Using Situated-Action Networks to Visualize Complex Learning

Alejandro Andrade, Indiana University, laandrad@indiana.edu

Abstract: The purpose of this paper is to show how Situated-Action Networks (SAN) is a viable visualization tool for representing the learning that happens as a movement from peripheral to more engaged forms of participation in a social activity. In taking action as the unit of analysis, persons, speech acts, and artifacts are considered as the nodes in a functional model of learning. This method is illustrated by a simple proxy where a graduate student takes the place of a pilot apprentice who learns how to land an airplane in a computer simulation. Findings show that the SAN's sociographs and the learning trajectory plot display the interconnected complexity of the learning that occurred at various forms of action as a result of participation. Further work and limitations of this method are discussed.

Keywords: activity theory, distributed cognition, situated learning, social network analysis

Introduction

The purpose of this paper is to explore a visualization tool called *Situated-Action Networks*. Inspired by sociocultural approaches to learning (Brown, Collins, & Duguid, 1989; Cole, 1996; Lave & Wenger, 1991) and through the analytical lenses of activity theory (Engeström, 1987; Roth, 2007; Vygotsky, 1978) and distributed cognition (Hutchins, 1995; Hutchins & Klausen, 1996), Situated-Action Networks can offer new insights as to how the learner's traces can be analyzed to generate complex visualizations of the learning process. Similar to Epistemic Network Analysis (ENA, Shaffer et al., 2009), Situated-Action Networks focus on the observable interactions in various forms of social activities. In shifting the unit of analysis from cognitive processes to social activity, this form of visualization accounts for the social dynamics, tracing the learner's trajectory of participation from the periphery to a more engaged form of participation (Lave & Wenger, 1991). Building on Social Network Analysis (Wasserman, 1994), various descriptive network indices can provide useful information about the learner's level of participation in the social activity. I illustrate such a model with the use of a simple proxy of participatory learning where a graduate student takes the place of a pilot apprentice who learns how to land an airplane in a computer simulation.

Learning as a process

We take *action* as the unit of analysis. According to activity theory (Engeström, 1987; Leont'ev, 1974; Roth, 2007; Vygotsky, 1978), an action stands as the mediational link between a subject and an object. In this dialectical sense, actions are processes that connect two qualitatively different expressions of the same unit: in actions, subject and object are two constitutive elements that cannot be considered independently from each other (Roth, 2007). The idea that both subjects and objects are opposite but mutually constitutive parts of the same unit has important consequences for what, and how, subjects and objects are included in a representation of the learning process. In our analysis, subjects and objects are both represented as nodes in a network. This idea is not new, however, and some antecedents can be found in Latour's Science and Technology studies (Latour, 1987) and some educational approaches inspired by his ideas (see for instance Shaffer & Clinton, 2006). Also, the theory of distributed cognition takes the system of cognition of both persons and objects as the unit of analysis (Hutchins, 1995; Hutchins & Klausen, 1996).

Two main kinds of actions are envisioned in SAN analysis: a) instrumental and b) communicative actions. Instrumental actions refer to the use of artifacts – as a transformative material process (e.g., hammering a nail) or as the consumption of a source of information (e.g., reading a ruler or a scale). Communicative actions refer to linguistically achieved (i.e., verbal or non-verbal) intersubjective actions. In this case, language *is* the object. Vygotsky (1986) conceived of language as a tool that shared the same ontological status as that of a physical tool. For Cole (1996), this ontological status is clarified by using the term "artifact" to refer to both linguistic communication and material tools.

