Alternative representation for CHAT systems: Uniting the subject-object division

Hendranus Vermeulen
hendranus@gmail.com
http://people.cs.uct.ac.za/~hvermeul

1. Introduction

Cultural-historical Activity Theory (CHAT) offers us a unique lens to look at human activity in context. As such it is not a theory which aims to evaluate the “objective” truth of the social world, aiming to empirically prove or falsify testable hypotheses. CHAT is closer to ethnomethodology’s notion of theory. It is an attempt to make concrete sociology’s theoretical constructs with reference to the social world (Halverson, 2002). These constructs are then evaluated not on how “true” they are but how well they classify and provide insights to the socio-cultural activities they are applied to. These evaluations are in themselves cultural activities which can take the form of academic rhetoric, publications, etc. It is hoped that this short exposition provides such an opportunity for this alternative representation of CHAT systems to be evaluated by the CHAT community. Before getting to the alternative representation for CHAT systems I would like to briefly describe the context in which it came about. This context will illustrate how it was developed as a response to practical challenges. This will also make the activities represented more meaningful. If you are impatient or pressed for time, you can click: Go to straight to Ø representation.

2. Context

I have been applying CHAT as an analytical lens in the construction of collaborative 3D environments. These learning environments are modeled on the ethnomethodologies of ill-structured problem solving practices. Ill-structured problems challenge us to move away from the restrictive definition of cognition as computation and give us practice to drive theory with.

CHAT is an excellent framework for describing human activity in context and demonstrates situated and distributed cognition in action. CHAT and the activity triangle representation developed by Engeström (Engeström, 1987) are useful for describing learning activities; and informing the design of learning environments, which aim to simulate these activity systems (Jonassen & Rohrer-Murphy, 1999). The environments I have been developing aim to afford learning through facilitating design practices in these rich and interactive environments; i.e. epistemic games (Shaffer, 2005) and (Shaffer, Squire, Halverson, & Gee, 2004). I have subsequently developed an evaluation method for collaborative 3D environments based on the CHAT lens. I am currently writing up this thesis for my PhD at the Department of Computer Science, University of Cape Town. CHAT is ideal for describing learning as a process rather than using “objective” measures to gauge learning. The “real world” CHAT analysis provides not only the task analysis for informing design, but also a benchmark with which to compare the CHAT analysis, of the learning activity in the 3D environment.
2.1. Analyzing activities in film studies

The ill-structured problem considered was the Long Take (LT) assignment. The LT in film terms refer to a particular type of take (recorded shot). It is most generally identified by its extended duration. A LT is where you could have cut together a whole lot of short shots, but you decide not to for a functional reason. LT has become a convention within film and within this context serve specific functions in film production, i.e. eliciting a particular experience from the audience. The LT assignment forms part of the curriculum for third year film and media students enrolled in the film production program at the University of Cape Town. I video recorded the student’s during the production activity of the LT; and developed the following CHAT system triangle in an attempt to represent the complex design activity observed (Figure 1):

The production of a successful LT is the outcome of the LT activity. The success of the outcome depends on the teacher and audience. The outcome of the LT activity is therefore dependent on value judgments of others. This object produced by the LT activity is the LT. The LT as an object goes through various transformations: story, plan, rehearsal, take, voice over and edited LT. Although there was a slow linear progression from story to the final edited LT, the interactions between these two “stages” were highly iterative, with each of the “stages” feeding back on one another requiring rework of earlier “stages”.

![Figure 1](image-url)
The object of the LT begins as the story discussed by the students, represented and communicated using words, gestures and illustrations. The story is the core concept of the production, it is the message that the product (LT) is intended to deliver to the audience. The story then becomes an instrument along with the location, resources and the properties of these to develop a plan embodying the story in an executable form. It is important to note here that the transformation of the object is not only from story to plan. The plan also becomes an instrument for subjects to reconsider the story. The transformation of the object is highly iterative moving both directions from story to plan and plan to story. The plan becomes the instrument for the rehearsal. The rehearsal is the next “stages” of the LT object’s transformation. The plan is executed during the rehearsal. Problems and oversights are identified and this then feeds back into the plan or story which is again transformed. For brevity the description of interrelations of other transformations of the LT object will be omitted.

