The Six Elements and the Causal Relations Among Them

Brenda Laurel

One of Aristotle’s fundamental ideas about drama (as well as other forms of literature) is that a finished play is an organic whole. He used the term organic to evoke an analogy with living things. Insofar as a whole organism is more than the sum of its parts, all of the parts are necessary for life, and the parts have certain necessary relationships to one another. He identified six qualitative elements of drama and suggested the relationships among them in terms of formal and material causality.¹

I present Aristotle’s model here for two reasons. First, I am continually amazed by the elegance and robustness of the categories and their causal relations. Following the causal relations through as one creates or analyzes a drama seems to automagically reveal the ways in which things should work or exactly how they have gone awry. Second, Aristotle’s model creates a disciplined way of thinking about the design of a play in both constructing and debugging activities. Because of its fundamental similarities to drama, human-computer activity can be described with a similar model, with equal utility in both design and analysis.

Table 38.1 lists the elements of qualitative structure in hierarchical order. Here is the trick to understanding the hierarchy: Each element is the formal cause of all those below it, and each element is the material cause of all those above it. As you move up the list of elements from the bottom, you can see how each level is a successive refinement—a shaping—of the materials offered by the previous level. The following sections expand upon the definitions of each of the elements in ascending order.

Enactment

Aristotle described the fundamental material element of drama as “spectacle”—all that is seen. In the Poetics, he also referred to this element as “performance,” which provides some basis for expanding the definition to include other senses as well. Some scholars place the auditory sense in the second level because of its association with music and melody, but, as will be seen in the next section, it is more likely that the notion of melody pertains to the patterning of sound rather than to the auditory channel itself.

One difference, probably temporary, between drama and human-computer activity is the senses that are addressed in the enactment.² Traditionally, plays are available only to the eyes and ears; we cannot touch, smell, or taste them. There are interesting exceptions. In the 1920s, for instance, director David Belasco experimented with using odors as part of the performance of realistic plays; it is said that he abandoned this approach when he observed that the smell of bacon frying utterly distracted the audience from the action on stage. In the mid-1960s, Morton Heilig invented a stand-
<table>
<thead>
<tr>
<th>Element</th>
<th>In Drama</th>
<th>In Human-Computer Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action</td>
<td>The whole action being represented. The action is theoretically the same in every performance.</td>
<td>The whole action, as it is collaboratively shaped by system and user. The action may vary in each interactive session.</td>
</tr>
<tr>
<td>Character</td>
<td>Bundles of predispositions and traits, inferred from agents' patterns of choice.</td>
<td>The same as in drama, but including agents of both human and computer origin.</td>
</tr>
<tr>
<td>Thought</td>
<td>Inferred internal processes leading to choice: cognition, emotion, and reason.</td>
<td>The same as in drama, but including processes of both human and computer origin.</td>
</tr>
<tr>
<td>Language</td>
<td>The selection and arrangement of words; the use of language.</td>
<td>The selection and arrangement of signs, including verbal, visual, auditory, and other nonverbal phenomena when used semiotically.</td>
</tr>
<tr>
<td>Melody (Pattern)</td>
<td>Everything that is heard, but especially the melody of speech.</td>
<td>The pleasurable perception of pattern in sensory phenomena.</td>
</tr>
<tr>
<td>Spectacle (Enactment)</td>
<td>Everything that is seen.</td>
<td>The sensory dimensions of the action being represented: visual, auditory, kinesthetic and tactile, and potentially all others.</td>
</tr>
</tbody>
</table>

Table 38.1. The six qualitative elements of structure in drama and in human-computer activity.

alone arcade machine called Sensorama, which provided stereoscopic filmic images, kinesthetic feedback, and environmental smells—for example, on a motorcycle ride through New York City, the audience could smell car exhaust fumes and pizza. Sensorama's problem was not that it addressed the wrong senses; it simply happened at a time when the business community couldn't figure out what to do with it—pinball parlors were monolithic, and it would be several years before Pong kicked off the arcade game industry.