The conception of learning that we bring to bear here can be defined as a *micro-genetic* change in the pattern of actions. A pattern of actions represents the regularities of the learner's actions as well as the actions from all the other participants in the situation. That is, the connections between subjects and objects through actions in a regular manner. A micro-genetic change can be considered as a small developmental change that occurs as a consequence of an appropriation of the social practices (Cole, 1996; Roth, 2007; Vygotsky, 1978). In this sense, learning would not be necessarily a change in the apprentice's mind, but rather a change in her participation – because the situation later displays a rather different pattern of actions. Therefore, learning is not observed as the acquisition of any kind of knowledge, but of a *knowing how* to go about doing things in practice (Greeno, 1998). This way of understanding learning has important consequences for how we should represent it. At the level of the analytical model, a learning process is better represented by a functional, rather than a structural, model (Roth, 2007). A structural model represents the elements in a situation, such as participants and instruments; for instance, a student and some weights, a notebook, and a scale. A functional model represents the actions that take place in that situation; for instance, the weighting and the note-taking. Because the pattern of actions in a situation can develop and change after participation (i.e., after the learning occurs), the process of learning is captured by two or more functional models displayed at different points of time. The changes in participation that occur in small periods of time (e.g., one or two hours) are regarded as micro-genetic changes (Vygotsky, 1986).

Situated-action networks

Our model uses Social Network Analysis (Wasserman, 1994) in a broad sense. In our analysis, nodes represent participants, artifacts, and speech acts. Actions constitute the links between nodes; that is, one action is represented by a link between two nodes. In SAN, the number of nodes depends on the number of participants, artifacts, and speech acts in a situation. In addition, the links between nodes are quantified by the amount of similar actions belonging to the same category (i.e., links that relate the same person to the same speech act category or the same instrument). We assume that the situational and organizational constraints would allow us to quantify the occurrences (cf. Schegloff, 1993) of these types of actions that relate persons, speech acts, and artifacts in a learning activity. The method can be outlined in four steps. The first step is to create a raw data matrix where nodes and actions are recorded throughout segments of time (e.g., 1 minute or 10 second intervals). Second, a number of segments is subset depending on the starting and ending point of a socially meaningful activity (e.g., a 45 min class or a 10 min class activity). Third, an adjacency matrix is produced from the raw data matrix. And, fourth, a graphic representation is created.

For instance, take a simple scenario where two children are solving the Hanoi-Tower problem collaboratively. The two children, the materials they use, and the types of utterances they make, are all represented by nodes in this situated-action network. The actions are links between nodes, that is, one action links a subject to the object of that action. In the example, let's consider what types of nodes and actions are present. Let's suppose that child A tells her partner "Move it to the left", pointing at one of the wooden blocks; then child B moves the block to the left. In this particular case, there would be one node per child, one type of utterance, and one wooden block. This simple case depicts a sequence of two actions, one that links the utterance by child A and one that links the moving of the block by child B, or in network notation, g = ((Child A, Talk), (Child B, Move)). One can examine this interaction between child A and B more closely and include a link between the utterance and child B because this utterance was intended to her. In network notation, g = ((Child A, Talk), (Talk, Child B), (Child B, Move)) (see Figure 1).

$$
\begin{array}{r} \\ d1\!: Children\ A \\ g = d2\!: Children\ B \\ d3\!: Talk \\ d4\!: Move \end{array}
\begin{array}{c} \begin{array}{cccc} d1 & d2 & d3 & d4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}
$$

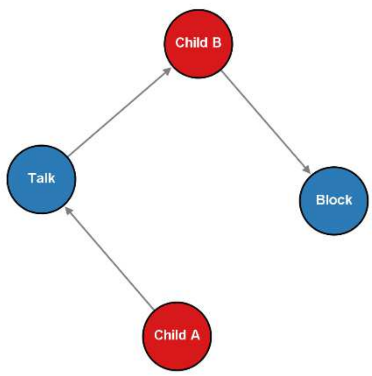


Figure 1. Matrix and graphical representation of a functional model

In comparing several of these representations over various activities, these networks reveal informative changes in the pattern of actions. The thickness of the lines can be used to represent the density of the actions between pairs of nodes. Various network indices (e.g., degree and centrality of nodes) can supplement the analysis, and can be used as a proxy of the learner's level of participation in the social activity. For instance, in our simple scenario, let's suppose that child B was learning how to solve the Hanoi-Tower problem, and child A was the more capable other (see Figure 2). In the first attempt, child A gave a lot of directives as to what actions were to be taken and child B moved the blocks as instructed. In a second attempt, both children were talking and moving the blocks equally. And in a third attempt, the roles were inverted in the sense that it was now child B who was able to tell child A what actions were required to solve the problem and child A moved the blocks.