3. Challenges identified

3.1. Challenge 1: The object’s transformation is not represented

The first challenge encountered using the CHAT triangle to represent the LT activity is that it did not represent the transformation of the object. The LT is in the object position, it is static. In Figure 1 I have attempted to illustrate the various “stages” of the LT object by indicating them next to the object. However, it does not represent the transformation of the object of the activity, particularly where the transformed object changes position in the activity system to bring about transformation. The best solution I had come across has been using more than one activity system to represent antecedent activities (University of Helsinki, 2006), i.e. representing each “stage” as a sub-activity. If you are using Acrobat Reader 6 or later you can click on Figure 1 to make the Shockwave file active. Then you will be able to navigate to sub-activities by clicking on the various stages of the LT object’s transformation. This solution however introduces additional challenges.

3.2. Challenge 2: Static representation does not illustrate temporality

It is difficult representing temporal relations using the triangle model, that is meaningful and reflect time span or duration of activities (Mwanza, 2002). It is essential to keep in mind that the static representations imply a dynamic iterative process. Animations have been added to represent the transformation (movement) of the object in the activity’s process. However, even the interactive triangle still represents rather static snapshots of the sub-activities. Relationships between sub-activities are also difficult to represent.

Hoping to find a better way of represent the dynamic transformation of the LT object, I decided to narrow the focus of my analysis to the second of the sub-activities of the LT activity system, namely the plan. This would provide a more tightly defined activity system for analysis. The reduced scope would hopefully reduce the activity’s complexity, making it easier to represent; and facilitate the implementation of the resulting 3D environment.
3.2.1 Reducing scope: Focusing on the planning stage

The LT assignment emphasized the planning of the choreography between the camera person, actor, props and cinematic space. Through this skillful choreography the character and story would be revealed. The planning of this choreography lies at the heart of the LT activity and provides a more contained activity system for analysis. The activity system producing the plan is illustrated in Figure 2 below:

The planning activity system centers on the production of the plan as object. The students participating in the LT assignment develop the plan for embodying the story in the selected location using their bodies, props, the physical space and a floor plan. The production of the plan is divided into three roles: director, camera person and actor. The students draw on the codes and conventions of the film community and are constrained by the rules of the long take exercise.

The three roles played by the students (director, camera person and actor) are particularly significant for the production of the plan. The students literally think through the plan using their bodies as objects for manipulation. The actor is acting the character of the story improvising their actions. The camera person is framing the actor and action; working out the movement of the camera in the location’s space. The director observes this interaction and provides feedback and instructions to the actor and camera person regarding performance, cinematography, blocking, props, etc. The interaction between these three roles is best described as choreography. Like a dance the participants perform individually, yet their continual interaction and feedback develops the activity as one movement.
3.3. Challenge 3: Interactions between participants are not represented

This dynamic interaction determining the plan, lies at the heart of the LT activity system; and was not adequately represented in the plan activity system. It was therefore decided to attempt to represent the planning activity represented in (Figure 2) using three activity systems. Each of the activity systems representing one of the activities performed by students in their respective roles (Division of labor): camera person, actor and director. Figure 3 attempts to represent an integrated representation of these three activity systems:

![Figure 3](image)

*These are vector images, so you can use the Acrobat Reader’s zoom tools to get a closer look.*

3.4. Challenge 4: Difficulties in representing the shared object

Despite the practical challenge of arranging and fitting all the elements on this page we are also faced with the challenge of representing the shared object, i.e. plan. Each of the activity systems associated with a role has a unique object that the subject of that activity system is transforming. The actor is using their body to produce an act, the object of the actor’s activity. The shot is the object the camera person is producing using the camera as instrument to frame the act. The director produces their vision using the actor and camera person as instruments (Red circle Figure 3). Although it could be interpreted that directors use people as instruments, this significantly reduces the contribution the camera person and actor was seen making in the planning activity. All three parties interact dynamically; reacting and building on one another’s ideas and suggestions. It is truly a collaborative activity where all participants were contributing equally to the production of the plan.

