Massachusetts Institute of Technology
Artificial Intelligence Laboratory
April 1977

AI WORKING PAPER # 145

A COMPUTATIONAL THEORY OF ANIMATION

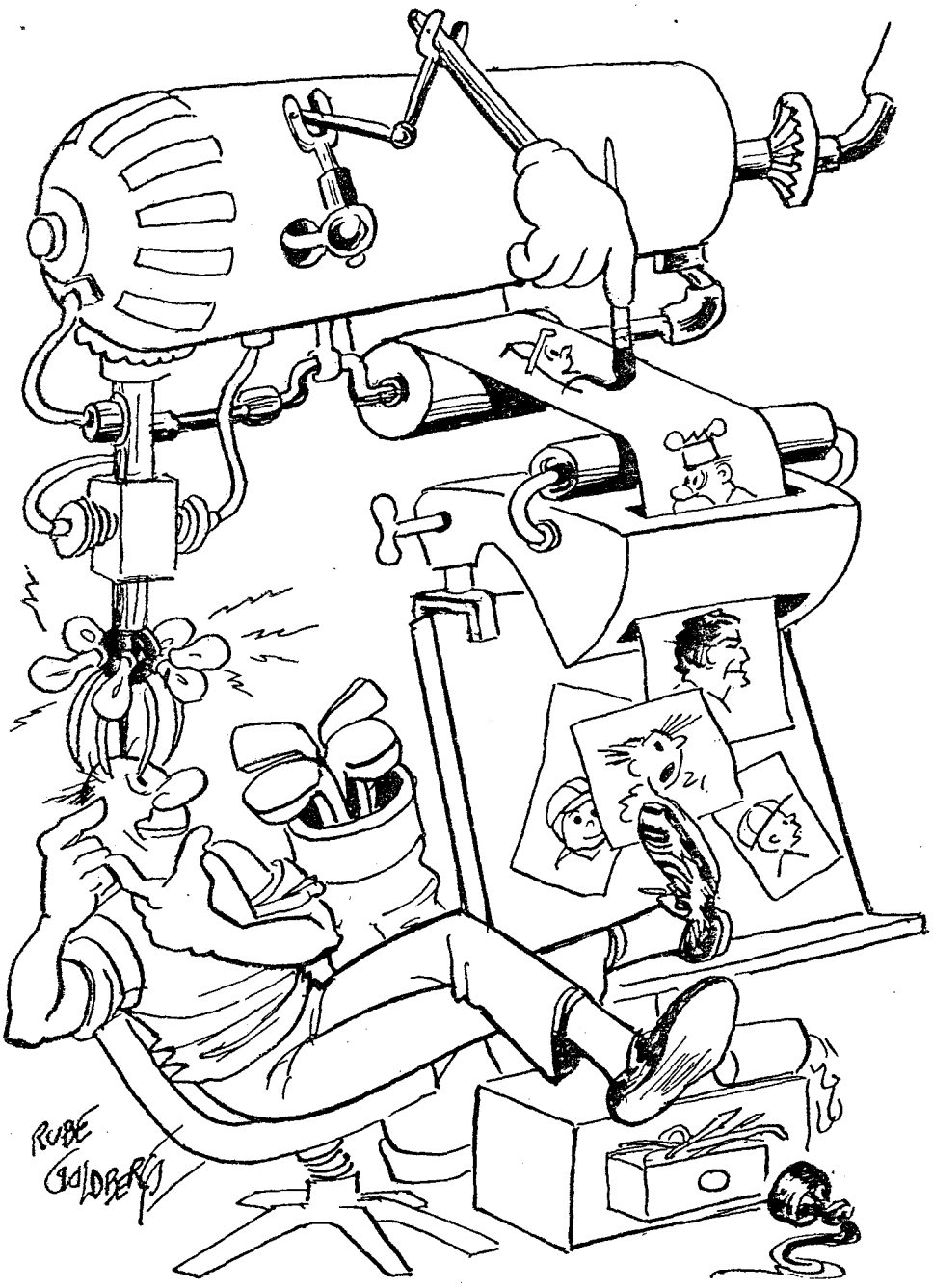
by

Kenneth M. Kahn

Abstract

A system is proposed capable of generating narrative computer animation in response to a simple script. The major problem addressed is how to imbed into the system some of the knowledge that animators use when creating animation. Infinitely many animated films can fulfill a single script. The system is faced with the problem of how to make a *good* one by making decisions in very under-constrained situations. This paper is a total revision of AI Working Paper 119.

The author of this work is supported by an IBM Fellowship. The research described herein is being conducted at the Artificial Intelligence Laboratory, a Massachusetts Institute of Technology research program.



CARTOONIST'S EASY FUTURE—GET THE IDEA AND THE MACHINE DOES THE REST.

TABLE OF CONTENTS A Computational Theory of Animation

I. Introduction and Overview

I intend to implement a computer animator, i.e. a computer program, imbedded with some of the knowledge of an animator, capable of generating computer animated films in response to sketchy scripts. The proposed system will contain several knowledge bases enabling the system to create motions, shapes and colors that convey the intended personalities, moods and interactions. The animation will be produced using an actor-based animation language which I am developing concurrently. [Kahn 1976] The main emphasis of this research is to contribute to artificial intelligence, however, I hope that contributions to computer graphics, animation and aesthetics will also result.

A) My view of aesthetics

Aesthetics is concerned with the nature of the process of creating art, the work of art itself, and the experiencing of art.[1] The system I plan to implement will be a working model of the first, the nature of the creative processs. A computational view offers little insight into the nature of the work of art and computers will need to be able to see before attempts to enable them to experience works of art should be attempted. The quality of a work of art or aesthetics can profitably be broken into the following aspects:

1) Structural. This part is concerned with the overall scheme, structure, story or script. In the proposed system this will be provided by the user and will define the task to be performed.

---

1 While some of things I will say may apply to film, static visual arts and the non-visual arts, the scope of this paper is limited to the aesthetic and affective aspects of animation.

2) Emotional. This is where I see the system being the strongest. The system will know of the moods, feelings, personalities and relationships that motions, shapes, colors, and their interactions and transformations can convey.

3) Perceptual. This is the part that is concerned with what looks pleasant or beautiful. For example, flicker (i.e. alternating dark and light frames) for many people is perceptually unpleasant.[2] A more positive example is that bright and harmonious colors are typically more pleasant than muddy, dull or clashing colors. The perceptual aspect includes more traditional aesthetic concerns such as composition and balance. An important component of this is the handling of global interactions of each aspect such as the color or shape.

4) Symbolic. The symbolic aspect of a work of art depends upon a cultural set of associations between what an image depicts and other concepts. I plan to investigate this aspect the least since it requires too much knowledge of the real world to be imbedded into the system.

B) The kinds of animation I expect the system to produce

I intend to limit the kinds of animation that the system will produce (at least initially) to narrative, non-representational object oriented films. That is the system will not know anything about the appearance of real-world objects, such as trees, faces, etc., but instead will be limited to using only more abstract geometrical objects, such as circles, blobs, triangles, spirals, etc.. This

---

2 This example, as with most in this paper, is culturally dependent and can be considered correct only when all else is normal. These assumptions could easily be overridden since any of the other aesthetic aspects (structural, emotional or symbolic) or the context may dominate.

limitation is to avoid the need to imbed in the system tremendous amounts of knowledge. For example, if the system were to use eyes then it would need to know how they move, vary, express emotions, and are combined with the rest of a face. It can be argued that by restricting myself to non-figurative images I am making it harder on myself to have my characters express emotions. Many facial and body gestures are universal, often the gestures are even common to all primates [Darwin 1872]. I am still considering their use since they are relatively well understood and described, however I expect that the necessary added complexity will keep faces and bodies out of the implementation. Animation is hard enough to teach a computer without adding the problems of teaching a computer to draw.

I have chosen narrative animation because a "story" gives the system a lot of structure to work within. The words "story" and "narrative" should be interpreted in their most general sense. I do not plan to get involved in the problems of having the system understand stories in a natural language. An example of a script is presented in a later section.

C) How I hope to go about this

I plan to attack the problem of creating a system capable of producing a wide variety of such films in response to simple stories by investigating separately the uses of the speed, path, size, shape, color and texture of objects, to convey personality, mood and emotions. Only after reasonable success with this will the interactions be taken into account. Initially, I intend to concentrate on the first two aspects. Eric Martin, an animator and teacher, has emphasized the dynamics of animation in statements like, "Animation displays process" and "movement has as much integrity as drawing". He gives class exercises to produce films with only one small dot that

does not change color or shape or size. It is very instructive to see how much can be expressed this way. The dot's character can be conveyed as happy, burdened, neurotic, frightened or whatever. Certain areas of the space can be characterized simply by the way the dot behaves. If, for example, the dot avoids an area, slowly approaches it and suddenly runs away, one tends to think of the area as frightening or dangerous.

When single dot films are mastered, another dot (identical to the first) will be added. Very complex interactions corresponding to friendship, love, hate, fear, dominance, rebellion, etc. will be portrayed. For example, one might ask the system to produce a dot movie of a neurotic dot (conveyed by erratic, jerky movements) that comes across a happy dot (which moves rhythmically in smooth curves with a gentle velocity) and after initial trepidation on both their parts, they slowly merge. The resulting dot is happy but after a while the neurotic personality takes over.

Color could be added to the dots to help give them character. The happy dot might have a bright red color, the neurotic one might be a dark green color (or better still, change color occasionally). In same way that I plan to make dot movies concentrating on motion to debug my handling and understanding of it, I plan to also make color movies. The problem would be to see what can be conveyed when limited to changing the color of a few uniformly colored regions.

Shape could also be used to enhance the characterization. Round objects are typically friendlier than pointed ones, irregular objects stand out more (i.e. have more character) than symmetric ones. (Of course, context can influence this, a square among blobs is the oddball). Shape films could also be explored where a shape could suddenly or gradually (using interpolation routines) change into another.

D) Interactions and conflicts

The next problem is the interactions of the various aspects of the film. A slow rhythm chosen for a section of a film to emphasize the seriousness of the section might conflict with a character's energetic impulsive dynamics. Various solutions are possible, if, for example, there are other characters they might be used to retain as much as possible of the mood of the slow rhythm and the contrast might even enhance the energetic character's personality. If the slow rhythm is considered important enough, the character's motions could still be put into synchronization with the rhythm. Suppose the desired rhythm is a simple three second beat, then the character's motions might be modified to be a small multiple, e.g. a 1 or 1.5 second beat.

Other kinds of conflicts are possible. The motion of an object may be so over bearing as to wash out any desired effects of the colors. In such situations either the motion clues should be underplayed or the color changes exaggerated. Conflicts can arise because two rather similar motions are being used to convey different personalities. In such cases the differences between the motions usually need to be increased. This could be done in either of two ways. One is to generate a difference description and then to modify each motion to widen that gap. For example, wild joyfulness might call for fast, curvy, irregular motion and neurosis might call for fast, straight, irregular motion. If both these conditions need to be conveyed simultaneously (for different individuals) then their difference, in this case straight versus curvy, would be exaggerated. The other way such conflicts might be resolved is based on the fact that there are often a few different ways of the portraying the same mood or personality. The choice between them could be partially based on the largest difference between the two motions.