In addition, the novice's learning trajectory toward more engaged, expert-like participation can be displayed by plotting each network's center of mass in a 2-dimensional plot using the first two principal components. A principal component analysis is a statistical method that uses the eigenvalues from the adjacency matrix to create a linear combination of the nodes so as to cluster the most important ones (i.e., the nodes that carry the most weight) in the first dimension and a second linear combination with the second most important nodes for the second dimension. Each dimension is named after the dominant node that accounts for the most weight in the function. A 2-dimensional plot uses the first two principal components to represent the linear combination of the two most important clusters of nodes with the greatest weights. Because these dimensions are metric free, the numbers on the axis are arbitrarily set to have a mean of 0 and a standard deviation of 1. Using the principal component scores from the adjacency matrix, a center of mass is calculated as the multivariate centroid for a person in a particular iteration of the activity. For instance, the progression of the center of mass from our previous example would show that the density of actions of child B progresses (throughout the various iterations of the activity) towards that of child A in the first attempt (see Figure 3). As can be seen, the first center of mass for child B in (a) is high on movements but low on talk, and the progression shows how the importance of movements goes down whereas the importance of talk goes up toward the third center of mass (b). To illustrate the application of Situated-Action Networks, I bring the case of a graduate student who learns how to land an airplane in a computer simulation with the support of a more knowledgeable other.

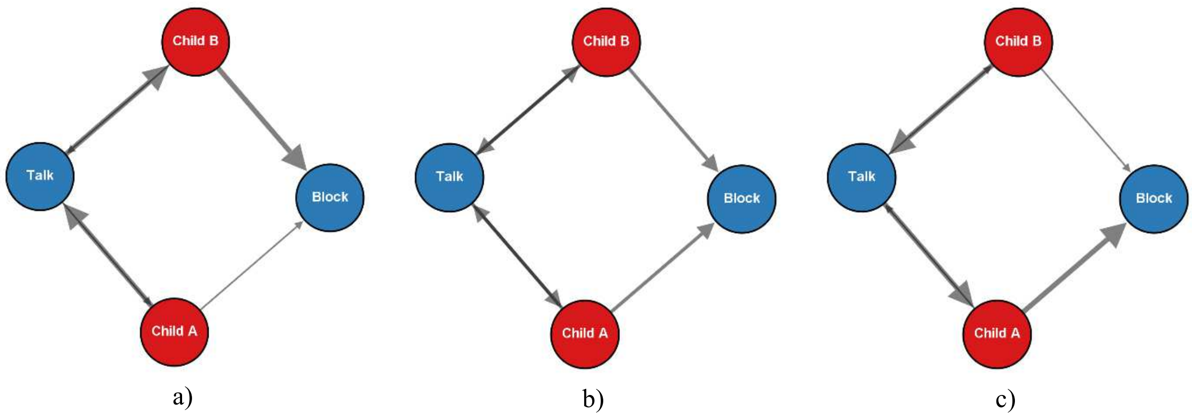


Figure 2. A sequence of situated-action networks in a simple scenario

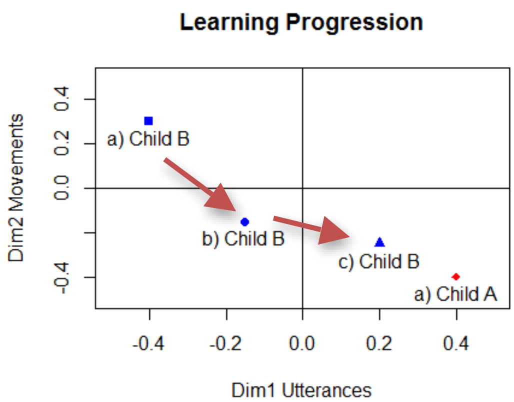


Figure 3. Child's B learning progression represented by the trajectory of his center of mass

