2
Related Work

Introduction

  The problem of constructing characters is an ancient problem that many fields have struggled with. Computer graphicists and artificial intelligence researchers have been facing the issue since the birth of their respective disciplines, but both of these are young fields. I'll talk a bit at the end of this chapter about some relevant work that has been done in the fields of CG and AI, but first we'll look to the Media Arts for information and inspiration. Many of the problems that character constructionists faced (and continue to face) in the analog domain carry over into the digital domain. As I mentioned in the last chapter, I spent a fair number of years working in the theater, and in the first part of this chapter I'll draw on my own personal experiences as an actor thinking about character, and briefly mention some sources of information and inspiration from animation, model kits, comics. My point in discussing these issues here is to familiarize the reader with the process of character construction in other domains. We need to understand and respect those processes as we try to map out similar capability in the digital domain. I will summarize what we've seen at the end of the section.

  In the second half of this chapter I'll review some of the relevant literature in CG and AI, with an emphasis on work I feel is particularly relevant to the direction I've taken in this work.

Character Construction in the Media Arts

An Actor Prepares

Before beginning this section, I must point out an important constraint hanging over the actor's shoulder during the entire character construction process: the story. As an actor, and not a script writer, my emphasis here is on the character, not the role within the story. The role is what's written, the character is the particular instantiation of that role for a given performance situation. To be more explicit: The character is always constrained by the role.

The role makes a very specific contribution to the story, to which all is subservient to. The character must work within the bounds of the role, which itself must work with in the bounds of the story. As we discuss the options available during the character construction process, always keep in mind that the choices available are always constrained by the role and its place in the story. While there is a time and a place to discuss the process of creating a role and designing a story, this is not it. With that caveat, we turn our attention to the process an actor, upon being cast in a role, must go through to create the character they will inhabit.

Mirror, Mirror...

When first presented with a role, an actor faces several dilemmas. In my experience, before all else, the actor must first step into the character and look in the mirror. What does the character look like? How does the character hold itself? What does it sound like? In order to begin to start answering these questions, the actor looks away from the mirror and at the material; the *script* and the *backstory*.

Do we only know of this character from the story/set-piece we're about to see the character in? If not, what backstory (see pp 47-62 of **Seeger90** for a nice discussion of this) do we have on the character? Assuming it's consistent (i.e. in the case of comic books or films, it's usually not), what of the back story informs the construction of this character for this piece? Since we can't add to the story, what might we want to add to the backstory that will help create this particular character, that is consistent with what will unfold in the story?

Historically, the look of the character is largely determined by the actor's body. As media move more and more into the digital domain, this becomes less of a central issue, but it's still an important one. In animation this has always been less of a concern, but has

remained important. In animation, the "actor" is a combination of the animator drawing the scene and the voice talent doing the character's voice. Many an animated character's look is directly informed by the voice talent that makes them speak. Also, many animators work with a mirror by their side, so that they can observe some part of themselves in motion as they go about some activity, which they then interpret and apply to the character they're constructing.

Demeanor: A Character's Potential for Action

Once the gross form of the character is sketched out, an actor turns to the question of comportment: how does the character hold itself? Is it hunched over, as if sickly or perhaps sneaky? Does it stand ramrod straight, like a frightened schoolboy or with the frozen menace of a Marine saluting?

All characters have a face, even if it's not a human one with two eyes a nose and a mouth. What does this character's face look like? Is there an amused sneer, a vague look of discomfort, a sheen of sweat, the contorted features of an upcoming burst of rage, a look of benign happiness or a blank stare?

It's important to realize that none of these are questions of action, really, but a question of how the character looks with regard to its potential for action.

Action!

The next question is what happens when the character first opens its mouth, or reaches for something, or waves, or takes a step. Many times the dynamics of the character are exactly at odds with its demeanor: the ramrod straight Marine who can't control his high pitched laughter and giggling, the little old lady with the bass voice, the innocuous looking fat man who saunters across the room with the aplomb of James Bond.