E) How the system will achieve effects

The system needs to know at least one way to achieve each effect it may want to achieve. The descriptions of the means of accomplishing desired effects must be manipulable in many ways to facilitate the kinds of solutions to the interaction problems just mentioned. The means/effects descriptions should consist of a small set of stereotypical actions, each action itself should be described in such a manner that it may be easily modified. The system will need to be able to exaggerate or down play any aspect of the action. In many cases, it will need to be able to impose a rhythm upon the stereotype, or to stretch or shorten the length of the component actions of the stereotype.

The system also needs some knowledge of how to implement the suggestions for achieving a desired effect. As an example, suppose the system wants to create a character that is simple, orderly, and dull. Associated with each of those adjectives may be a short list of suggestions. Let us just consider the suggestions relating to shape. Now objects will have a list of properties, such as whether an object is closed or not, or the number of sides it is composed of, which, if any, axes it is symmetric about, how sparse or busy it is, etc.. The suggestions for "orderly" might be that the object should be closed, be composed of a small number of sides, symmetric (perhaps about both the x and y axis), etc.. Dull may suggest much the same thing with maybe the additional suggestions that the object should be simple, that none of its component lines cross, and that a good example of a dull object is a square. Dull may also suggest that the object not be very different from most of the other objects in the film. So at this point the system has a list of desired attributes of the object it is about to create. They may conflict or be redundant. There may be other restraints on the choice, such as the requirement that the object contrast significantly

with another object which has already been determined to be say a pentagon. Another restriction may be that the system is being asked to be original. This might mean it should not make this decision as it did previously (to do this it would have to remember having resolved a similar situation earlier) or more simply it may mean that suggested examples (such as "dull"'s suggestion to use a square) be avoided if possible.

Once the system has decided that it wants an object, it must generate the code to realize it. Suppose it has decided it wants a closed, symmetrical, simple, four-sided figure, then it will have to match the desired object against procedures that can draw it. In this case, "Poly" will match since it will be described as drawing closed, symmetrical, n-sided objects.[3] Associated with "Poly" will be information as to how to use it to achieve particular aims. Among this information will be the advice that states if you want a simple n-sided figure then call "Poly" with an angle equal to 360 divided by n.

The above example would indicate that each graphical procedure will need to have machine-understandable specifications. Hopefully such specifications will be simple, a description of the kinds of objects it can draw and advice as to how to use it to draw particular objects. Since I intend to limit the system to simple geometrical shapes and a few blobs and squiggles this should not be too complicated. In addition to specifications for graphical procedures, it will also need specifications of programs that transform, move, or manipulate objects. This investigation into the representation and use of the knowledge about display procedures might lead to significant contributions to computer graphics techniques. Currently, computer graphics is done without the

---

3 "Poly" is the name of the prototypical Logo graphics program. It draws polygons by moving a "turtle" forward, then turning an angle and repeating.

machine having any understanding of the procedures that it is executing. The usefulness of this self understanding is one thing I hope the system will demonstrate.

F) Global Constraints

At any point in the decision-making process the system will have to make many rather arbitrary decisions. Should it make this object red or orange, should this object be a square or a hexagon, should this object follow this trajectory starting from the left, right, top or bottom, should this object be medium or large sized? Sometimes the system will be unable to find any reason to choose one over the other. Hopefully, in such cases the choice will be inconsequential. More often, however, there will be more global considerations that will constrain the system's choices. For various reasons, the system may have any of the following constraints to operate under:

1) variety, that is the choice is restricted, if possible, to those alternatives not already chosen for other objects.

2) simplicity, this applies at all levels, the shape of the objects and their interaction, the composition, the interactions between objects etc.. Many people place a very high value on simplicity or economy that produces complex effects.

3) originality, to the system that will mean both making decisions that are different from past decisions and ignoring, if possible, specific suggestions or examples when alternatives exist.

4) film length, the choice of motions and interactions of objects will be influenced by any constraints as to the desired running time of the film.

5) rhythm, a desired rhythm will also influence the dynamics of the film.

6) mood, a desired mood will affect the choice of most of the parameters.

7) contrast or lack therof, often it will be desirable to have an object be distinguished and once its characteristics are determined this places constraints on the others.

8) tension, often can be created by generating expectations by following some rule, pattern, or stereotyped sequence and then at the appropriate moment breaking it.

9) exaggeration or lack therof, if a scene is to be emphasized then the actions and interactions need to be taken to extremes.

10) coherence, usually coherence of theme, style, mood, etc. gives the work a certain maturity or sense of purposefulness.

11) obviousness, often effects should be obvious, other times they should be subtle depending upon the audience, the desired affect, etc..

II. An Imaginary Scenario

The problems of creating stories for the system to animate is beyond the scope of this research. Instead I will be making up and borrowing stories for it to do. Myths, folk tales and fairy tales are likely sources. For the following imaginary session with the system, the user will be trying to get the system to make a film based on a very simplified version of "Cinderella".

There are at least four aspects of the script that need to be communicated to the system. It needs descriptions of the characters, of their relationships, of the plot or action and of restrictions as to the length, overall mood, pace, level of complexity and originality.

A) Description of the Characters and their Relationships

First we will enter descriptions of the characters, as follows:[4]

```
(CREATE CINDERELLA
        (PHYSICAL-DESCRIPTION (AND BEAUTIFUL SHABBY))
        (PERSONALITY (AND GOOD FRIENDLY HARD-WORKING SHY))
        (ROLE-IN-STORY MOST-IMPORTANT))

(CREATE STEP-MOTHER
        (PHYSICAL-DESCRIPTION UGLY)
        (PERSONALITY (AND MEAN SELFISH STRONG EVIL)))
```

Next we will define the relationship between Cinderella and her step mother and then continue defining the other characters.

```
(CREATE (RELATIONSHIP STEP-MOTHER CINDERELLA)
        (AND DOMINATES HATES))

(CREATE (RELATIONSHIP CINDERELLA STEP-MOTHER)
        (AND OBEDIENT TOLERANT))
```

---

4 Readers unfamiliar with the notation of internal machine representations of knowledge should skim the expressions that follow, if they are difficult to comprehend. Think of them as bits of knowledge in a very stylized and restrained language.

```
(CREATE FAIRY-GODMOTHER
        (PHYSICAL-DESCRIPTION (AND PRETTY MAGICAL))
        (PERSONALITY (AND GOOD KIND STRONG)))

(CREATE (RELATIONSHIP CINDERELLA FAIRY-GODMOTHER)
        (AND POLITE GRATEFUL))

(CREATE (RELATIONSHIP FAIRY-GODMOTHER CINDERELLA)
        (AND PROTECTIVE GENEROUS HELPFUL))

(CREATE PRINCE
        (PHYSICAL-DESCRIPTION (AND BEAUTIFUL STRONG))
        (PERSONALITY (AND GOOD STUBBORN DETERMINED)))
```

Notice that by using the adjective "strong" for the prince's physical description rather than for his personality I am indicating that only his appearance need convey strength and that no commitment is made regarding behavior. This contrasts with the step mother's strength specified earlier.

```
(CREATE (RELATIONSHIP-BETWEEN CINDERELLA PRINCE) LOVE)
```

To aid this exposition, I will ignore the other characters, the step sisters, king, queen, and servants.

B) The Description of the Plot

Next I shall enter in the plot. The personalities and relationships that are established in the following scenes were described above.

```
(CREATE (SCENE INTRODUCTION)
        (ESTABLISH (AND (PERSONALITY CINDERELLA)
                        (EMOTIONAL-STATE CINDERELLA (JOY SLIGHTLY-POSITIVE))
                        (PERSONALITY STEP-MOTHER)
                        (EMOTIONAL-STATE STEP-MOTHER (JOY SLIGHTLY-NEGATIVE))
                        (RELATIONSHIP-BETWEEN CINDERELLA STEP-MOTHER))))
```

```
(CREATE (SCENE KEPT-APART)
        (SEQUENCE: (CINDERELLA WANTS (CINDERELLA MEETS PRINCE))
                   (CONVEY (PREVENTED-BY STEP-MOTHER (CINDERELLA MEETS PRINCE)))
                   (ESTABLISH
                    (AND (EMOTIONAL-STATE CINDERELLA (JOY VERY-NEGATIVE))
                         (EMOTIONAL-STATE STEP-MOTHER (AND (JOY POSITIVE)
                                                           (PRIDE POSITIVE)))))))

(CREATE (SCENE STEP-MOTHER-GOES) (DISPLAY (EXIT-BY STEP-MOTHER)))

(CREATE (SCENE INTRODUCE-FAIRY-GODMOTHER)
        (ESTABLISH (PERSONALITY FAIRY-GODMOTHER)))

(CREATE (SCENE CINDERELLA-BEAUTIFIED)
        (AND (CONVEY
              (CAUSED-BY FAIRY-GODMOTHER
                         (CHANGE (PHYSICAL-DESCRIPTION CINDERELLA SHABBY)
                                 (PHYSICAL-DESCRIPTION CINDERELLA ELEGANT))))
             (ESTABLISH (EMOTIONAL-STATE CINDERELLA (JOY VERY-HIGH)))
             (ESTABLISH (EMOTIONAL-STATE FAIRY-GODMOTHER (JOY MODERATE)))))
```

By separating the two "establishes" they are forced to happen sequentially since what I am trying to capture here is that the fairy godmother becomes happy because Cinderella is happy. Note also that in the scene introducing the fairy godmother the system is not requested to establish her relation to Cinderella. This is because it becomes apparent during the Cinderella "beautification" scene and the following scene.

```
(CREATE (SCENE NO-LONGER-KEPT-APART)
        (CONVEY FAIRY-GODMOTHER
                (UNDOES
                 (FROM-SCENE KEPT-AWAY
                             (CONVEY
                              (PREVENTED-BY STEP-MOTHER
                                            (CINDERELLA MEETS PRINCE)))))))

(CREATE (SCENE PRE-MEETING)
        (AND (ESTABLISH (EMOTIONAL-STATE CINDERELLA (ANTICIPATION HIGH)))
             (ESTABLISH (AND (PERSONALITY PRINCE)
                             (EMOTIONAL-STATE PRINCE
                                              (AND (JOY LOW)
                                                   (LONELINESS HIGH)))))))
```

```
(CREATE (SCENE MEETING)
        (AND (ESTABLISH (RELATIONSHIP-BETWEEN PRINCE CINDERELLA))
             (ESTABLISH (AND (EMOTIONAL-STATE PRINCE (JOY VERY-HIGH))
                             (EMOTIONAL-STATE CINDERELLA (JOY VERY-HIGH))))))

(CREATE (SCENE JUSTICE)
        (AND (CONVEY (AWARE STEP-MOTHER (TOGETHER PRINCE CINDERELLA)))
             (ESTABLISH (EMOTIONAL-STATE STEP-MOTHER (JOY VERY-LOW)))))
```

I simplified the ending partly to keep this section from getting even longer and also because there are problems with representing the second part of the story. The system will not know about particulars such as glass shoes, or more seriously it might not know about the concept of "fitting". One possible abstraction of the ending might be for Cinderella to begin to lose her bright colors and run off leaving a small part of her that retains its bright colors. The prince then goes around trying to find Cinderella, finds her in her dull state, does not recognize her, until Cinderella is presented with the piece, becomes colorful again and reestablishes her relationship with the prince and their happiness.

