knows that if she agrees to this request, A will find out that she has a one-way ticket, and will infer that she is planning a breakup. To execute the hiding strategy, B must therefore (repeatedly) refuse A's request. A then grabs B's ticket. At this point, B's strategy to hide her desire to break up with A has failed. Indeed, her Runaway strategy has failed. However, she still has the same objective: to get out of her relationship with A. She now switches strategies, to taunt A with her one-way ticket to the Bahamas. She tries this several times, but A does not take the bait.

The proof of coherence: The formal proof is simple. First (condition 1), we demonstrate that B's objective is motivated by the backstory. This follows from some commonsense domain axioms on relationships and nonconfrontational tendencies.

Next (condition 2), we demonstrate that at any point, all current objectives of B are scene objectives or are generated from the scene objectives. Since we deal only with B's single scene objective, this is trivial.

Next (condition 3), we demonstrate that B pursues her current strategy only when it lines up with her current objective and only when she knows the strategy hasn't failed. Assume that A's grabbing the ticket occurs between *Sa* and *Sb*. Until *Sb*, B's current strategy is the Runaway strategy. At *Sb*, B realizes that this strategy has failed, and switches to the Taunt strategy, which still lines up with her current objective.

Next (condition 4), we demonstrate that B performs actions only if they are part of her current strategies. We consider each of B's 6 actions. Her first 3 actions, which happen before *Sb*, are refusing A's request to hand over the ticket. These are part of the Hiding strategy, which is itself part of the Runaway strategy. Her next 3 actions, which happen after *Sb*, are part of the Taunt strategy/action. *This completes the proof sketch.*

5. Related Work

There have been two previous studies relating AI and Stanislavskian theory. El-Nasr(2004) develops an interactive narrative architecture, using techniques of user modeling and behavior analysis that are based on Stanislavskian theory. The focus is on constructing new stories rather than on analyzing existing scenes. Hoffman(2006) considers how one might apply Stanislavskian theory to construct robots that interact with humans. A primary focus is the physical actions and gestures that a robot would perform. Neither El-Nasr nor Hoffman works in formal logic; neither addresses the notion of coherence.

There are clear connections between our work and first, the long tradition, dating back to (Schank & Abelson 1977), of story understanding using knowledge of an agent's goals and plans; and second, the work toward a declarative theory of reactive planning (Traverso *et al.* 1996). These lines of research are less formal than our theory, and do not address the notion of coherence. There are also connections between our work and that of (Sandewall 1998) who examines within a formal logic how agents might recognize and recover from plan failure.

6. Conclusions and Future Work

We have presented a formal theory of Stanislavskian scene analysis. Our theory builds upon previous theories of action, but adds several new elements, including concepts of strategy, strategy failure, and coherence.

Future work includes first, developing a more detailed theory of strategy failure; and second, extending the notion of scene analysis to multiple agents.

In the long term, we wish to attempt to formalize a later stage of the SAP: determining which of a character's emotions are supported by a scene analysis. We believe that this requires much preliminary work in developing a formal structure for representing and reasoning about emotions.

Acknowledgements: I am grateful to Laurence Gewirtz for guidance and helpful discussions, and to Ernie Davis for his comments on an earlier draft of this paper.

References

Davis, E. 1998. The naive physics perplex. *AI Magazine* 19(3):51–79.

Davis, E. 2004. Knowledge and communication: A first-order theory. *Artificial Intelligence* 166:81–140.

El-Nasr, M. S. 2004. A user-centric adaptive story architecture: borrowing from acting theories. In *Advances in Computer Entertainment Technology*, 109–116. ACM.

Gewirtz, L. 2007. Ten questions for scene study. http://laurence.gewirtz.googlepages.com.

Goldman, A. I. 1970. *A Theory of Human Action.* Englewood Cliffs: Prentice-Hall.

Gordon, A. S. 2004. The representation of planning strategies. *Artif. Intell.* 153(1-2):287–305.

Hagen, U. 1991. *A Challenge for the Actor.* Scribner.

Hoffman, G. 2006. Acting lessons for artificial intelligence. In *Proc., 50th anniv. summit of AI (AI-50).*

Kahan, S., and Rugg, K. W. 1998. *Introduction to Acting.* Allyn and Bacon.

Kripke, S. 1963. A semantical analysis of modal logic i: Normal modal propositional calculi. *Zeitschrift fur Mathematische Logic und Grundlagen der Mathematik* 9:67–97.

Lewis, R. 1991. *Advice to the Players.* Theatre Comm. Grp.

Mateas, M. 2002. *Interactive Drama, Art and Artificial Intelligence.* Ph.D. Dissertation, CMU.

McCarthy, J. 1998. Elaboration tolerance. In *Working Papers, Commonsense-1998.*

Miller, R., and Morgenstern, L. 1997. The commonsense problem page. http://www-formal.stanford.edu/leora/commonsense.

Moore, R. C. 1980. Reasoning about knowledge and action. SRI Technical Report 191.

Sandewall, E. 1998. Logic based modelling of goal-directed behavior. In *KR'98*, 304–315.

Sandewall, E. 1999. *Logic Modelling Workshop.* http://www.ida.liu.se/ext/eta/lmw.

Schank, R., and Abelson, R. 1977. *Scripts, Plans, Goals, and Understanding.* Lawrence Erlbaum Associates.

Stanislavski, C. 1936. *An Actor Prepares.* New York: Routledge.

Stanislavski, C. 1983. From the production plan of othello. In Cole, T., ed., *Acting: A Handbook of the Stanislavski Method, revised edition.* New York: Three Rivers Press.

Stein, L. A., and Morgenstern, L. 1994. Motivated action theory: a formal theory of causal reasoning. *Artificial Intelligence* 71(1):1–42.

Traverso, P.; Giunchiglia, E.; Spalazzo, L.; and Giunchiglia, F. 1996. Formal theories for reactive planning systems. In *AAAI'96 Workshop on Theories of Action, Planning, and Control.*