In an attempt to represent an alternative configuration (Figure 4 bellow) was developed, arranging the three activity systems around a shared object. The shared object (highlighted in red) is presented in the centre.
3.5. Challenge 5: Difficulties in representing interaction between activities

Using multiple activity systems to represent the sub-activities of an activity system (as in figure 3 and 4), introduces the problem of indicating the relationships between these sub-activities; and also the interaction between these sub-activities objects and mediators. For example in the LT we find that space plays a very important role in the co-ordination of the activity. The contradictions between the camera person’s framed space, versus the actor’s performance space, versus the director’s cinematic space (location, props, etc.); constrains and drives the plan to be developed. These between activity system contradictions are also difficult to represent, however, very important in driving the larger activity system.

Mwanza (Mwanza, 2002) developed the Activity-Oriented Design Method (AODM) with the aim of providing practitioners with a structured and replicable method for applying CHAT. Although the AODM techniques and resulting visual representation enable us to break apart the components of the LT activity system; and draws our attention to the contradictions between these components they, do not illustrate how these contradictions support the larger activity.
4. Considering the alternatives

Wells identifies similar challenges in his exposition of the co-construction of artifacts and meanings: “By now it will be abundantly clear that it is no easy matter to represent diagrammatically the multidimensional complexity of this—or any other—activity. In the first place, the temporal sequence of the different types of constitutive action requires more than one diagram, if only to show how artifacts created as the outcome of one action become mediators in subsequent phases... Even more difficult is to show the shifting relationships among the various mediational means that the girls employed to achieve the goals of the different phases of the activity.” (Wells, 2002).

**Figure 5**

Wells developed an interesting solution (Figure 5): “For the sake of simplicity of representation, this version of the diagram represents a dyad engaged in joint activity. However, there is no limit, in principle, to the number of participants who may be jointly involved. Each participant is represented by an expanded triangle of the kind created by Engeström (Engeström, 1987). All the participant triangles (just the two in this version) share a common base, indicating that the co-participants are members of the same community and subject to the same rules and division of labor. Most importantly, they are acting on the same object, which is represented in the center of the diagram—though, of course, this does not mean that they construe it in exactly the same way. Each also
brings to the activity his or her personal kit of resources, drawn from the pool available within the community as a whole.” (Wells, 2002).

In principle there is no limit to the number of participants involved in Well’s representation, however, practically we are limited to the dyad. It would be difficult to clearly represent three participants, since one of the participants will cover the others’ activity systems. This is the case in 2D. If we construct the activity system in 3D, it would provide an additional dimension to fit participants into. With this in mind I developed the following 3D collaborative activity system (Figure 6) using a 3D modeling application:

![Image of a 3D model](Figure 6)

*If you are Using Acrobat Reader 7 or later you can activate the 3D model by clicking on Figure 6. Once activated you can click-and-drag on the model to rotate (default), walk, pan or zoom.*

When modeled in 3D the activity system is most economically represented as an octahedron. The octahedron when viewed from the “Front” represents a hexagon, which is really a CHAT triangle with the Subject, Object and Community points expanded through the sides of the triangle. Although the 3D representation is interesting and provides us with a simpler geometry; it does not simplify our interaction with the representation and enable us to more clearly represent collaborative activities. On the contrary it seems to complicate matters.
5. The fundamentals of the problem

Having run out of alternatives, I started thinking that maybe these challenges derive from a more fundamental problem underlying the CHAT triangle. Could it be that the problems I have been encountering derive from attempting to represent an inherently multidimensional phenomenon in Cartesian space (n-Space)? We all learn to visualize 'space' as the geometrical space of three dimensions; having the x, y and z axes of the Cartesian co-ordinate system. Could it be that the old dualistic division between x and y, mind and body, knower and known, object and subject, lie at the heart of my troubles?