C) Constraints upon the Film and Individual Scenes

All that remains for us to complete the script is to give descriptions of the constraints we want the system to conform to. The list of possible constraints was given in a previous section. These constraints can be applicable to the entire film, or only to particular scenes. Suppose we enter the following constraints of the entire film:

```
(FILM-CONSTRAINT CINDERELLA-FILM
                 (AND
                  (VARIETY-LEVEL MEDIUM)
                  (COMPLEXITY-LEVEL LOW)
                  (ORIGINALITY LOW)
                  (FILM-LENGTH (INTERVAL (MINUTES 4) (MINUTES 8)))
                  (RHYTHM MEDIUM-FAST)
                  (MOOD SCENE-DEPENDENT)
                  (COHERENCE HIGH)
                  (OBVIOUSNESS HIGH)))
```

Most of these constraints are typical for fairy tales and also make the system's behavior easier to describe since it will not be trying to be subtle or original. The film will not have any dominant mood, instead each scene will specify its own mood. The scene constraints follow:

```
(SCENE-CONSTRAINT INTRODUCTION
                  (AND
                   (MOOD NORMAL) ;this is the default if omitted
                   (RHYTHM NORMAL);this is relative to the film's rhythm
                   (LENGTH LONG)))

(SCENE-CONSTRAINT KEPT-AWAY
                  (AND
                   (MOOD (CHANGES HOPEFUL DEPRESSING))
                   (RHYTHM (CHANGES FAST SLOW))
                   (LENGTH LONG)))

(SCENE-CONSTRAINT STEP-MOTHER-GOES-AWAY
                  (AND
                   (MOOD DEPRESSING)
                   (RHYTHM SLOW)
                   (LENGTH SHORT)))

(SCENE-CONSTRAINT INTRODUCE-FAIRY-GODMOTHER
                  (AND
                   (MOOD DEPRESSING)
                   (RHYTHM PICKING-UP)))
```

```
(SCENE-CONSTRAINT CINDERELLA-BEAUTIFIED
                  (AND
                   (MOOD (CHANGES DEPRESSING JOYOUS))
                   (RHYTHM PICKING-UP)
                   (LENGTH LONG)))

(SCENE-CONSTRAINT NO-LONGER-KEPT-AWAY
                  (AND
                   (MOOD (AND JOYOUS ANTICIPATORY))
                   (RHYTHM FAST)
                   (LENGTH SHORT)))

(SCENE-CONSTRAINT PRE-MEETING
                  (MOOD ANTICIPATORY)) ;length and rhythm are "normal"

(SCENE-CONSTRAINT MEETING
                  (AND (MOOD (VERY JOYOUS))
                       (RHYTHM FAST)
                       (LENGTH LONG)))

(SCENE-CONSTRAINT JUSTICE
                  (AND
                   (MOOD JOYOUS)
                   (LENGTH LONG)))
```

D) Initial Elaboration of the Character Descriptions

As the script now stands, without the added complexities of closer adherence to the original fairy tale, the system is faced with a difficult enough task animating it. The system will begin expanding the script by expanding independently the descriptions of each of the characters to the extent possible, noting any conflicts. Then it will go through the list of conflicts and under-determined features trying to resolve them based upon desired similarities or differences between the characters. Next the system will go through the script a scene at a time, filling in the descriptions of the dynamics of the characters. Sometimes the character descriptions will be modified in the process of conveying certain effects. For example, in order to emphasize

Cinderella's transition from "shabby" to "elegant" the system will go back to Cinderella's initial appearance and resolve any previous conflicts or undecided features in the favor of shabbiness. Only when the scenes are expanded will the system finalize the details of the appearance and typical dynamics of the characters to the extent necessary to actually realize them.

We will follow the system in attempting to realize the above script. The first task it faces is to establish Cinderella's physical appearance and typical dynamics. Physical appearance is defined by the shape, size, line, color and texture of an object. It is based upon the beauty, complexity, niceness, interestingness, strength and coherence of the character. One's typical dynamics are characterized by their typical rhythm, curvature of path, speed, smoothness, decisiveness and the like.

To decide upon Cinderella's physical appearance each descriptor of her is poled for suggestions. As we have described her, we will need to look at the suggestions of "beautiful", "shabby", "good", "friendly", "shy", and "hard-working". Strengths of the suggestions are provided to help resolve conflicts. As it turns out only the first four descriptors have any suggestions about appearance. The suggestions might be represented in the system as follows:

```
(SUGGESTIONS-OF BEAUTIFUL
                (FOR SHAPE
                     (AND
                      (REGULAR (STRENGTH MODERATE))
                      (CLOSED (STRENGTH HIGH))
                      ((COMPLEXITY-LEVEL MODERATE) (STRENGTH LOW))
                      ((STEREOTYPES: (AND STARS SPIRALS SNOWFLAKES))
                       (STRENGTH HIGH))))
                (FOR SIZE) ;nothing, could be left out
```

```
(FOR LINE
     (OR
      (STRAIGHT (STRENGTH MODERATE))
      (AND
       (SMOOTH (STRENGTH HIGH))
       (LONG (STRENGTH WEAK))
       (GENTLY-CHANGING-WIDTH (STRENGTH WEAK))
       (CURVY (STRENGTH HIGH)))))
(FOR COLORS
     (AND
      ((NUMBER SEVERAL) (STRENGTH MODERATE))
      (HARMONIOUS (STRENGTH HIGH))
      ((OR (PASTEL BRIGHT)) (STRENGTH HIGH))
      ((AT-LEAST ONE (OR PASTEL BRIGHT)) (STRENGTH HIGH))))
(FOR TEXTURE
     (AND
      (REGULAR (STRENGTH MODERATE))
      ((ELEMENTS (SIZE SMALL)) (STRENGTH MODERATE)))))
```

Before I go on and present the suggestions of the other descriptors I should defend the arbitrariness of these suggestions. Clearly beautiful objects with only one or two colors, with jagged lines, with an irregular shape can be created. The suggestions are not intended to be complete, nor as hard and fast rules. As we shall soon see, the system will occasionally make decisions contrary to the suggestions. It should also go without saying that these suggestions are not infallible, ugly objects can be constructed following these suggestions. Nonetheless, these suggestions are very helpful.

The form of the suggestions from "shabby" are not very different but are presented so that later sections using these suggestions are clearer.

```
(SUGGESTIONS-OF SHABBY
                (FOR SHAPE
                     (AND
                      (IRREGULAR (STRENGTH MODERATE))
                      (ASYMMETRICAL (STRENGTH HIGH)))))
```

```
                    (FOR LINE
                         (AND
                          (RAGGED (STRENGTH HIGH))
                          (THIN (STRENGTH WEAK))))
                    (FOR TEXTURE (IRREGULAR (STRENGTH MODERATE)))
                    (FOR COLORS
                         (AND
                          (MUDDY (STRENGTH HIGH))
                          ((STEREOTYPES: (AND BROWNS TANS))
                           (STRENGTH MODERATE)))))

(SUGGESTIONS-OF FRIENDLY
                    (FOR SHAPE
                         (AND
                          (NON-POINTED (STRENGTH MODERATE))
                          ((NOT (COMPLEXITY-LEVEL HIGH)) (STRENGTH WEAK))))
                    (FOR LINE (CURVY (STRENGTH MODERATE)))
                    (FOR SIZE ((NOT LARGE) (STRENGTH MODERATE)))
                    (FOR COLORS ((NOT BLACK) (STRENGTH MODERATE))))

(SUGGESTIONS-OF GOOD
                    (FOR COLORS
                         (AND
                          ((NOT BLACK) (STRENGTH HIGH))
                          ((STEREOTYPE: WHITE) (STRENGTH HIGH)))))
```

"Friendly" and "good" have a few suggestions for appearance and are more important in determining a character's behavior which will be discussed later. At this point a merge of the suggestions is made noting any conflicts. The sources of suggestions will be needed to resolve conflicts and will be kept even when there are no conflicts. The system's description of Cinderella at this point is:

```
(CINDERELLA APPEARANCE
          (SHAPE
           (AND
            (CLOSED (FROM BEAUTIFUL))
            ((COMPLEXITY-LEVEL MODERATE) (FROM BEAUTIFUL))
            (ASYMMETRICAL (FROM SHABBY))
            (NON-POINTED (FROM FRIENDLY))
            ((NOT (COMPLEXITY-LEVEL HIGH)) (FROM FRIENDLY))
            (CONFLICT:
             (REGULAR (FROM BEAUTIFUL)) (IRREGULAR (FROM SHABBY)))))
          (SIZE
           (CONFLICT: ((NOT LARGE) (FROM FRIENDLY))
                      (LARGE (FROM MOST-IMPORTANT))))
          (LINE
           (AND
            (THIN (FROM SHABBY))
            (CURVY (FROM FRIENDLY BEAUTIFUL))
            (LONG (FROM BEAUTIFUL))
            (GENTLY-CHANGING-WIDTH (FROM BEAUTIFUL))
            (CONFLICT: (RAGGED (FROM SHABBY)) (SMOOTH (FROM BEAUTIFUL)))))
          (COLORS
           (AND
            (SEVERAL (FROM BEAUTIFUL))
            (HARMONIOUS (FROM BEAUTIFUL))
            ((AT-LEAST ONE (OR PASTEL BRIGHT)) (FROM BEAUTIFUL))
            ((STEREOTYPE; (AND BROWNS TANS)) (FROM SHABBY))
            ((NOT BLACK) (FROM (AND GOOD FRIENDLY)))
            ((STEREOTYPE: WHITE) (FROM GOOD))
            (CONFLICT: (PASTEL (FROM BEAUTIFUL)) (MUDDY (FROM SHABBY)))))
          (TEXTURE
           (AND
            ((ELEMENTS (SIZE SMALL)) (FROM BEAUTIFUL))
            (CONFLICT:
             (REGULAR (FROM BEAUTIFUL)) (IRREGULAR (FROM SHABBY))))))
```

A few comments are necessary here. In order to avoid a proliferation of disjunctive suggestions, they are typically resolved at this stage. When all but one branch of the disjunction causes conflicts that branch is taken. In the two cases of disjunctive suggestions here (from beautiful regarding color and line) the resolution is more complex. For line the choice is between

straight lines and smooth long curves with gently changing widths. The latter is preferred because only one part of it conflicted, the smoothness, while there was agreement from friendly about the curvature. The conflict between shabby's muddy colors and beautiful's pastel or bright colors was chosen in favor of pastel since pastel is closer to muddy than bright. The strengths of suggestions are also used to prefer conflicts with weak suggestions. The stereotype suggestions would have been removed if the film was constrained to be original. (If it was expected to be very original it might keep around the stereotypes but reclassify them as possibilities to avoid.)