At the same time that Heilig was thinking about multisensory arcade games and movie theatres, the development of new genres of participatory theatre accelerated. Such artists as Judith Melina and Julian Beck of the Living Theatre, Robert Wilson, Peter Brook, Jerzy Grotowski, and John Cage experimented with performances that began to dissolve the boundaries between actors and audience by placing both in the same space. Wilson, Cage, Josef Svoboda, and others produced works that integrated filmic and photographic images, musical instruments, and machines in novel ways.

In the 1980s, these trends toward increasing the sensory dimensions of audience participation gave rise to works where the audience could touch the actors and scenery and move about freely in the performance space. For example, in *Tina and Tony's Wedding*, a contemporary "interactive" play, the audience is invited to follow the actors around from room to room (kinesthetic), to touch props and sit on the furniture (tactile and kinesthetic), and to share in a wedding banquet (taste and smell). Another notable example is Chris Hardman's Antenna Theatre, where audience members move around a set prompted by taped dialogue and narration heard through personal headphones. A spate of site-specific interactive plays and "mystery weekends" in the late 1980s enjoyed a fair amount of commercial success. Contemporary performance art shares many of the same origins.

It is interesting that the development of this theatrical genre has been concurrent with the blossoming of computer games as a popular form of entertainment, and I speculate that computer games have in some ways served as a model for it. In fact, it is in the areas that dramatic entertainment and human-computer activity are beginning to converge that pan-sensory representation is being most actively explored. When we examine that convergence, we can see ways in which human-computer activity has evolved, at least in part,
as drama's attempt to increase its sensory bandwidth, creating the technological siblings of the kind of participatory theatre described above.

The notion of "interactive movies," which has gained popularity in the late 1980s, has its roots in both cinema and computer games, two forms that combine theatre and technology. Earlier works were relatively isolated. These include the productions of Lanterna Magica in Czechoslovakia and an "interactive movie" that was shown in the Czech Pavilion at the 1967 World Expo in Montreal, Canada, in which the audience was allowed to influence the course of the action by selecting from among several alternatives at a few key points in the film (however, it is rumored that all roads led to Rome—that is, all paths through the movie led to the same ending). The idea of interactive movies has been rekindled and transformed into a bona fide trend by advances in multimedia technology. Likewise, there were early experiments in interactive television in the mid-1970s (such as the failed Warner QUBE system). Interactive TV had to await similar technological advances before finally becoming a 1990s buzz-word.

In drama, the use of technology to create representations goes at least as far back as the mecanique of the ancient Greeks. Cinema as a distinct form diverged from drama as the result of the impact of a new performance technology on form, structure, and style. In complementary fashion, computer games can be seen to have evolved from the impact of dramatic ideas on the technology of interactive computing and graphical displays. Computer games incorporate notions about character and action, suspense and empathy, and other aspects of dramatic representation. Almost from the beginning, they have involved the visual, auditory, and kinesthetic senses (you need only watch a game player with a joystick to see the extent to which movement is involved, both as a cause and effect of the representation).

At the blending point of cinema and computer games are such new forms as super-arcade games like Battle Tech and sensory-rich amusement park installations like Star Tours. These types of systems involve the tactile and kinesthetic senses; some are investigating the inclusion of the other senses as well through both performance technology and direct stimulation to the nervous system [Rosen and Gossler, 1987]. "Virtual reality" systems increase intensity through techniques described as sensory immersion—instead of looking at a screen, for instance, a person is surrounded by stereoscopic sounds and visual images delivered through earphones and "eyephones." Through the use of special input devices like specially instrumented gloves and suits, people may move about and interact directly with objects in a virtual world. Interestingly, the first virtual reality systems and applications were developed for nonentertainment purposes like computer-aided design, scientific visualization, and training. Home computers and home game systems are not far behind these expensive, special-purpose systems in their ability to deliver multisensory representations.

The element of enactment is composed of all of the sensory phenomena that are part of the representation. Because of the evolutionary processes described above, it seems appropriate to say that enactment can potentially involve all of the senses. These sensory phenomena are the basic material of both drama and human-computer activity; they are the clay that is progressively shaped by the creator, whether playwright or designer.