CHAT opposes Cartesian dualism: “A difficulty with the classical AI view from the perspective of a cultural-historical approach to mind is that it adheres to a strong Cartesian separation between what is outside and what is inside the mind” (Cole, 1997). CHAT (Cole & Engeström, 1993), distributed cognition (Hutchins, 1995), situated learning (Lave & Wrenger, 1991), mediated action (Wertsch, 1998); all adhere to the notion that the mind is intimately related to the body; and the rest of the “world” for that matter. People are situated in cultural, historical and ecological contexts. We think from being-inside this situated existence and our thoughts derive from “being” inside this relationship. Cognition can therefore not be separated from the “world” and limited to the cranium; our thoughts are related to our instruments, culture, and ecology. Cognition derives from the relationship between us and the “world”. But what is the nature of this relationship? According to the abovementioned theories cognition develops from human action in situ and is mediated by cultural tools. This process can be represented using the CHAT triangle (Engeström, 1987). The subject acts to produce the (cognitive) object using instruments. The process is supported by the community and their historically developed practices, division of labor and the rules of the activity. Could it be that the challenges we face in representing collaborative activities result from the CHAT triangle (Engeström, 1987) positing the object and subject in opposing co-ordinate positions in Cartesian space?

It has been argued that CHAT and other theories of distributed cognition do not place enough emphasis on the role of the object in mediating activity. Should the object be granted the same ontological status as the subject as argued by (Latour, 1996a) and (Shaffer & Clinton, in press)? It would seem as though objects are more and more becoming active participants in human activities. In the 3D learning environments that I am designing, all user interactions are “controlled” by this object. The representation of the user in the “world”, what they can do, their goals and progress are all “determined” by the interactions that I have programmed into the simulated environment inside the computer. However, I don’t want to pursue this argument here. If you are interested please see the discussion (Hutchins, 1996), (Latour, 1996a), (Engeström, 1996) and (Latour, 1996b). If the resulting representation lends itself to be a useful toolforthought (Shaffer & Clinton, in press) for using to think and theorize about the ontological status of objects, that would be great. However, what I am concerned with here is solving the practical problem of representing collaborative activities. These ideas have however influenced my thinking and considerations regarding the object-subject relationship as opposing forces united in a process transformation.
6. Creative problem solving

So how otherwise could one represent the relationship between subject and object? Thinking in binaries has become so entrenched in our western thought that it has become quite difficult to think otherwise. I decided to look East in search for alternative representations that might capture the dynamic transformative relationship between object and subject.

![Yin Yang Symbol](image)

The “Yin Yang” symbol originates in ancient Chinese philosophy and metaphysics. It represents the unity of opposites, describing the nature of all things in the universe as two primal opposing but complementary forces. The dark side - “Yin” means shady; and represent the night and dormancy, while the white side – “Yang” means sunny; and represent the day and activity. Like day and night the one cannot exist without the other. They are represented as opposites, yet united in occupying the same space. They exist in a process of continual transformation; both Yin and Yang consuming yet supporting the other. This process is ideally held in balance; as one increase the other decreases. The “Yin Yang” therefore also represents balance. In each of the opposing forces we also find a small dot representing a bit of the opposite force. “Yin” can consequently transform into “Yang”; and “Yang” into “Yin”. Like the day turning into night and then into day, etc. The “Yin Yang” is therefore a great symbol for representing transformation, process and change. The “Yin Yang” has been popularized in the West as a Taoist symbol and now forms part of our popular culture. Could I use the “Yin Yang” and its paradoxical connotations of unity in opposites to represent the relationship between object and subject?

6.1. The Ø Representation

If we unify the Symbol S (representing the subject) and O (represent the object) in Cartesian space, by placing them on top of one another, we get Ø which looks strikingly like an inverted “Yin Yang” symbol. I created the “Ø” character to denote the paradoxical subject-object coupling; and signify the new representation I was developing. I set out to develop the Ø representation using the “Yin Yang” symbol and its symbolism to illustrate the dynamic relationship between object and subject in a united process of transformation. I decided to start by representing unmediated action (Figure 7) bellow:
6.2. Unmediated Activity