All that remains at this stage of the script elaboration is the description of the typical dynamics of Cinderella. This will provide tendencies and defaults as to her behavior, thereby helping to convey her personality. Usually a character's typical dynamics are modified by the actions required of it. The emotional state of a character will also greatly modify the dynamics of a character, the characteristic dynamics will provide a reference base. These modifications are discussed later. Typical dynamics are generating in the same manner as the physical description, i.e. suggestions from the descriptors of the character are combined and simplified with conflicts being noted. The result of this process for Cinderella is:

```
(CINDERELLA TYPICAL-DYNAMICS
            (AND
             ((SELDOM STILL) (FROM HARD-WORKING))
             (REPETITIVE (FROM HARD-WORKING))
             (DELIBERATE (FROM HARD-WORKING))
             ((AVOIDS STRANGERS) (FROM SHY))
             ((ATTRACTED-TO FRIENDS) (FROM FRIENDLY))
             ((PATH: (LONG CURVES)) (FROM GRACEFUL))
             (RHYTHMIC (FROM GRACEFUL))
             ((LOW ACCELERATION) (FROM GRACEFUL))
             (CONFLICT:
              (SLOW (FROM (AND SHY FRIENDLY)))
              ((NOT SLOW) (FROM HARD-WORKING)))))
```

Disjunctive suggestions will not be discussed since they are handled as they were for the generation of the physical description. "Graceful" is not one of the original descriptors of Cinderella, it was suggested by "beautiful".

The elaboration of the step mother's description will proceed similarly. She will be ugly, pointed (mean), large (strong), simple (different), dark (evil), few colors (ugly and different), reds and browns (evil). The descriptor "different" was not given in the original script, the system generated it by comparing the descriptions of the characters. Of the four characters defined the step mother is the only one that is evil or ugly and so is clearly different. "Strong" suggests that her texture be solid or regular. Since she is different from the others, when the texture conflict in Cinderella's physical description is re-examined, it will be decided in favor of an irregular texture. The step mother's typical dynamics would be swift (strong), decisive (strong), and jagged (ugly, evil) movement. To increase the contrast between Cinderella and her step mother, Cinderella's conflict between slow and faster motion suggestions will be decided in favor of slow.

E) Expansion of the Introduction Scene

The first scene, called "Introduction", is where Cinderella's personality and emotional state, her step mother's personality and emotional state and the relationship between the two are established. The personalities are conveyed by both the characters' typical dynamics and certain tendencies in particular situations. The typical dynamics have already been determined as described above. An example of a tendency is Cinderella's tendency to avoid strangers due to her shyness. The only aspect the two characters' emotional state mentioned in the script is their joy levels. "Joy" has a list of suggestions much in the same way that descriptors like "beautiful" or "friendly" did as described earlier. Joy's suggestions are represented as:

```
(SUGGESTIONS-OF JOY
                (ANY-OF ;i.e. at least one of the following
                 (FOR MOTION
                      (AND (BOUNCY (STRENGTH MODERATE))
                           (SWINGING (STRENGTH MODERATE))
                           (FAST (STRENGTH HIGH))
                           (CURVY (STRENGTH MODERATE))
                           (RHYTHMIC (STRENGTH HIGH))))
                 (FOR COLORS
                      (OR
                       (A-CYCLE
                        (AND VERY-FAST SHORT)
                        (BEST-WHEN (COLORS BRIGHT))
                        (STRENGTH MODERATE)))
                      ((CHANGE (NOT JOYOUS) JOYOUS)
                       (PREREQUISITE:
                        (ESTABLISHED
                         (CONVENTION (COLORS JOYOUS))))
                       (STRENGTH MODERATE)))
                 (FOR SIZE
                      (A-CYCLE
                       (AND FAST SHORT EXTREME)
                       (BEST-WHEN (SPARSE SCREEN))
                       (STRENGTH WEAK)))))
```

One of the ways of conveying joyfulness using color makes use of a convention. A convention is established by picking certain characteristics of the colors, such as the brightness or redness, and then changing the colors as if that were a method of conveying joy. Initially this is always done in conjunction with at least one other method. If whenever a character becomes joyful, it moves in a joyous manner and becomes brighter then after a few instances (the number is largely dependent upon the "obviousness" level of the film as described earlier) the color change can be used without the motion.

The intensity of the joy can be established by the number of different methods used (the

stronger ones counting more) and by the intensity with which a particular method is carried out. For example, one way of conveying slight joy would be a slightly bouncy, reasonably fast, slightly curvy rhythmic motion. Extreme emotions can be conveyed by using many methods and exaggerating each one. In addition, there are additional suggestions for extreme emotions such as trembling or exploding.

Darwin [Darwin 1872] in writing about the expression of emotions in man and animals puts forth a principle which he calls "antithesis". The idea is that if a particular facial expression or series of movements are used to express a particular emotion then the opposite emotion is convey by the opposite expression or movement. For the few cases that I have investigated this principle is very useful for conveying emotions by the kind of animation the system will produce. For example, to convey sadness we need only negate the description of joy. A reasonable way to convey sadness is to use a leaden, constrained, slow, straight, arhythmic motion, the opposite of the joyful motion described above. The suggestions of sadness can concisely be described as:

```
(SUGGESTIONS-OF SADNESS
                (NEGATE JOY
                        (EXCEPT
                         (FOR COLORS
                              (A-CYCLE
                               (BEST-WHEN (COLORS (OR DULL DARK)))
                               (STRENGTH WEAK))))))
```

The principle does hold for the use of a color cycle to convey sadness since the negation of the joyful version indicating very slow changes in a long cycle. The exception clause is necessary only because the advice as to when this is good advice differs and this method seems less likely to succeed than its joyful counterpoint. The advice as to when this is most applicable is based upon

the perhaps universal attitude towards bright as joyful and dark as sad. The difference in strength is probably due to the visually grabbing aspect of the joyful version.

In the first scene Cinderella is slightly happy and her step mother is slightly sad. The system must choose either the motion, colors, or size or some combination of these to convey this. Since the desired film should be obvious motion is chosen since it is strongly suggested and contrasts most with the typical dynamics of the characters. Since the film's complexity level is low, convention establishing is avoided and only one method is used to convey the emotion.

The remaining task in the elaboration of the initial scene is to establish the relationship between Cinderella and her step mother. We need to convey that Cinderella is dominated by, is obedient towards, and tolerant of her step mother. That her step mother hates her will be left as a tendency to hurt Cinderella in later scenes. Again the notion of suggestions from the relation type appears to be adequate. For example, to represent what the system needs to know about dominance we have:

```
(SUGGESTIONS-OF (ACTOR-1 DOMINATES ACTOR-2)
                (ANY-OF
                 (ACTOR-2 IS-OBEDIENT-TO ACTOR-1)
                 (FOR MOTION
                      (OR
                       ((ACTOR-1 PUSHES ACTOR-2)
                        (AND REPETITIVELY FAST)
                        (NECESSARY-CONDITION
                         (NEVER (ACTOR-2 PUSHES ACTOR-1)))
                        (STRENGTH MODERATE))
                       ((ACTOR-1 OBSCURES ACTOR-2)
                        REPETITIVELY
                        (BEST-WHEN (RARELY (OCCURS OBSCURING)))
                        (NECESSARY-CONDITION
                         (NEVER (ACTOR-2 OBSCURES ACTOR-1)))
                        (STRENGTH MODERATE))))
                 (FOR POSITION
                      ((USUALLY (ACTOR-1 ABOVE ACTOR-2))
                       (STRENGTH WEAK)))
                 (FOR ORDER
                      ((USUALLY (EARLIER (ACTOR-1 DOES X)
                                         (ACTOR-2 DOES X)))
                       (STRENGTH WEAK)))
                 (FOR SIZE
                      ((ACTOR-1 BIGGER-THAN ACTOR-2)
                       (STRENGTH WEAK)))))
```

Paraphrasing the above, it states that to convey that someone dominates another the dominating character can either push or obscure the other, or usually be above the other, or usually do some action first, or to be larger than the other. Except for the obscuration method these methods are common sense. They are the type of clues that observers of animal behavior look for to understand the social rankings of say a baboon troop. Obscuration may be a convention that should be established just as the color convention needs to be established to convey joy. It remains to study the reactions of viewers to decide what methods are valid without establishing conventions first.

The suggestions for conveying tolerance and obedience are:

```
(SUGGESTIONS-OF (ACTOR-1 TOLERATES ACTOR-2)
                (FOR MOTION
                     ((ACTOR-1 STAYS-WITH ACTOR-2)
                      (STRENGTH MODERATE))))

(SUGGESTIONS-OF (ACTOR-1 IS-OBEDIENT-TO ACTOR-2)
                (FIND TASK-1
                      SUCH THAT
                      (AND
                       (ACTOR-1 CAN-DO TASK-1)
                       (NOT (ACTOR-2 LIKES (ACTOR-2 DOES TASK-1)))))
                (OR
                 ((ACTOR-2 COMMUNICATES-TO ACTOR-1 (VICINITY-OF TASK-1))
                  (NECESSARY-CONDITION
                   (AND
                    (NOT (EQUAL (VICINITY-OF ACTOR-1)
                                (VICINITY-OF TASK-1)))
                    (NOT (EQUAL (VICINITY-OF ACTOR-2)
                                (VICINITY-OF TASK-1)))))
                  (STRENGTH WEAK))
                 ((SEQUENCE:
                   (ACTOR-2 DOES TASK-1)
                   (ACTOR-2 COMMUNICATES-TO ACTOR-1 UNKNOWN-MESSAGE)
                   (REPETITIVELY (ACTOR-1 DOES TASK-1))
                   (BEST-WHEN (INCREMENTAL TASK-1))
                   (STRENGTH HIGH)))
                 ((SEQUENCE:
                   (ACTOR-2 DOES (SMALL-PART-OF TASK-1))
                   (ACTOR-2 COMMUNICATES-TO
                            ACTOR-1
                            (OR (NEXT-STEP-OF TASK-1)
                                UNKNOWN-MESSAGE))
                   (ACTOR-1 DOES TASK-1))
                  (NECESSARY-CONDITION (DECOMPOSABLE TASK-1))
                  (STRENGTH MODERATE))
                 ((SEQUENCE:
                   (ACTOR-2 COMMUNICATES-TO ACTOR-1 UNKNOWN-MESSAGE)
                   (ACTOR-1 DOES TASK-1))
                  (STRENGTH WEAK))))
```

The essence of the conveyance of obedience is for the obedient character to perform some task that the other orders done. The alternatives are for the commander to either give a demo

then gesticulate, do part of the task then gesticulate, simply gesticulate or point towards the area where the task is to be performed, or to simply gesticulate. The system is faced with the same problem as Yugoslavian animators that can not have there characters communicate verbally since they want their product to be comprehensible to speakers of many different languages. Their solution is to have their characters utter sounds that while clearly non-linguistic convey the impression that it is form of communication. In an analogous manner the communication of "unknown messages" is accomplished by gesticulating, by changing in a pseudo-communicative manner the position, orientation, colors, shape, texture or size of the "talking" character. Only if the message relates to one of those aspects will it not be faked. For example, if the first method above is used then the vicinity of the task will be communicated by back and forth movements in the direction of the area.