Once the basics of the character have been sketched in, we turn to the specifics of activity that the character will be asked to perform. Does it physically have everything it needs? If the character is supposed to be a private detective, does it have a gun? Does it know how to use it? How does it load it? Does it always leave it unloaded or loaded? Is the safety always on? Is this a problem? How about a pack of cigarettes? How does it light them - does it always get it lit the first time, or not?

At this point, we're still building up the character. If we suddenly realize that our character needs some particular prop, we might look in the script or backstory to see if

there's any information we can use. If not, we might invent some that was consistent with the script and backstory: "This gun was my partner's. Sometimes it sticks, but it's all I've got left of Marty, so I always keep it in my back holster. I like the fact that Marty's still covering my back…"

Building a Character

Once the basics of the character are set, the actor can finally pick up the script and begin to look at what specific things they're going to be asked to do. At this point, the abilities of the character are foremost on the mind of the actor. What exact tasks is the character expected to do? Should they be performed hesitantly, woodenly, excitedly, carelessly, hurriedly, or nonchalant? All of the character that has thus been created; the look, the demeanor, the weird tick when he says "porridge", all of these come into play as the actor begins enacting the tasks the character must perform. Realizing that the character must dance well in a scene in the second act, the actor realizes that some of the fumbling behavior he was thinking of before is now inappropriate. He modifies the way the character carries himself, but now transfers the business to the character's nervous habit with his left hand and a yo-yo, which keeps getting tangled.

Let's Take it Again, From the Top…

I once played the role of Argan, the crotchety, cuckolded husband in Moliere's "Imaginary Invalid". As the play opens, the scripts calls for Argan to be "sitting alone at a table in his bedroom, adding up his apothecary's bills with counters, and talking to himself as he does". He launches into a long monologue on his ailments and the variety of medicines and treatments he's been charged for. As is customary in Moliere, there is some stage direction, but not much.

After reading the script several times and thinking about the actors cast in the other roles, I began my own character construction process. Moliere had written the role of Argan, but now it was my task to create a particular character, my own instantiation of Argan. As I usually do, I began constructing the character for the role by figuring out how he spoke. I started reading the character's lines, and found myself stooping over a bit as I did so. After reading through a few times, I began bobbing my head a little, as if with palsy, as I read the lines over and over, sometimes repeating a line several times in different variations of the voice, sometimes stooping over more, sometimes less. We had originally blocked the scene with the character sitting in his wheelchair, giving his speech. As I

started reading the lines aloud, though, I started shuffling around, stopping and restarting, sometimes gesturing and sometimes glaring. The script didn't call for any particular activity or props; the point of this exposition in the play is to introduce the character and set up the farce about to ensue.

I soon realized, though, that a five minute monologue right at the start with a doddering old man talking about getting enemas and complaining about the cost of his doctor's visits wasn't going to quite have the effect we wanted unless we could grab the audience's attention. I began searching for a "bit of business" for the character to do that worked in the context of the scene.

I eventually came up with a bit that worked just right. Near the beginning of the monologue, Argan is talking about one particular noxious concoction: "a good purgative and tonic concoction of fresh cassia with Levantine senna, etc., according to Monsieur's prescription, to expel and evacuate the gentleman's bile" (**Moliere**, pp 434-435). For the performance, I pulled several (carefully wrapped) fresh eggs out of Argan's housecoat and broke them into a glass, which I contemplated for a period during the course of the next bit of the monologue, and then gulped them down in a sputtering gulp. This never failed to get an audible "urgh!" from the audience, and helped get them to believe that Argan really did take his medicine, and didn't just talk about it. This bit of business (Argan pulling vile concoctions out of various places and drinking them with impunity) was worked into several scenes, and was quite helpful in shaping my particular interpretation of the role.

The important point here is this: What my Argan looked, sounded and behaved like were created over the course of many rehearsals, where one change fed back on another. Discovering that the voice I used for Argan broke when I said a certain word caused me to emphasize it more. That caused me to reassess his facial expression as he spoke. That in turn modulated the pacing of his speaking, which was punctuated by his gestures. Because he was always looking for his pills and evil libations, he was always wearing housecoats with many pockets. Because he voice failed sometimes and because he was lame, he had a cane, which he stamped loudly on the ground to be heard. In a discussion with his brother this cane suddenly became a sword. The script didn't call for this prop by name, but it made sense in the context of the verbal duel the two engaged in.