Figure 7 represent the most basic relationship between the object and subject using the Θ representation, i.e. unmediated interaction. Unmediated interaction occurs in activities which are not mediated by cultural instruments. The subject acts directly on the object, for example an organism (subject) foraging for food (acting) in the environment (object). According to the ecological approach to perception (Gibson, 1986) the organism’s perception is an active process that selects the relevant information (smell of food) from the environment necessary for its survival. The environment (object) “makes” these opportunities available to the organisms inhabiting it. The environment properties (protein richness) is said to afford the organism, which have evolved to find and exploit, these opportunities. This blurs the distinction between environment and organism. In Figure 7 the object and subject are placed in opposing positions, yet united in a process of mutual constitution (indicated by the red arrows). I have also swapped the object-subject’s usual positions of subject left and object right, to emphasize the movement (circulation) of the dynamic transformative relationship between object and subject. The Θ representation provides a much more dynamic representation of an activity system (challenge 2). Having made some progress, I decided to complicate matters by representing mediated activity using the Θ representation (Figure 8), as illustrated bellow:
6.3. Mediated activity

In the above representation (Figure 8) the Θ is used to represent mediated activity. The activity revolves around the subject transforming the object. The Θ representation highlights the Θ interaction; it is the focus of the lens. In Figure 7 the process of mutual constitution is indicated by the red arrows. I have also labeled these arrows to try and describe the mediation-action coupling at work in the transformative interaction between the subject and object. I am also using the term “externalize” (Piaget, 1962) and “internalize” (Vygotsky, 1978). Although these terms were defined in context of play, I hope to use them to illustrate the synergy at work between the subject and object. Externalization and internalization are but two poles of the same process (cognition), linking the “outside” and the “inside”. Coercion and volition is used to describe the autonomy paradox inherent in all Θ’s.

In context of mediated action the Θ transformation can also be interpreted to represent *interobjectivity* as defined by (Latour, 1996a). Cultural objects and the infrastructures they constitute enfold the intentions of the subjects that design them. However, these objects and infrastructures also constrain subjects’ interactions. These cultural objects both “outside and inside” our heads play a important role in our cognition and activities (Latour, 1996a). This *interobjectivity* can be inferred from the Θ transformation; illuminating the active role objects play in influencing our thinking and behavior.

The mediators of the activity are arranged around the Θ transformation. I am using the traditional categories of the CHAT system triangle to label the mediators. They are arranged around the periphery of the Θ indicating the mediatory and constraining
(contradiction) roles mediators play in driving the Θ transformation; and so the activity. I have positioned the instruments next to the subject, indicating that this is what the subject uses to transform the object. The outcome is the goal mediating the production of the object and therefore positioned next to it. The other mediators I have arranged to be positioned next to the “body” of the subject or object they relate to. One of the advantages of using this arrangement is that there is more space to position mediators, but more importantly, mediators that do not fall strictly in one category or change category during the activity can be accommodated. Each of the mediators is also in themselves a minimized Θ. In an attempt to make this labeling clearer, I applied the Θ representation to the director (subject) producing the vision (object) activity, in Figure 9 bellow:

6.4. Director’s activity system

![Diagram of Director’s Activity System]

The director plays a very important role in the production of a film. They are responsible for the artistic vision of the film; and guiding the whole production team in realizing this vision. The director externalizes their vision by directing and giving instructions to the other members of the team. The outcome of the activity is a successful choreography that captures the vision. As the choreography unfolds in the: cinematic space, story, the actor’s performance, the camera person’s cinematography, blocking, consideration of the audiences’ perspective, etc. these mediators give shape to the director’s vision. This provides the director with a more concrete object to reflect on and internalize; and utilize to make decisions and give instructions about the unfolding choreography. The director’s vision is realized through this dynamic and highly iterative process. When considering collaborative activities, however, we need to expand the mediatory loop to make space for others to join the activity; as illustrated in Figure 10 bellow:
6.5. Expanded Mediation

Figure 10

Figure 10 utilizes the CHAT triangle’s categories to represent the expanded Θ including the mediators in a feedback loop (red arrows) of shared mediatory space or context as described by (Dourish, 2004). Figure 11 illustrates the expanded Θ representing the example of the camera person producing the shot:
6.6. Collaborative planning activity