It remains to see how these suggestions are used to convey the relationship between Cinderella and her step mother. To convey the dominance of the step mother over Cinderella the system will accept the suggestions relating to the position, size and order of acting. They all add to the obviousness of the conveyance and do not add to the complexity of the film as using two ways to convey joy would have. This new requirement that Cinderella be smaller than her step mother will resolve the conflict between "friendly"'s suggestion that she not be large and "important"'s suggestion that she be large. Since her step mother is large, Cinderella is now determined to be medium sized. All the suggestions relating to position, size and order are weak suggestions and so the system needs to follow one of the motion suggestions to make the relationship more obvious (if the desired level of obviousness were less the system would not need to rely on the motion). The choice between pushing and obscuring is difficult to make. If there were differences in the

strengths of the suggestions, the complexity of the actions, or the restrictions that are associated with the methods (by the "best-when" and "necessary-condition" clauses) then it would opt for the simplier or more obvious method. Since pushing takes less time than obscuring it is chosen since the amount of time for this scene is limited. The scene is described as long, however, since the entire film is supposed to be between four and eight minutes and there are eight other scenes many of them long the maximum time for this scene is a little more than one minute. If it were not that the conveyance of obedience also needs to be done in this scene the other motion, obscuring, would have been picked.

The time constraints also play a role in the choice of which method to portray Cinderella's obedience towards her step mother. Obviousness would suggest the strongest suggestion, i.e. the "demo" method. This requires much time, however, since the task must be performed several times. The moderately strong suggestion where the step mother does a small part of the task and Cinderella finishes is not excessively time consuming and is more strongly suggested than the others and so is chosen. It remains now only to choose a task. There are three possibilities; she can either build an object, destroy one, or push an object to another location. The last one is chosen since it takes the least amount of time and yet is decomposable. Cinderella's obedience is conveyed by the following sequence of actions:

1) Her step mother goes up to a large object

2) She slowly pushes it a small way towards some point

3) She stops and moves back and forth towards the point

4) Cinderella comes over and slowly pushes it all the way to the point, pausing occasionally.

The reason the object is large, is pushed slowly and Cinderella occasionally pauses is to convey the difficulty of the task in order to make the episode's meaning more obvious. If the variety level of the film were high then the use of pushing for both the task and the conveyance of dominance would be avoided by choosing a different task.

F) The Expansion of the Kept Apart Scene

In the next scene Cinderella is prevented by her step mother from meeting the prince. The script says that first the fact that Cinderella wants to meet the prince should be conveyed, then that she is prevented from doing so by her step mother. As a result of this she becomes very sad and her step mother is happy and proud (gloating). In addition, the scene should be long and its mood should change from hopeful to depressing and the rhythm should slow down. The reason I described the mood as changing like this is to help the audience identify with Cinderella. If the overall mood corresponds with Cinderella's happiness the viewer (if susceptible to the overall mood of a scene) will tend to feel the sadness or joy at the same times that Cinderella does. Maybe the system should know this trick to create empathy with a character and use it when requested.

There are at least two ways the system will be able to convey that a character wants to do something. One is to do a small first part of the desired action as soon as the prerequisites are satisfied. Another way is to create an imaginary scene (a fantasy or dream) of the character and as the action is completed the character becomes very happy and then the dream is over and the film returns to the "real" world. Making the transition to the dream can be accomplished in several ways; a stereotypical one being to zoom into the character and then dissolve into the dream. The dream method is very time consuming and despite the fact that this is a long scene there is the

second part to do and so the first method is chosen. The first part of this scene is now defined to be the prince making an entrance far from Cinderella and her step mother, and then Cinderella moving towards him only a small amount.

In order to convey that someone prevents some action from occuring the system will expand the action into its major components and prerequisites. The preventer then does something to cause at least one of the components to be unfulfillable. The description of someone meeting someone else is:

```
(SUGGESTIONS-OF (ACTOR-1 MEETS ACTOR-2)
                ((ACTOR-1 MOVES-TO (VICINITY-OF ACTOR-2))
                 (PREREQUISITE: (AND (ON-STAGE ACTOR-2)
                                     (VISIBLE ACTOR-2)))
                 (STRENGTH HIGH)))
```

For the step mother to prevent Cinderella from meeting the prince she can either make it so that Cinderella cannot move to the vicinity of the prince, or she can make the prince not be on stage or not visible. Looking at the methods for accomplishing any of these, the system finds that to cause someone not to be on stage the best method is to "push" them off and that this method is best when the pusher either dominates the other or is described as being stronger. This is not the case with the prince and the step mother and so does not appear as attractive as the methods for causing the prince to not be visible. If the step mother was described as "magical" then she could cause him to disappear, Since she is not, she can cover him with something that will obscure him. The other possibility, causing Cinderella to be unable to move to the prince, is much closer to the original fairy tale but is rejected due to the complexity and long duration of the methods of accomplishing this. Later we will force the system to choose this last possibility. The second part

of this scene will depict the step mother covering up the prince with a dark solid object that is larger than him. Cinderella then wanders around not finding him and she becomes sadder while her step mother becomes happy. In order to emphasis Cinderella's extreme sorrow the system will use not only "sadness"'s motion suggestions but also the color and size cycle suggestions.

To convey the mood or rhythm of a scene to system can use background color, scenery, various cinemagraphic techniques such as fast cutting, close-ups, and the like, or simply rely on the emotional state and rhythm of the characters. The last method is preferable since it is simplest however it will not work for this scene since the characters diverge greatly. The next simplest method is to use the background color which is sufficiently obvious to be acceptable. During this scene the background color will be bright and change with a fast pulse initially. Later the colors will darken the pulse will slow.

G) Cinderella Beautified Scene

After Cinderella is prevented from meeting the prince, her step mother leaves. The fairy godmother appears and her personality is conveyed. Next the system must convey that the fairy godmother causes Cinderella's physical description to change from shabby to elegant. The first set of suggestions inspected are those for conveying that an actor caused an event. Here Michotte's work on the perception of causality [Michotte 1963] is helpful. A simplified and partial description of those suggestions is:

```
(SUGGESTIONS-OF (CAUSED-BY ACTOR-1 EVENT-1)
                (OR
                 (ANY-OF
                  ((SEQUENCE:
                    (ACTOR-1 COMMUNICATES NONSENSE)
                    (HAPPENS EVENT-1))
                   (STRENGTH WEAK))
```

```
((SEQUENCE:
  (ACTOR-1 TOUCHES (MAIN-OBJECT-OF EVENT-1))
  (HAPPENS EVENT-1))
 (STRENGTH MODERATE)))
((SEQUENCE:
  (ACTOR-1 PUSHES (OBJECTS-OF EVENT-1))
  (HAPPENS EVENT-1))
 (BEST-WHEN (PART-OF EVENT-1 LOCATION-CHANGE))
 (STRENGTH HIGH))))
```

The most obvious method is pushing the objects associated with the event, however its "best-when" clause is not satisfied and so the choice the system now faces is whether to use only the "touches" method or both gesticulating and touching. The scene length is long and obviousness is high so it decides to do both.

The next problem for the system is how to change the physical appearance of Cinderella. Changes can either be abrupt or gradual, gradual changes are clearer since there is little question of whether someone disappeared and another appeared or if someone abruptly changed appearance. Since this is a long scene gradual change is feasible. The way this gradual change will be accomplished is by interpolating the earlier description into the later one. This is accomplished by discovering in a suitable form the difference between the two descriptions, and then to the first description the difference is slowly added. Interpolation between shapes is a common problem in computer graphics. I have extended these techniques to the interpolation of colors, textures, and line. The dynamics of the transition from the first to the second appearance still need to be specified and the scene constraints relating to the mood and rhythm are appropriate here. The speed with which the interpolation occurs could slowly pick up, with a rhythm that goes from depressing to joyous. To make the transition even more obvious the system will exaggerate the difference between the two appearances. To do this it need only give the suggestions of

"shabby" higher priority for the initial appearance and likewise for "elegant" in the second appearance. This will resolve many of conflicts in the partial description of Cinderella that was generated earlier. Her shape and texture will be irregular, her line jagged and her colors muddy.

H) Cinderella's Appearance Before

After all the scenes have been expanded the system will then determine the appearance of the characters to a level that the graphics routines can realize. This was postponed in anticipation of new suggestions and conflict resolutions that occur during scene elaboration. The decision to decide all conflicts in favor of shabby, for example; occured in the process of elaborating the beautification scene. The shape at this point is described as closed, asymmetrical, non-pointed, irregular, and moderately complex. This description is matched only by the "blob" procedure. Asymmetrical and irregular shapes are the forte of blob. The complexity of a blob is easily controllable. To avoid points the blob procedure need only be called with no turns greater than sixty degrees. The actual appearance of Cinderella is realized by the following procedure call:

```
(DRAW-BLOB 'CLOSED
           'ASYMMETRICAL
           'NO-CROSSINGS
           12 ;the number of "lines"
           15 ;minimum angle of any turn
           60 ;maximum angle of any turn
           60 ;minimum line length
           90 ;maximum line length)
```

Draw-blob is a graphical procedure that with controlled use of random number generation draws a blob to the above specifications. The third argument, "no-crossing", is necessary in order to give the shape a texture which is part of her description. The number of lines is twelve so that

the blob is moderately complex. The maximum angle is sixty degrees in order to keep the blob from being "pointed". The minimum and maximum line length are the result of the line being described as long. The minimum angle is the standard default for blob since there is nothing in the description to indicated its value. The low originality level permits the use of such defaults.

The line is described as being thin, curvy, long, ragged, with a gently changing width. The line drawing routine should be able to determine how to draw such lines without much difficulty. Gently changing width might be defined to grow five percent every inch for four inches followed by four inches of shrinking five percent. Thin would be defined to be only one raster point wide.

Her size is described as "medium". Simplicity and obviousness both suggest that the screen not be too cluttered, however obviousness also does not want objects to be too small to notice. Since there is no scenery and there is at the most only three characters on the screen at any one time, the system can easily pick a reasonable percentage of the screen to be the average size of the characters, in this case about five percent. Since Cinderella is smaller than the average, she would be about half this size. Since one of the reasons that she is not as large as her step mother is because her step mother dominates her (the other reason being that she is friendly), Cinderella will grow a little when her step mother is not around. This is part of "dominance"'s suggestion relating to size.

Regarding Cinderella's texture it has been decided that it be irregular and consist of many small elements. The texture expert knows that irregular textures can be generated by using a random number generator to decide for each position for the element whether to draw it or not. A typical small texture element is small squares and the lack of originality desired will make this suggestion acceptable.

The colors of Cinderella have been determined at this point to be several, harmonious, drab, with a least one pastel or bright color. Three specific suggestions that have been collected at this point are browns, tans, and white. Lack of originality suggests that all specific suggestion be accepted. Since several colors are called for three browns and three tans are added to the list. All that remains is one pastel or bright color. The requirement of harmony implies that the color not be very different from those already chosen. The pastel or bright color that best meets the requirement is orange, being related to brown and tan. Another pastel or bright color would be generated if the number of colors were not already at the high end of what "several" means. The suggestions for white and a pastel or bright color were strongly made and so white and orange will be twice as frequent as the other colors.