Summary

My point in going into relatively exhaustive detail about the character construction an actor may go through is to give the reader unfamiliar with this general process and an idea of how iterative and intertwined this process is. Here's one way to think about the process:

- the model's a given; the actor's body was cast in the role
- the model has an enormous number of degrees of freedom; the role or context informs the variables of the model
- which do we use?
- which do we constrain their values to a range?
- what abilities do we bring to bear? What processes will be used in the generation of behavior?
- iterate back over these till done

What Have We Learned?

In addition to acting, I looked at a variety of other domains in the "media arts". In comics (**McCloud93**), there is a rich history of character construction, both in a single issue of comic, over a long story arc, and also over decade-long adventures of continuing characters in consistent universes. Also, there was the curious ubiquity of a phenomenon called "retconning" ("retroactive continuity"), which means essentially where a character's past is changed to fit a new story that's being told. Two famous examples of this are the John Byrne retelling and reinterpretation of Superman (Byrne87) and the Frank Miller reinterpretation and future history of Batman (Byrne86), which also reached back and rewrote history.

Cartooning and animation, of course, from it's very inception has been about character, even when it hasn't been about story. The Disney, or "classical", approach to animation is well-documented (**Thomas81**, **Thomas87, Lasseter87**), although there are certainly other valid approaches (**White88**, **Solomon89**, **Jones89**, **Laybourne79**, **Hart94**, **Culhane88**, **Adamson75**). Either way, such techniques as "model sheets" (showing the character in a variety of poses) and "bibles"(detailing what a character would and wouldn't do in a variety of situations, as well as pointers of how to draw and animate the character) are used even in so-called "limited animation" (**Groening93**).

In addition to their primary audience of adolescents, models and kits have found a active audience with adults who appropriate characters from a variety of media, recast discontinued models, and sculpt and sell "garage kits" of both known and unknown

characters (**Dutt92, Bruegman94**, **Webb94**).

By looking at these and other areas, I was able to see a few things about character construction that were readily applicable to the problem of building 3D semi-autonomous characters in the digital domain:

- Character construction is iterative, therefore characters must be composed of malleable media; whatever the "stuff" they're made of, it needs to retain its plasticity, as it is constantly being reformed, reshaped, remolded during the construction process.
- The character's body, props, demeanor, and abilities are all intimately tied together. During the process of constructing a character, changing one of these almost certainly has ramifications on others. Making the character hold itself stiffly changes the way it is dressed and the way it walks across a room, giving a character a ring to twist might give him an outlet for a nervous tick that was previously expressed by tapping his foot, etc.
- The character builder needs to be able to easily move back and forth between their roles of creator and character: treating the character first as golem, then as avatar, and back again, ad infinitum. be more specific here...
- prototypes: A character can be (and probably will be) part stereotype, part caricature. Building on known frameworks or taxonomies helps to speed the process of creation, and derivative and hackneyed components of a character can be replaced over time. Character construction is a process; not since Athena sprung full formed from the forehead of Zeus has a character appeared fully formed and realized.
- Collaboration is a fact of life in the creative process of character construction. The collaboration might be with other creators of the same character (i.e. the penciller, inker and writer of a comic book character), or with creators of previous incarnations of the same or similar roles (i.e. the actor who played the role previously), or with others involved in the scene the character is in (i.e. the director of the piece).
- reusability: One person's character is another person's raw material. In the same way that fans appropriate television characters (Jenkins92), and model makers reshape figurines (Dutt92), and comic artists "retcon" superheroes (Byrne87, Miller87), we must be prepared for the appropriation and reuse of our character and its constituent parts.

Task Level Animation

Zeltzer discusses a three part taxonomy of animation systems: guiding, animator level, and task level. Guiding includes motion recording, key-frame interpolation, and shape interpolation systems. Animator level systems allow algorithmic specification of motion. Task level animation systems must contain knowledge about the objects and environment being animated; the execution of the motor skills is organized by the animation system. The work in this dissertation is a set of components needed for a task level animation system.