**Pattern**

The perception of patterns in sensory phenomena is a source of pleasure for humans. Aristotle described the second element of drama as "melody," a kind of pattern in the realm of sound. In the Poetics he says that "melody is the greatest of the pleasurable accessories of tragedy" [Poetics, 1450b, 15–17]. The orthodox view is that "spectacle" is the visual dimension and "melody" is the auditory one, but this view is problematic in the context of formal and material causality. If the material cause of all sounds (music) were things that could be perceived by the eye (spectacle), then things like the vibration of vocal cords and the melodies of off-stage musicians would be excluded. On the contrary, all that is seen in a play is not shaped solely by the criterion of producing sounds or music (although this may have been more strictly true in the performance style of the ancient Greeks than it is today). The formal-material relationship does not work within the context of these narrow definitions of music and spectacle.

In the previous section, we have already expanded spectacle into all sensory elements of the enactment. The notion of melody as the arrangement of sounds into a pleasing pattern can be extended analogically to the arrangement of visual images, tactile or kinesthetic sensations, and probably smells and tastes as well (as a good chef can demonstrate). In fact, the idea that a pleasurable pattern can be achieved through the arrangement of visual or other
sensory materials can be derived from other aspects of the Poetics, so its absence here is something of a mystery. Looking “up” the hierarchy, it could be that Aristotle did not see the visual as a potentially semiotic or linguistic medium, and hence narrowed the causal channel to lead exclusively to spoken language. Whatever the explanation, the orthodox view of Aristotle’s definitions of spectacle and melody leaves out too much material. As scholars are wont to do, I will blame the vagaries of translation, figurative language, and mutations introduced by centuries of interpretation for this apparent lapse and proceed to advocate my own view.

The element of pattern thus refers to patterns in the sensory phenomena of the enactment. These patterns exert a formal influence on the enactment, just as semiotic usage formally influences patterns. A key point that Aristotle made is that patterns are pleasurable to perceive in and of themselves, whether or not they are further formulated into semiotic devices or language; he spoke of them, not only as the material for language, but also as “pleasurable accessories.” Hence the use of pattern as a source of pleasure is a characteristic of dramatic representations, and one which can comfortably be extended to the realm of human-computer experience.

**Language**

The element of language (usually translated as diction) in drama is defined by Aristotle as “the expression of their [the characters] thought in words” [Poetics, 1450b, 12–15]. Hence the use of spoken language as a system of signs is distinguished from other theatrical signs like the use of gesture, color, scenic elements, or paralinguistic elements (patterns of inflection and other vocal qualities). In the orthodox view, diction refers only to words—their choice and arrangement. That definition presents some interesting problems in the world of human-computer activities, many of which involve no words at all (e.g., most skill-and-action computer games, as well as graphical adventure games and graphical simulations). Are there elements in such nonverbal works that can be defined as language?

When a play is performed for a deaf audience and signing is used, few would argue that those visual signs function as language. The element of language in this case is expressed in a way that takes into account the sensory modalities available to the audience. A designer may choose, for whatever reason, to build a human-computer system that neither senses nor responds to words, and which uses no words in the representation. Hardware configurations without keyboards, speech recognition, or text display capabilities may be unable to work with words.

In human-computer activities, graphical signs and symbols, nonverbal sounds, or animation sequences may be used in the place of words as the means for explicit communication between computers and people. Such nonverbal signs may be said to function as language when they are the principal medium for the expression of thought. Accordingly, the selection and arrangement of those signs may be evaluated in terms of the same criteria as Aristotle specified for diction—for example, the effective expression of thought and appropriateness to character.

**Thought**

The element of thought in drama may be defined as the processes leading to a character’s choices and actions—for example, to emotion, cognition, reason, and intention. Understood in this way, the element of thought “resides” within characters, although it can be described and analyzed in aggregate form (the element of thought in a given play may be described as concerned with certain specific ethical questions, for example). Although it may be explicitly expressed in the form of dialogue, thought is inferred by both the audience and the other characters (agents), from a character’s choices and actions. In his application of a theatrical analogy to the domain of artificial intelligence, Julian Hilton puts it this way: “What the audience does is supply the inferencing engine which drives the plot, obeying Shakespeare’s injunction to eke out the imperfections of the play (its incompleteness) with its mind.” [Hilton, 1991]

If we extend this definition of thought to include human-computer activities, it leads to a familiar conundrum: Can computers think? There is an easy answer. Computer-based agents, like dramatic characters, do not have to think (in fact, there are many ways in which they cannot); they simply have to provide a representation from which thought may be inferred.