I) Cinderella's Appearance After

To determine Cinderella's appearance after her fairy godmother beautifies her the system must remove all descriptors of Cinderella's appearance that are based on suggestions from "shabby". It must also reconsider all conflicts that were decided in "shabby's" favor. Then suggestions from "elegant" must be added. This is preferable to the alternative method of starting from scratch deciding Cinderella's appearance as if she had originally been described as elegant and not shabby. Decisions regarding her appearance that are unrelated to the question of her shabbiness or elegance might be remade differently, causing aspects of her appearance to change that are misleading.

The suggestions of "elegance" that are taken into consideration here are:

```
(SUGGESTIONS-OF ELEGANCE
                (NEGATE SHABBY
                        (EXCEPT
                         (FOR SHAPE
                              ((COMPLEXITY-LEVEL HIGH) (STRENGTH MODERATE)))
                         (FOR LINE
                              (THIN (STRENGTH HIGH)))
                         (FOR TEXTURE
                              (ELEMENTS (SIZE VERY-SMALL)) (STRENGTH HIGH))
                         (FOR COLORS
                              ((AND VERY-MANY BRIGHT) (STRENGTH HIGH))
                              ((STEREOTYPE: RAINBOW)
                               (STRENGTH MODERATE))))))
```

In forming the new appearance description of Cinderella conflicts which are easily resolved by compromise are encountered. For example, "beautiful" suggests moderate complexity and "elegance" suggests "high complexity" so they compromise upon moderate-high complexity. Similarly for the suggestions of very many colors and several colors is compromised to many.

After beautified Cinderella's shape is described as moderately highly complex, regular, symmetrical, and closed. Stereotypical suggestions are stars, spirals and snowflakes. Spirals are ruled out since they are typically neither symmetrical nor closed. Snowflakes are not possible with long lines and Cinderella's line is described as thin, smooth, gently changing width, curving and *long*. Since the requested originality level is low, stars are acceptable. Moderately high complexity suggests a star constructed of between twenty and forty lines. The interpolation expert suggests that the number of lines in Cinderella's appearance after be an integer multiple of her appearance before. This simplifies the task of interpolating smoothly between the two appearances. Since her previous appearance is drawn with twelve lines the number of lines now should be either 24 or 36.

Twenty four is chosen since Cinderella's size is middling and in order to create perceptibly complex shapes they must be rather large. Even so Cinderella will be slightly larger than before "beautified" in order to accommodate her added complexity.

Cinderella's texture is straight forwardly implementable. It is regular and composed of small to very small (another compromise) elements. A typical small elegant shape is a diamond and is chosen in this case.

She is to have many bright colors. A stereotypical suggestion is to use the colors of the rainbow, a spectrum. Many suggests that there be from ten to twenty five colors. The interpolation expert again requests the number of colors here be a multiple of the number of colors in her shabby appearance. Her earlier appearance consisted of eight colors, however, two of those colors are used twice as often as the others making it equivalent to ten colors. The choice is now between ten or twenty colors of the rainbow. Since she is not large a large number of colors cannot "fit" into a small area and so ten colors are chosen.

If the system were extended to make use of cinemagraphic techniques such as zooms, close ups, fast montage, pans and the like then the decisions regarding the number of colors or lines in Cinderella's appearance would have been different. Since the beautification scene is important (this can be inferred from its length, maybe it should be described explicitly if these techniques are available to the system) and since the major event is Cinderella's appearance change a close up would be chosen. Since she will practically fill the screen the number of lines and colors will be decided in favor of larger numbers.

J) Criticisms of the First Pass

When the film is finally complete (or upon request a scene could be viewed as it became ready) it will be seen by the user. Ideally the hardware would be fast enough to show it fast enough. Since it is not, the movie will be filmed frame at a time and then seen. Stills from the film would be viewable without filming and the user will be able to see stills from the various scenes. Let us suppose that after either viewing the film or the stills the user has a few complaints. One very reasonable complaint would be that the scene in which the step mother prevents Cinderella from meeting the prince was portrayed in manner too foreign from the original fairy tale. Since there will be no natural language interface, the user will communicate this dissatisfaction by asking for a review of the high-level decisions made in that scene. Among them will be the decision between conveying the prevention by using the method of hiding the prince, chasing the prince off the screen, or preventing Cinderella from moving towards him. The first method was chosen, to indicate that the latter method should be used the user need only copy the internal form of that method and indicate that it be given a very high priority. If the user's criticism were more vague, then he or she should copy the form of the hiding method indicating that it be given the lowest possible priority. This would correspond to saying of the film that that scene was done poorly, that the prince should not have been hidden, without giving any alternatives.

In either case the system will begin to "re-think" the scene. If the latter form of criticism were given the system would need to choose between Cinderella being prevented from moving or her step mother chasing the prince off the stage. The later method is still inappropriate for the relationship between the prince and the step mother. The other method while more time

consuming and complex has no viable competitors and so is chosen. For a character to prevent another from moving to a particular region there are three possible methods in the system:

1) leave a trail behind that is impenetrable (a fence)

2) leave images of self behind as one circles the "prisoner" (metaphorically standing guard)

3) intercept any movements towards the "forbidden" region (actually standing guard)

4) tack or tie down the character.

The third method is ruled out by the following scenes where the step mother leaves and the fairy godmother undoes whatever is done here. The most obvious and most easily undone method is the first and so it is chosen. This change will propagate to the scene where the fairy godmother undoes the prevention. Previously she revealed the prince's locations, now she will follow the fence erasing as she goes.

Other sorts of criticism are possible. For example, the user may not be pleased with the blob-like appearance of Cinderella in the first few scenes. Suppose the user thinks that the line, colors, and texture are sufficient to convey the transition from shabby to elegant and want to disable "shabby"'s suggestions regarding shape. This will result in her shape being regular, symmetrical and moderately complex. This will lead to a star, but a simplier one than is used after she is "beautified".

III. Control Structure

The basic process the system will be engaged in is making a series of decisions to refine or substantiate a structure initially given by the user. The choices the system makes are usually only partially constrained by the kinds of considerations already mentioned. Even when the structure provided by the user is relatively specific the system is still left with apparently arbitrary decisions such as whether a particular object will be a hexagon or octagon, red or orange, or move slowly or very slowly.

One can imagine a whole range of strategies for ordering the refinements or decisions the system will make. The system could operate in a top-down, breadth-first fashion, refining on each pass only that which is necessary. The advantage here is that coherence will be greater. The other extreme, a depth-first exploration would correspond to refining one branch down to the level of object specifications and then expanding the next branch with the recently added constraints. The problem with this approach is that the elements defined earlier will tend to be arbitrary while the later ones might be too constrained, even to the extent that back-up is necessary.

My current tendency is towards the former approach where decisions are made only when necessary. With either approach there remain many ordering decisions. Should certain aspects be developed first, such as motion and size, and only after they have been developed the other aspects such as shape, color and texture be expanded? Should one character be developed completely first based on some ordering of the aspects and then the others expanded. Or should all the characters' motion be decided, then their size, then their shape, etc.? Should the order vary depending upon the structure constructed so far and the desired emphasis?

This searching through a large space may remind one of earlier artificial intelligence programs. Decision trees were constructed and search heuristics applied to solve problems. The paradigm has subsequently shifted from search-oriented systems to knowledge-oriented systems. Why then am I describing the problem as one of searching through a very large space of possible shapes, colors, sizes, and motions for an object rather than knowing what to do? With art there are no very well-defined goals but there are many possible "solutions." Knowledge of all sorts is being used to make these decisions but the choices are still under-constrained. The best methodology that I can think of in such cases is to make those decisions first that are the easiest, the most constrained, or the ones most likely to provide reasonable constraints for the rest of the decision process. It is hoped that there will be little interaction between the more specific decisions, so that the color of a friendly object can be chosen without regard to its size. The coherence of the object will be the result of the more global constraints and choices. The size and color, for example, of an object need only be based on considerations such as the object is intended to be friendly, undistinguished, that the predominating colors in the film are blues and greens, and that the screen is never to get very cluttered. The determination of the size, however, need not be based on the color and conversely, so that interactions of the low level aspects can be avoided.

IV. Numbers, both Real and Random

This may be a minor or incidental point but I hope to avoid as much as possible using numbers, both random ones and garden variety ones. My objections against numbers are partly phenomenlogical, that is they don't feel right; and mostly computational, that is much economy can be gained by restricting the system to the following "numbers": "a very little bit", "a little bit", "a little", "a medium amount", "a lot", "an awful lot". These "numbers" would be used in the descriptions of motions, sizes, distances, and intensities. The advantages of such numbers is that they permit assertions to be made about the numbers themselves (such as consistently using the number "an awful lot" will result in a wild, confusing film), enable crude cross-aspect comparisons (hopefully "a little bit" of motion will "correspond" to "a little bit" of size), and ease the task of making decisions relative to the context ("a lot" of size corresponds to a smaller number if all the other objects are small or if the entire screen is small). Of course, at some stage these numbers will need to be converted to ordinary numbers and relatively simple procedures could accomplish this.

My objections against random numbers are based on a belief that good art is not arbitrary, but the product of thought. One reason I dislike producing art with random numbers is that to my knowledge the only other attempts to create automatic picture generating systems relied heavily on random number generators. Good art is carefully constructed and thought out, usually each element is there for a reason. A random number generator is as likely to produce a "Mona Lisa" as are a million monkeys to type "Hamlet". This is not to say that there are not inconsequential decisions that have to be made that could just as well be made by the toss of a coin. In such cases, however, one can just as easily ignore the random number generator and produce as good a film.

Also, I am not saying that there is no place for noise in art. A slightly perturbed square may in some cases be more desirable than a perfectly drawn square. Some very useful texture patterns are basically different noise levels. A drunkard's walk may be disturbed by random perturbations. These uses of randomness, however, are very controlled. A drunk does not walk forward a random amount, then right or left some random angle, and then repeat. There is a basic pattern that is disturbed but not destroyed by randomness. If, as in the earlier example, the slightly perturbed square is desired then there must be a reason, perhaps to make it look more like a square drawn by a human in a hurry. Again in such cases, the random perturbations must be very controlled. Also any reliance upon randomness will make it much more difficult to demonstrate and test the usefulness of the knowledge in system.

V. The Structuring of the Space of Each Aspect of an Object

When the system is determining the speed, path, shape, color and texture of each object, its task would be greatly simplified if each aspect were properly structured. Intuitive spaces are desirable, so are standardized ones. The way the space is structured or described greatly influences the ease of maneuvering and reasoning within it.

A) Size

Size is the easiest aspect to structure. It can be standardized by using something like the greatest inscribed circle, the least circumscribed circle or their average. This way a square of size "a lot" will be roughly the same as a line of size "a lot". This approach contrasts with the way the size of an object is usually controlled in computer graphics. A description of an object is given in terms of its coordinates or as horizontal and vertical changes from the last position of a "pen". This object can then be scaled. A standard turtle procedure for drawing, say polygons, is one which one controls not the absolute size but only the length of each side. These alternatives do not facilitate inter-object size comparisons and so are inadequate for the system's use.