In a task level animation system, there are several kinds of planning activity that can go on. In this work, I am concerned only with the lowest level of planning— what Zeltzer called motor planning (**Zeltzer87**). Motor planning is similar to the kind of problem solver Minsky calls a difference engine (**Minsky86**).

> *"This reflects current notions of how animal behavior is structured in what we call an expectation lattice, in which motor behavior is generated by traversing the hierarchy of skills selected by rules which map the current action and context onto the next desired action."* (**Zeltzer87**)

Since I am concerned with this lower level of activity, I refer to the behavior of the characters I'm building as "semi-autonomous", since they are being directed explicitly by some higher level control system (a human or some planner). They do act and react autonomously within the given context of a task, but the selection of the original task is left to some other system.

Some Interesting Actors and Characters

Using forward kinematic techniques, Zeltzer showed a biped with many degrees of freedom that could walk over uneven terrain (**Zeltzer84**). His system was a step towards an animation system that allowed interaction at the task level, although the available motor skills of the animated figures were limited to forward locomotion.

Girard's PODA system has creatures that can walk, run, turn, and dance using kinematics and point dynamics (**Girard85**). Again the emphasis in this system is on the animation of legged locomotion, and allowing the animator control over its creation. Autonomy of the animated creatures is not the goal, rather intelligent and artistic control by the animator is.

Sims designed a system for making creatures that, using inverse kinematics and simple dynamics, could navigate over uneven terrain (**Sims87**). This system was notable in that the notion of "walking" was generalized enough that he could generate many different kinds of creatures that all exhibited different behavior very quickly. More recently, Sims has developed a system for quickly prototyping creatures embodying a set of physically-based behaviors by breeding them (**Sims94**). He presents a genetic language that he uses to describe both the shape and the neural circuitry of the creatures. His work is most interesting in the context of building systems in which creatures are bred by using aesthetic decisions as fitness functions. This work, more than any other, shows the power of genetic techniques when applied to complex computer graphic character construction problems.

Reynolds describes a system based on the actors model of distributed computation for animating the behavior of flocks and herds (**Reynolds82**, **Reynolds87**). The use of the actor model allows for a great amount of flexibility, but the communication overhead between actors imposed for their particular application is non-trivial (O(n2)).

Also of note are Miller's snakes and worms, which use relatively simple notions about the motion of real snakes to generate quite interesting motion (**Miller88**). The locomotion is controlled by a behavior function which allows the snake to be steered towards a target.

One of the most ambitious animated creatures to date is a dynamic hexapod that was developed here in the Computer Graphics & Animation Group at the MIT Media Lab by McKenna and Zeltzer (**McKenna90A**, **McKenna90B**). They demonstrated an articulated figure with 38 degrees of freedom, that uses the gait mechanism of a cockroach to drive a forward dynamic simulation of the creature moving over even and uneven terrain. It is an example of how successfully biologically-based control schemes can be adapted for computer animation. A virtual actor hexapod that uses the same gait controller and exhibits several simple behaviors has been also been demonstrated.

More recently, Xiaoyuan & Terzopoulos demonstrated a framework for the animation of fish that provides "realistic individual and collective motions with minimal intervention from the animator."(**Xiaoyuan94**) Their fish animations are remarkable life-like in their behavior, and show the power of combining state-of-the art behavior animation techniques with high quality shape and shading.

Even more recently, Blumberg and Galyean demonstrated a "directable" dog

character that can respond autonomously, in real-time to user input in the context of a larger, scripted narrative activity. (**Blumberg95**)

Animation Systems

Badler et al. describes a system for translating NASA task protocols into animated sequences that portray astronauts performing specified tasks in a space station work environment (**Badler91**). The focus of their research is concerned more with portraying and evaluating human motor performance for specified tasks, or for instructing agents in the performance of tasks, rather than the development of architectures for representing and implementing virtual actors. More recently (**Badler93**), they discuss the larger issues in building virtual humans, with particular emphasis on interacting with the virtual humans through spoken natural language. Their system is far and away the most comprehensive with regard to modeling human figure motion, although they are not concerned with the autonomous reactive behavior of the virtual human in its environment, but rather realistic human motion that can be used predictively for human factors studies.