When a folder on my Macintosh desktop opens to divulge its contents in response to my double-click, the representation succeeds in getting me to infer that that’s exactly what happened—that is, the “system” understood my input, inferred my purpose, and did what I wanted. Was the system (or the folder) “thinking” about things this way? The answer, I think, is that it doesn’t matter. The real issue
is that the representation succeeded in getting me to make the right inferences about its “thoughts.” It also succeeded in representing to me that it made the right inferences about mine!

Thought is the formal cause of language; it shapes what an agent communicates through the selection and arrangement of signs, and thus also has a formal influence on pattern and enactment. The traditional explanation of how language serves as material for thought is based on the overly limiting assumption that agents employ language, or the language-like manipulation of symbols, in the process of thinking. This assumption leads to the idea that characters in a play use the language of the play quite literally as the material for their thoughts.

I favor a somewhat broader interpretation of material causality: The thought of a play can appropriately deal only with what is already manifest at the levels of enactment, pattern, and language. Most of us have seen plays in which characters get ideas “out of the blue”—suddenly remembering the location of a long-lost will, for instance, or using a fact to solve a mystery that has been withheld from the audience thus far. The above theory would suggest that the interjection of such thoughts is unsatisfying (and mars the play) because they are not drawn from the proper material. Plays, like human-computer activities, are closed universes in the sense that they delimit the set of potential actions. As we will see in the discussion of action below, it is key to the success of a dramatic representation that all of the materials that are formulated into action are drawn from the circumscribed potential of the particular dramatic world. Whenever this principle is violated, the organic unity of the work is diminished, and the scheme of probability that holds the work together is disrupted.

This principle can be demonstrated to apply to the realm of human-computer activity as well. One example is the case in which the computer (a computer-based agent) introduces new materials at the level of thought—“out of the blue.” Suppose a new word processor is programmed to be constantly checking for spelling errors and to automatically correct them as soon as they are identified. If the potential for this behavior is not represented to you in some way, it will be completely disruptive when it occurs, and it will probably cause you to make seriously erroneous inferences, to perhaps think “something is wrong with my fingers, my keyboard, or my computer.” The computer “knows” why it did what it did (“thought” exists) but you do not; correct inferences cannot be made. A text message, for instance, or an animation of a dictionary with its pages turning (language), could represent the action as it is occurring.

Other kinds of failures in human-computer activity can also be seen as failures on the level of thought. One of my favorite examples is a parser used in several text adventure games. This particular parser did not “know” all of the words that were used in the text representation of the story. So a person might read the sentence, “Hargax slashed the dragon with his broadsword.” The person might then type, “take the broadsword,” and the “game” might respond, “I DON’T KNOW THE WORD ‘BROADSWORD’.” The inference that one would make is that the game “agent” is severely brain-damaged, since the agent that produces language and the agent that comprehends it are assumed to be one in the same. This is the converse of the problem described in the last paragraph; rather than “knowing” more than it represented, the agent represented more than it “knew.” Both kinds of errors are attributable to a glitch in the formal-material relationship between language and thought.

Character and Agency

Aristotle maintained that the object of (i.e., what is being imitated by) a drama is action, not persons: “We maintain that Tragedy is primarily an imitation of action, and that it is mainly for the sake of the action that it imitates the personal agents” (Poetics, 1450b, 1–5). In drama, character may be defined as bundles of traits, predispositions, and choices that, when taken together, form coherent entities. Those entities are the agents of the action represented in the plot. This definition emphasizes the primacy of action.

In order to apply the same definition to human-computer activities, we must demonstrate first that agents are in fact part of such representations, and second, that there are functional and structural similarities between such agents and dramatic characters.