B) Speed

Here I could borrow from physics and characterize the motion of an object during a scene by its velocity and acceleration. Were I to adhere even more strongly to mechanics I would merge speed with path and deal only in velocity vectors. This does not, however, match the way I conceptualize motion. Often the speed and path need to be modified separately (the path may be right, but the speed too slow, for example). Perhaps the speed will often change in complex ways

as a function of time. It may also be important to quantize speed and acceleration so that motion descriptions do not become too complex.

C) Path

Here there are two problems combined; what the position of an object is and what course it will take. The position of an object in a standard turtle program is not too easy to control, one can really only control where the turtle starts. I expect to use some rule of thumb such as consider the position of an object as a point that corresponds to the "center of mass" of the object. The "center of mass" can easily be approximated by the average position of all the vertices of the object. This makes it easy to center one object inside another, to have one object circle another, etc.. Other special points need to be associated with objects, such as a point about which an object grows or turns. Center of mass provides a convenient default.

The course or path of the center of the object could be characterized as a two dimensional cartesian space or as a list of turtle commands. While I may use the latter I think a richer more structured, more intuitive space is needed. A path can be characterized as either smooth or jagged, long or short, closed or not, symmetrical or asymmetrical, etc..

D) Shape

I hope that shape and path can share the same descriptive mechanism.

E) Color

From what little I've read it seems that an intensity-hue-saturation space for color is more intuitive and closer to our perceptual mechanism (for example, linear interpolation between colors

appears more linear) than a red-green-blue space. It is difficult to think of colors in terms of how much red, green or blue they have. It is usually easier to think of a color as being so bright or intense, as being so saturated (i.e. how much white is in the color), and its hue.

F) Texture

The textures I plan to have available to the system are few. One kind is simply different noise levels. Another is repeated shapes or simple patterns such as show every other point, or every other horizontal line. Certain aspects can be characterized such as a texture's intensity (the percentage of points turned on in the area), its orientation, its density (i.e. the distance between elements relative to their size) or its complexity. Part of the characterization of a texture built by the repetition of a shape would be the characterization of the shape itself.

G) Other aspects of objects

There are other aspects of objects that I have not mentioned. Orientation and rotation are important and probably are similar to position and speed. The line that delineates an object can also be controlled for various effects. It can have varying width or color. It can be smoothed. Dotted and dashed lines are also possible.

VI. Expectations of the viewer

Playing with the viewer's expectations is important. Little is worse than a very predictable film. Good films create expectations, stretch them, break them at crucial moments or satisfy them in unexpected ways. Expectations create a point of focus that can be used to the filmmaker's end. Many of the expectations are dependent upon the sophistication and the idiosyncrasies of the viewer. Whether the system will be capable of playing with expectations or this difficult task will be left to the user who creates the initial structure is a question that requires more thought.

VII. Evaluation

There are at least two different aspects of this research that entail tricky questions of evaluation. The first is whether the films the system produces are any good and connect in some way to the initial structure that inspired them. The other is concerned with how the system should be evaluated as a contribution to AI, computer graphics, animation or aesthetics regardless of the quality of the films.

A) Evaluation of generated films

Originally I conceived of the task of the system as the generation of films that, in addition to being good, followed the script accurately. Two different experiments could then be used to test this accuracy. One would be to have the viewers describe the film without knowing what its intention was and then comparing the viewers' descriptions with the original script. The other would be to tell the script, show the film, and then ask for criticism. I still plan to carry out these experiments, however I don't think they are crucial to the evaluation of the films. The feedback from the viewers will hopefully be very useful. If the responses to a group of films is positive, consistent but not what was intended then presumably the system could be modified to produce those kinds of films when the effect they created is desired and to try something else to produce the kind of film originally desired. It should be kept in mind, however, that the system's main goal is to produce *good* films, not simply films that match the script.

But how are we to know if the films are any good? Some very sufficient but unnecessary criteria would be, winning an academy award, getting rave reviews from critics, or being a great box office hit. A more likely criterion would be to receive favorable criticism from "experts", in

particular, human animators. Failing this the evaluation can fall upon the reader. The thesis is the point, the films only the product. The reader of the thesis is constantly evaluating the research and is likely to consider him or herself enough of an expert to judge the films. If this all fails there is still one leg of defense to stand on. The system could be based upon a very good model of the generation of animated films and due to its limited knowledge produce lousy or childish films. This argument should be avoided since it depends upon faith on the readers' part that ignorance is its only flaw.

B) Evaluation of the system

Were the system to produce award-winning films, I would be very happy, however, I would not consider that proof of the system's worth. It might have produced the film in a very *ad hoc* and limited manner. The quality of the reasoning that the system uses to create the films is much more important than the final product. Additionally, I would hope that the system would at least remember each major decision it makes and a justification for it. I do not intend to build any self-explanation abilities into the system, but it will remember in an accessible form the information that such abilities would depend upon. Despite my decision to limit the expertise of the system to narrative, object-oriented, non-representational films, I expect the system to be rather versatile. It should be able to respond to a wide variety of scripts. It should be able to re-do a film with additional constraints. (Such as "make the film a little bit more sedate and calm" or "make it shorter and lighter".)

Even in the case that the films it produces are lousy, it would not constitute proof that the system is lousy. The system might be based on very sound and general principles, its reasoning

may be very clear, it may be very versatile, and still produce dull, ugly, or silly films. If this were the case and it seemed plausible that its flaws were in the content of its data base, which was either too limited or just plain incorrect, then the system could still be considered a success. One should remember that the primary purpose of implementing such a system is to demonstrate, via a working example, a computational theory of animation. To lessen the chance that its knowledge base will be inadequate, however, I have been, and intend to continue, becoming an expert, i.e. an animator. Towards this end filmmaking is my minor, and I have made several films both in a conventional manner and by computer.

Another hard question is how to evaluate the construction of a working system in terms of possible contributions to artificial intelligence. What is the general applicability of a system designed to produce a very limited class of animation? The problems of exploring a very large space with such poorly defined goals and under-constrained choices have been not been investigated in AI to my knowledge. I claim that it is an important cognitive activity that occurs in situations other than the one discussed here, generating art. Presumably living one's life is an example. The goal's are poorly defined, part of the problem is generating the goals (as is also probably the case with art), and the choices that one has as to what to do next are not sufficiently constrained by the reality of the situation or one's goals.

The following is a list of other ways in which this research might contribute to AI:

1) by providing a computational view of aesthetics. While aesthetics is probably not as important as other areas being explored by AI researchers such as understanding natural language, it is an important consideration in research, design, programming and many other endeavors.

2) by constructing a computational epistemology for a real world endeavor by cataloging and structuring the knowledge of an animator.

3) by representing, structuring, and cataloging adjectives (such as friendly, sharp, red, etc.). While the use the definitions of adjectives will be generative, hopefully the extension of the definitions to be useful for interpretation will not be great.

4) by providing a better understanding of what it means to be creative or original

5) by testing Carl Hewitt's actor formalism [Hewitt 1975] on a big, real problem. The mapping of the objects in my system into actors seems straightforward, and the experience will help answer questions about the usefulness of actors.

6) by providing a better grip on how to manipulate, exaggerate, negate and transform stereotypical descriptions.

VIII. Objections

I am well aware that in committing myself to this research I am treading dangerous waters. Computer scientists typically feel that aesthetics and animation have little to do with valid *computer science* research. Many objections of a completely different sort are likely from the art world. A computational theory of animation appears to be at variance with commonly held views of art as intuitive, imaginative, unpredictable, non rule-based, and requiring the creator to have a soul or innate talent. I think that an AI view of these problems would be very enlightening to the art world and should be formulated. Hopefully a working system will be very useful in meeting these kinds of objections and more.

A) Lack of perceptions

The system will be perceptually limited to input from a key-board. It will be blind both to the world and its own films. One may wonder how such a perceptually-limited system could be expected to produce visual art. A partial answer is that perception is a process of creating symbolic structures from external stimulus and therefore the system needs only those symbolic structures and not a perceptual mechanism. Its knowledge base will, to a large extent, be a condensed and simplified version of the structures that I have built up over the years as a perceiver. There are historical cases of great artists or musicians who became blind or deaf and yet were able to continue producing great works of art or music. Nonetheless, it is quite useful to see what one has produced. The system will maintain a symbolic structure describing the film and its images as it creates them. This structure will compensate for most of the perceptual deficiencies of the system. This will not completely compensate however, for there are many cases were a creator does not

notice some feature or interaction until the product is created. For example, one film I made had a very complex shape growing and turning while simultaneously the colors were changing in subtle ways. The color changes turned out to be too subtle -- I still have troubles convincing people that they are really happening. The only solution of the need for feedback is to rely on a human observer, who gives perceptual criticisms. The system would then only need the ability to take such criticism into account, in this example by exaggeration of the aspect that became overwhelmed and down playing the other components.

B) Everyone sees differently

Another objection to the system is concerned with the perceptions of the viewer. The objection is that each individual perceives a film differently, so it becomes problematic as to what model of the viewer the system should have. Much of the knowledge of the system consists of facts such as "red is friendly", "square is dull", "curving, smooth motion is happy" etc.. These are culturally-dependent generalizations that occasionally fail due to the viewer's idiosyncratic experiences and associations. There are at least two retorts to this objection. One is that every film or work of art suffers from the same problem and yet there is enough of a common ground of associations, perceptions, and experiences for the work of art to communicate and affect its viewers. It is upon this common ground that the system's knowledge will be based. The other response to this objection is that the construction of idiosyncratic knowledge-based systems is not inconceivable. Each viewer could modify the knowledge base according to his or her own feelings and associations and then let the system produce the film. A knowledge-based system should be capable of conforming to each individual viewer's subjective preferences and personality.

C) A generative theory of art is not an interpretive theory

One may object that the proposed system will exemplify the process of generating art, and as such has little to contribute to the process of interpreting, perceiving, enjoying, or analyzing art. This objection may well hold true, it remains to be seen. It is more likely, however, that an understanding of generation will contribute greatly to understanding perception. Perhaps much of the knowledge base is used in a similar manner when interpreting as when generating. For example, the fact that "red is warm" may be used both in generating "warm" objects and in interpreting red objects as "warm". The mechanism which is used to generate coherent, rhythmical films is probably similar to that which perceives the same features. Also the system will in some sense be able to "interpret" the films it itself makes as organized along certain lines, as using certain techniques, as conveying a particular mood, etc..

Another defense of approaching the problem via generation rather than interpretation is that its easier and still answers many interesting questions about both the representation of knowledge and aesthetics. This research project is already very ambitious, to expect it to contribute to interpretation is unreasonable.

D) But it will be mechanical and predictable

A common criticism of films or art objects is that they appear too mechanically constructed, that they are too predictable. This is an especially common criticism of computer generated films. The motion is too smooth, it has little character. The shapes are too exact, they too lack character. The computer films that are exceptions to this are either driven by pseudo-random number generators or by humans. The former are often structureless and lack coherence since too many

decisions were made by "tossing a coin". The latter are often fine since they avoid being mechanical by having humans graphically communicate the motion or shape to the computer.

My system will hopefully avoid this difficulty by a process similar to that which a human goes through. The motion it will give objects will typically have character and complexity, unless of course it is trying to convey a certain mechanical personality of an object. Basic patterns of motion will be combined and modified to achieve a very definite character such motions are not likely to be mistaken as mechanical.