Reeves et.al (**Reeves90**) describe an animation system called Menv ("modeling environment") which is most notable because of its successful usage in the creation of some of the most compelling computer-based character animation to date (**Luxo Jr.**, **Red's Dream** , **Tin Toy**, **KnickKnack**). They discuss the concept of **articulated variables** in their modeling language, **ML**. ML is a C-like procedural language in which animation is effected by allowing certain variables to change over time. Although I didn't know of this work until after I'd designed my original system, this work and discussions with one of its authors (**Ostby94**) directly inspired and shaped my later work on **eve** (the modeling language in WavesWorld), as will be described in Chapter 4.

Strassmann (**Strassmann91**) built a system, Divadlo, that was in the same spirit as this dissertation (building a system that treated AI and CG as equal partners). While WavesWorld has veered heavily towards the CG side, Strassmann's work emphasized on a natural language interface to the system.

Another system that had a similar object-oriented perspective towards animation was the SWAMP system (**Baker92**), where she used the CLOS notion of streams as an abstraction of a control strategy. This is similar in spirit to what this work does with agents and articulated variables, but is less powerful, as she doesn't deal with the issues of blending the effects of complementary strategies.

Materials & Scene Description Languages

Cook first proposed a flexible tree-structured shading model that can be used for specifying complex shading parameters with a small number of parameters (**Cook84**). This work was especially interesting in showing the diverse ways that textures could be used to modify the shading of a given piece of geometry. This work was eventually extended by Hanrahan & Lawson (**Hanrahan93**) to a compiled language used by the RenderMan® Interface, a photo-realistic scene description protocol (**Pixar89**).

Sims work on his "evo" system (**Sims91**), in which 2D images are bred by a combination of a special purpose language, genetic techniques, and aesthetic fitness functions provided interactively by a user, is interesting in pointing the way towards future systems where a user iteratively develops an algorithmically generated image by high level interaction only. This work has obvious implications for breeding 2D procedural textures which could correspond to complex natural phenomena.

Recently, the issues of modeling natural phenomena by procedural texturing and modeling have been addressed by several researchers (**Ebert94**). Most of the examples in this book are given using the RenderMan Shading Language.

Artificial Intelligence

Minsky describes a theory in which a mind is composed of a society of interacting parts, each of which, considered by itself, is explicable and mindless, that he calls the *Society of Mind* (**Minsky87**). The work done by Travers for the Vivarium project here at the Media Lab contains good examples of systems of agents that are autonomous and exhibit interesting behavior (**Travers89**). His ideas are loosely based on Minsky's Society of Mind theory and model the behavior of groups of insects using perception sensors of the environment and agent-based representations of the state of each insect's "mind".

Agre and Chapman have developed a theory of general activity (**Agre87**). They argue that there are two kinds of planning, which can be referred to as capital-P Planning and small-p planning. They contend that much of AI research is on Planning, while what people actually do a lot more of is planning. This is similar to Zeltzer's discussion of motor planning as a subset of more general problem solving skills. Their work on Pengi is quite interesting because of their assertion that "we believe that combinatorial networks can form an adequate central system for most activity."

Wilson describes the animat problem, which seems to agree well with the ethological approach Zeltzer has long advocated:

> *"To survive in its environment, an animal must possess associations between environmental signals and actions that will lead to satisfaction of its needs. The animal is born with some associations, but the rest must be learned through experience. A similar situation might be said to hold for an autonomous robot (say on Mars or under the sea). One general way to represent the associations is by condition-action rules in which the conditions match aspects of the animal's environment and internal state and the actions modify the internal state or execute motor commands."* **Wilson87**

He describes a system using a classifier system (a variant of the Genetic Algorithm (**Goldberg89**)) to approach the problem of an animat in a 2D environment.