If we extend the mediational periphery beyond that of the individual Θ it makes space for other Θ’s to enter the shared context of collaborative activity. Figure 12 (below) represents the collaborative planning activity (as an alternative to the problematic Figures 3 and 4):

In the collaborative activity represented in Figure 12, we find each of the Θ’s’ individual objects of transformation represented; as well as the shared object (challenge 4). The plan is the shared object being transformed in the shared mediatory space. It is the focus of the collaborative activity and unites the three Θ’s action-mediation couplings in a shared mediation cycle. We see the interaction flowing from the shared mediation space, driving each of the Θ’s and their object transformation (red arrows). The shared mediational cycle also resolves the challenge of representing interactions between the activities’ mediators and objects (challenge 5). The transformation of each of the Θ’s in turn loop back to drive the shared mediation cycle. Each of the Θ’s is communicating with the others resulting in their shared intersubjectivity as described in (Kaptelinin & Cole, 1996). The interactions with mediators can be communicated (orange arrows) to other Θ’s and so effect their Θ transformation (challenge 3). The Θ’s interaction with mediators (red arrows) and communication between Θ’s (orange arrows) demonstrates how the individual and collaborative activities interact to create each other. An example of this would be if the director regards the space mediating the camera person’s framing and the actor’s performance as constraining. The director communicates this to the other participants; and so influence the production of the shot and act (individual objects), as well as the plan (shared object).
6.7. Re-representing the LT activity

Having resolved the challenges encountered in representing the collaborative planning activity, can we use the \( \Phi \) representation to resolve the initial challenge? Figure 13 (below) aims to do just that:

![Figure 13](image)

In an attempt to re-represent the LT activity system (Figure 1) using the \( \Phi \) representation, I have arranged the various “stages” of the LT activity as mediators around the Film Students (subjects) and the LT (object). The mediators are in themselves an \( \Phi \) couplings. They are however interlinked, effecting and being effected by the larger \( \Phi \) transformation. The “stages” can therefore be thought of as part of the “body” of the larger LT object; mediating and driving the transformation of the LT object. The transformation of the object is so represented (challenge 1). In “opposition” to the mediators transforming the LT, I have tried to arrange the (learning) outcomes we hope to be mediated along side the “body” of the subjects (film students). This aims to illustrate that for each of the mediators transforming the LT a corresponding skill could be learned. These outcomes are the practices required to perform a LT. These are the skills we hope the students will internalize by participating in the LT activity. Through the planning activity, for example, we hope the film students will learn choreography. Through the writing of the script they will learn narration, etc.
7. Conclusion

It has been the aim of this exposition to account for the development of the Θ representation in response to practical challenges faced trying to represent complex collaborative activities using the CHAT activity triangle (Engeström, 1987). In so doing I have not systematically thought through all the theoretical implications and connotations of the Θ representation. It is my sincere hope that this exposition could serve as a shared object to elicit such discussions and further refinements of our representations, communication and thinking about activities; and in particular collaborative activities.

I also hope that these ideas could contribute to the resolution of some of the challenges facing third generation CHAT by establishing the role of communication, multiple perspectives, power and diversity (Cole, 1988). The true test of the Θ representation lies in whether other members of the CHAT community find it useful in representing complex and collaborative activities. I have applied the Θ representation to the 5th Dimension activity system (engineered by Michael Cole and his colleagues) in the hope that it might demonstrate its usefulness (Figure 14 bellow):

![Figure 14](image-url)

In this attempt to represent the 5th Dimension using the Θ representation (Figure 14), I have expanded 4 of the Θ’s participating in the collaborative 5th Dimension activity system. The mediators are also arranged in opposing, yet superseding binaries (Brown & Cole); illustrating how these “opposites” are united in, and driving the transformative activity system of the 5th Dimension. Each of the participants has a different perspective in the collective activity. These perspectives are communicated between participants (orange arrows), resulting in intersubjectivities (blue) that regulate and co-ordinate the individual activities’ interobjectivities in relation to the collective activity (red arrows).
8. References


Dourish, P. (2004). What we talk about when we talk about context.