In a purely Aristotelian sense, an agent is one who takes action. Interestingly, Aristotle admits of the possibility of a play without characters, but a play without action cannot exist [Poetics, 1450a, 22–25]. This suggests that agency as part of a representation need not be strictly embodied in “characters” as we normally think of them—that is, as representations of humans. Using the broadest definition, all
computer programs that perform actions that are perceived by people can be said to exhibit agency in some form. The real argument is whether that agency is a "free-floating" aspect of what is going on, or whether it is captured in "entities"—coalesced notions of the sources of agency.

The answer, I believe, is that even when representations do not explicitly include such entities, their existence is implied. At the grossest level, people simply attribute agency to the computer itself ("I did this, and then the computer did that"). They also attribute agency to application programs ("My word processor trashed my file"). They often distinguish between the agency of system software and applications ("Multifinder crashed Excel"). They attribute agency to smaller program elements and/or their representations ("The spelling checker in my word processor found an error").

In social and legal terms, an agent is one who is empowered to act on behalf of another. This definition has been used as part of the definition of agents in the mimetic world. It implies that, beyond simply performing actions, computer-based agents perform a special kind of actions—namely, actions undertaken on behalf of people. It therefore also implies that some sort of implicit or explicit communication must occur between person and system in order for the person's needs and goals to be inferred. I think that this definition is both too narrow and too altruistic. There may be contexts in which it is useful to create a computer-based agent whose "goals" are orthogonal or even inimical to those of human agents—for instance, in simulations of combat or other situations that involve conflicting forces. Agents may also work in an utterly self-directed manner, offering the results of their work up to people after the fact.

For now, we will use the broader definition of agents to apply to human-computer activity: entities that can initiate and perform actions. Like dramatic characters, they consist of bundles of traits or predispositions to act in certain ways.

Traits circumscribe the actions (or kinds of actions) that an agent has the capability to perform, thereby defining the agent's potential. There are two kinds of traits: traits that determine how an agent can act (internal traits) and traits that represent those internal predispositions (external traits). People must be given cues by the external representation of an agent that allow them to infer its internal traits. Why? Because traits function as a kind of cognitive shorthand that allows people to predict and comprehend agents' actions [see Laurel, 1990]. Inferred internal traits are a component of both dramatic probability (an element of plot) and "ease of use" (especially in terms of the minimization of human errors) in human-computer systems. Part of the art of creating both dramatic characters and computer-based agents is the art of selecting and representing external traits that accurately reflect the agent's potential for action.

Aristotle outlined four criteria for dramatic characters that can also be applied to computer based agents [Poetics, 1454a, 15–40]. The first criterion is that characters be "good" (sometimes translated as "virtuous"). Using the Aristotelian definition of "virtue," good characters are those who successfully fulfill their function—that is, those who successfully formulate thought into action. Good characters do (action) what they intend to do (thought). They also do what their creator intends them to do in the context of the whole action. The second criterion is that characters be "appropriate" to the actions they perform; that is, that there is a good match between a character's traits and what they do. The third criterion is the idea that characters be "like" reality in the sense that there are causal connections between their thoughts, traits, and actions. This criterion is closely related to dramatic probability. The fourth criterion is that characters be "consistent" throughout the whole action; that is, that a character's traits should not change arbitrarily. The mapping of these criteria to computer-based agents is quite straightforward.

Finally, we need to summarize the formal and material relationships between character and the elements above and below it in the hierarchy. Formal causality suggests that it is action, and action alone, that shapes character; that is, a character's traits are dictated by the exigencies of the plot. To include traits in the representation that are not manifest in action violates this principle. Material causality suggests that the stuff of which a character is made must be present on the level of thought and, by implication, language and enactment as well. A good example is the interface agent, Phil, who appears in an Apple promotional video entitled "The Knowledge Navigator" (© 1988 by Apple Computer, Inc.). In the original version, Phil was portrayed by an actor in video format. He appeared to be human, alive, and responsive at all times. But because he behaved and spoke quite simply and performed relatively simple tasks, many viewers of the video complained that he was a stupid character. His physical traits (high-resolution, real-time human portrayal) did not match his language capabilities, his thoughts, or his actions (simple