Some might complain that being deterministic the system will be too predictable. It will be deterministic in the sense that, unless told to behave differently, it will produce the same film for the same script. The system, however, will remember the decisions it made in making the first film and, if so instructed, will try to make another that is significantly different from the first. One may still object that the system will be predictable, if one asks for a warm object it turns out red or orange, never black or blue. The system should be able to inject an unexpected or unpredictable element into a film without the use of random number generators. It could establish a pattern and then break it at a critical point. It could decide to use one aspect of an object as an odd distinguishing feature. For example, it might give a very pleasant friendly object a large pointed shape (presumably an unfriendly shape). Here the system would depend upon the motion, color, line, and texture to override the size and shape. If possible, the system will choose which feature or features to make anomalous for a character on some positive basis not only because it is expected that the intended character portrayal will not be weakened by the anomalous features. As in the example of the large pointed friendly object (a friendly giant), the atypical size and shape for friendly objects may have been chosen because the character was supposed to be a strong

friendly ally of another object. One of the goals of the system is that it have a good reason for every choice it makes.

E) Its got no soul, no imagination, no feelings

Many would object that the system could not produce good art since it lacks either soul, imagination or emotions. Soul is the hardest to contend with since I don't really know what it is. I expect the system will appear to have a rather distinct personality as a filmmaker, but does that have anything to do with soul?

Imagination is almost as tricky, the system will probably appear to be imaginative, to come up with "creative" solutions to problems but some people demand more than appearances. Will it really be creative or simply doing what I told it? I think the question is based upon a mistaken notion of creativity and I expect the system will do both.

Some people think that when an artist paints a depressing scene that he or she must be feeling depressed while painting it. It might indeed help to be feeling in a particular mood to create a consistent and strong visual representation of it. The system need not "feel", however, to *behave* the same as people who do.

F) Once you reduce it to a rule its no longer Art

People who say this typically do not understand the computational complexities of a procedural system using a knowledge base. There is probably truth in this objection if one thinks of a simple rule, or a few weakly interacting rules. If the process is too simple, too predictable, then indeed it is not likely to produce good art. The system I am proposing, however, will be based on a very large set of "rules" that interact in complex ways.

G) Need to understand psychology too well

It might be the case that in order to produce good art one needs a very good understanding of human psychology. One must understand how percepts affect people, how different kinds of expectations are created, etc.. The system will have some knowledge about this but it is possible that it will be inadequate. The system may be too ignorant to function properly, or maybe not. Much of the system's knowledge will be useful facts and techniques that are based upon an empirical understanding of human psychology. At this point I see no other way to resolve the question of whether this will be sufficient without actually attempting to build such a system.

H) Art is not well enough understood

There are two senses of this objection that may be true. One is that I don't understand art well enough and must be an artist or study art for many years to even attempt to form a theory of animation. I am trying to reduce this deficiency as best I can, by reading, by making films, by going to art museums and screenings, by talking to animators, etc. but it may be that this will prove inadequate. It is common in AI, however, for researchers who are not experts in language, medicine, electronics or perception to still try to apply AI to those fields.

A more serious problem would be the interpretation of this objection, as *no one* knows enough about art, at least in an explicit form, to attempt such a project. Many people believe that the skill to produce art is not understood, some even venture to claim that it is not capable of being understood. If, however, I do manage to create a convincing system then I would have made a contribution to the thesis that the knowledge involved in producing and appreciating art is symbolic and not very different from the other kinds of knowledge that people possess.

I) Good Art combines very many subtle aspects

One interpretation of this objection reduces, once again, to the possibility that I don't (or can't) know enough to tell the system enough to perform adequately. We'll see.

It should be clear, however, that the system will be capable of subtle effects. It might make the difference, in say color, between two objects very small, just barely perceptible. Or it may want to subtly create a schizophrenic character by down playing the obvious schizophrenic behavior and by avoiding obvious techniques such as quickly changing the shape or color of the character.

IX. Summing up

A system has been proposed that I hope will contribute to AI, animation, aesthetics, and computer graphics. The system will be based upon a computational view of generating art that emphasizes the need to structure the necessary knowledge so that it can be used in a very flexible manner and be easily modified and manipulated by the system. The emphasis will be on the facts, generalizations, rules, and procedures. What they are, how they should be represented, how they should be organized, etc.. This is not to say that no attention will be devoted to trying to discover very general and powerful principles or to heuristics for searching through the very large decision space,[5] only that I consider the problem of getting the needed knowledge into the system to be more important.

I keep seeing strong analogies between this research and research on natural language understanding. Linguists traditionally were interested primarily in the syntax and generation of language. They lacked a sufficiently powerful view of computation to handle adequately problems of semantics and meaning. When a computational view was applied to language by AI researchers the emphasis shifted towards the understanding and interpretation of utterances. Art historians and theoreticians have, for the most part, studied the problems of interpretation and criticism of art. They have failed, or more commonly not even tried, to formalize the knowledge or discover the rules behind the generation of art. Perhaps, in the same way that a computational view of language has aided linguistics and furthered the state of artificial intelligence, a computational view of aesthetics will aid in the understanding of art and artificial intelligence.

---

5 The notion of many sources of suggestions, sometimes competing and conflicting, sometimes reinforcing each other, and sometimes merging together is an example of a general technique that may be developed in the course of this research.

Comments of any sort very welcome.

I wish to thank the following people for comments and criticisms of earlier drafts of this paper: Carl Hewitt, Seymour Papert, Henry Lieberman, Marilyn McLennan, Peter Szolovits, J. C. R. Licklider, Mike Genesereth, Bruce Schatz, Bill Kornfeld, Marvin Minksy and Ira Goldstein. Most of the research reported herein was done while the author was supported by an IBM fellowship.

Bibliography

[Baecker 1969]
Baecker, R.
"Interactive Computer-Mediated Animation"
MIT EE Ph.D. Thesis 1969 MAC-TR-61

[Berlyne 1971]
Berlyne, D.
Aesthetics and Psychobiology
Meredith Corporation, New York, 1971

[Berlyne 1972]
Berlyne, D.
"Affective Aspects of Aesthetic Communication"
pp. 97-118
in Communication and Affect - A Comparitive Approach
ed. Alloway, T., Krames L., and Pliner P.
Academic Press, New York, 1972

[Franke 1971]
Franke, H.
Computer Graphics Computer Art
Phaidon Publishers, New York 1971

[Gips 1974]
Gips, J.
"Shape Grammars and their Uses"
Stanford AI Lab AIM-231, March 1974

[Goldberg 1974]
Goldberg, A.
"Smalltalk and Kids -- Commentaries"
Learning Research Group, Xerox Palo Alto Research Center, 1974 Draft

[Goldberg 1976]
Goldberg, A., Kay A. editors
"Smalltalk-72 Instruction Manual"
The Learning Research Group, Xerox Palo Alto Research Center, March 1976

[Goldstein 1974]
Goldstein, I. P.,
Understanding Simple Picture Programs,
MIT AI Laboratory AI-TR-294, September 1974

[Goldstein 1975]
Goldstein I., Lieberman H., Bochner H., Miller M.
"LLOGO: An Implementation of LOGO in LISP"
MIT-AI Memo 307, March 4, 1975

[Goldstein 1976a]
Goldstein I., Papert S.
"AI, Language and the Study of Knowledge"
MIT-AI Memo 337, 1976

[Goldstein 1976b]
Godlstein, I., Abelson H., Bamberger J.,
"LOGO Progress Report 1973-1975"
MIT-AI Memo 356, March 1976

[Gombrich 1961]
Gombrich, E.
Art and Illusion - A Study in the Psychology of Pictorial Representation
Princeton University Press, Princeton, New Jersey 1972

[Halas 1974]
ed. Halas, J.
Computer Animation
Hastings House, New York, 1974

[Hewitt 1975]
Hewitt C., Smith B.
"Towards a Programming Apprentice"
IEEE Transactions on Software Engineering SE-1, March 1975

[Kahn 1976]
Kahn, K.
"An Actor-Based Computer Animation Language"
Proceedings of the SIGGRAPH/ACM Workshop
on User-Oriented Design of Computer Graphics Systems
Pittsburgh, Pa., October 1976, in press

[Kahn 1977]
Kahn, K.
"Three Interactions Between AI and Education"
Machine Representations of Knowledge
ed. Elcock, E., Michie D.
Ellis Horwood Ltd and John Wylie (in press)

[Leavitt 1976]
Leavitt, R. editor
Artist and Computer
Creative Computing Press, New Jersey, 1976

[Michotte 1950]
Michotte, A.
"The Emotions Regarded as Functional Connections"
Feelings and Emotions
ed. Reymert M.
Hafner Publishing Company, New York, 1967

[Michotte 1963]
Michotte, A.
The Perception of Causality
Basic Books, Inc., Publishers, New York, 1963

[Minsky 1975]
Minsky, M.
"A Framework for Representing Knowledge"
The Psychology of Computer Vision
ed. Winston, P.
McGraw-Hill Book Company, New York, 1975

[Negroponte 1977]
Negroponte, N.
"The Return of the Sunday Painter or The Computer in the Visual Arts"
Future Impact of Computers and Information Processing
Dertouzos, M. and Moses J. ed.
in press 1977

[Newman 1973]
Newman W.
Principles of Interactive Computer Graphics
McGraw-Hill, New York, 1973

[Papert 1971a]
Papert S.
"Teaching Children Thinking"
MIT-AI Memo 247, October 1971

[Papert 1971b]
Papert S.
"Teaching Children To Be Mathematicians vs. Teaching About Mathematics"
MIT-AI Memo 249

[Parke 1972]
Parke, F. I.
"Computer Generated Animation of Faces"
Univ. of Utah Report CSc-72-120, June 1972

[Russett 1976]
Russett, R. and Starr C.
Experimental Animation, An Illustrated Anthology
Van Nostrand Reinhold Company, New York 1976

[Schank 1975]
Schank, R.
Conceptual Information Processing
North Holland Publishing Company 1975

[Simmons 1975]
Simmons, R.
"The Clowns Microworld"
Theoretical Issues in Natural Language Processing
ed. Schank R. and Nash-Weber B. June 1975

[Smith 1975]
Smith B. and Hewitt C.
"A Plasma Primer"
MIT-AI Working Paper 92, October 1975

[Sutherland 1963]
Sutherland, I.
"Sketchpad: A Man-Machine Graphical Communication System"
MIT Lincoln Lab TR-296 1963

[Tilson 1976]
Tilson, M.
"Editing Computer Animated Film"
University of Toronto, Technical Report CSRG-66, January 1976

[Tinbergen 1951]
Tinbergen, N.
The Study of Instinct
Oxford Uniersity Press, London, 1969

[Tolstoi 1898]
Tolstoi, Leo
Johnston, C., translator
What is Art
Henry Altemus Co, Philadelphia 1898

[Weir 1975]
Weir, S.
"The Perception of Motion: Action, Motives and Feelings"
Progress in Perception
Department of Artificial Intelligence Research Report No. 13
University of Edinburgh, December 1975