In work directed toward constructing autonomous robots, Maes has described the details of the connections among skills (competence modules in her terminology) for a "situated" agent (**Maes89**). In her action selection network, each motor skill has a set of preconditions (the condition list) that must be true in order for the skill to execute. In addition, there is a set of predictions about the state after the motor skill has executed: an add list of propositions expected to become true once the skill has executed, and a delete list of propositions that will no longer be true. Skills are interconnected through these preconditions, add and delete lists in the following ways: a skill S1, that, when executed, will make true the precondition for another skill S2 is called a predecessor node, and S1 may receive activation energy from S2. A skill S2 that has a precondition that will be made true by some other skill S1 is a successor of S1 and receives activation energy from S1. There are also conflicter relationships that correspond to inhibitory connections among nodes.

Importantly, Maes has introduced the notion of *spreading activation*, which provides for graded recruitment of motor resources-potentiation is not a binary switch, but a continuous quantity, so that a skill may be potentiated by varying amounts. This is also in agreement with the ethological account. The process of action selection takes into account the global goals of the agent, as well as the state of the world. Activation is spread to the skills from the goals and the state, and activation is taken away by the achieved goals which the system tries to protect. Activation is sent forward along the predecessor links, and backwards along the successor links; activation is decreased through the conflicter links, and each skill's activation is normalized such that the total activation energy in the system remains constant. If all the propositions in the condition list of a skill are satisfied

in the current state of the world, and that skill's activation energy is higher than some global threshold (as well as being higher than all the other modules in the network), that skill is invoked to perform its assigned action (thereby adding the propositions in its add list to the state and removing those on its delete list) and returns. If no skill is selected, the global threshold is reduced by some amount. Either way, the spreading of activation continues, as described above.

Rod Brooks has argued that AI should shift to a process-based model of intelligent systems, with a decomposition based on "task achieving behaviors" as the organizational principle (**Brooks86**). He described a subsumption architecture based on the notion that later, more advanced layers subsume earlier layers, in a sense simulating the evolutionary process biological organisms have undergone. He argues that AI would be better off "building the whole iguana", i.e. building complete systems, albeit simple ones, rather than some single portion of a more complex artificial creature. To this end, Brooks has spearheaded the construction of several successful (to varying degrees) mobile robots (**Brooks89**).

One example of a mobile robot based on the subsumption architecture was programmed by Maes to learn how to walk (**Maes90**). The algorithm was similar to the one previously described by Maes (and the one implemented in my SMVS thesis, **Johnson91**) with the addition of simple statistically based learning. In the chosen domain (hexapod walking), the algorithm proved appropriate and accomplished its goal, although it is unclear how well it scales or transfers to other domains.

Maes' early work on reflective systems (**Maes87**), coupled with work by Malone et.al. concerning the economics of computational systems (**Malone88**), is especially relevant when considering how to build systems that can modify purposefully modify themselves based on a notion of the computational milieu they are embedded in.

The work by Bates and his students on the Oz project at Carnegie-Mellon (**Bates91**) concerns similar topics addressed in this dissertation, although their approach is wildly different. Their work emphasizes the "broad but shallow" capabilities of their agents, but they say very little about how they plan to wed the interesting AI capabilities they have been developing with computer graphic systems. Their Woggles system is engaging, but their interest seems to lie in using it as a litmus test to show that they are on the right track rather than as any sort of framework to modify and improve the characters. The work

discussed and implemented for this dissertation is intimately concerned with the process of creating characters; while their work focuses on the artifact of their particular creatures. I argue in this dissertation that this is an important difference in approach, especially if we are to learn from our experience and expand our character construction abilities.

Blumberg's recent work on Hamsterdam (**Blumberg94**), on the other hand, is an excellent example of a successful marriage of ethologically inspired AI control systems hooked up to a sophisticated real-time graphics system. Blumberg emphasizes the ethological basis for his planner, and is concerned with building animal-like creatures that can be interacted with in real-time. In contrast, the work being proposed here is more concerned with iteratively building up scalable behaviors that are not necessarily wedded to real-time systems, but can adapt themselves at run-time to the computational capabilities of computing environment they find themselves in. Also, this work is more concerned with the general question of building virtual actors, whose internal mechanisms may or may not have a basis in ethology.