A brief example: Learning how to land an airplane

Methods

This example is a proxy of a naturalistic complex activity: the participation in the navigation and landing of an airplane in Microsoft Flight Simulator X. The goal of the learning activity was to support the learner's recognition and use of both the *navigation* indicator (GPS), and the *attitude* indicator (Artificial Horizon) to aid his performance in landing the aircraft. The purpose of the analysis was to visualize the progression of the learner's participation while toggling between the roles of the pilot and the copilot in the cockpit through repeated iterations of the activity. Seven activities were designed and conducted in a conference classroom, with a projector, a laptop, and a joystick (see Figure 4). The first activity was a pretest, the fifth activity a posttest, and the last two were near and far transfer activities. The total implementation lasted 2 hours. The participant was a convenience sample of a volunteer graduate student friends with the author, where the latter acted as the expert pilot because of his expertise with the simulator. The research question was: How does participation in several iterations of the activity shape the learner's participation in the activity, that is, his landing of a plane? The analyses were conducted with two main sources of data, a) talk and b) instrument use.

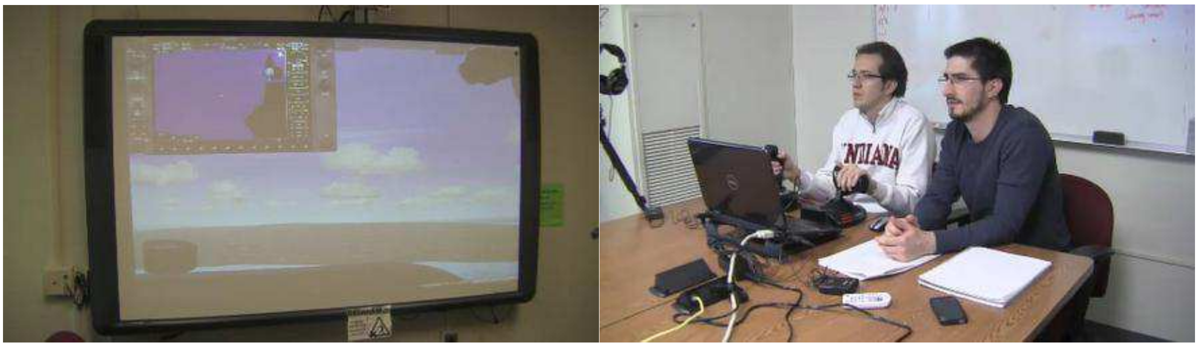


Figure 4. The research setting

Talk

Transcriptions were carried out to capture the textual talk at the expense of detailed nuances of pronunciation and timing. Our assumption was that the cockpit conversation would reveal the level of understanding of the piloting process (speed and power, altitude and speed, flaps and speed, etc.). Also, because these utterances have a dynamic and action-oriented nature (Schegloff & Sacks, 1973), grounded qualitative codes were produced to reveal the intended use of each speech act (see Table 1). Thus, an action would link the actor to a talk code – for instance, g = ((pilot, question), (copilot, answer), …, (pilot, reflection)).

Table 1: Grounded codes for the cockpit talk

Code	Definition	Example
Question	Asking a question	"What is the altitude I have to descend to?"
Answer	Answering a question asked	"Descend to 2000 feet"
Assent	Acknowledging a previous utterance	"Okay, descending to 2000 feet"
Caution	Cautionary statement	"Careful, you are a little too high"
Fact	Giving an objective fact	"Your speed is 70 knots"
Assistance	Providing direct assistance	"You are going too fast, I'll setting flaps to 20 degrees"
Prompt	Prompting a suggestion	"Make a right turn now"
Reflection	Pondering some facts and relations	"If I move the flaps down now then it will slow me down, right?"

Instrument use

Video data was analyzed to capture the density of the instrument use in five areas: a) pilot's gaze, b) joystick, c) throttle, d) flaps, and e) landing gear. The pilot's gaze was mostly directed at two distinct places: i) at the instrument panel or ii) at the screen where the simulation was projected. Our insight for coding the gazing behavior this way, perhaps due to the researcher's prior knowledge with the flight simulator, was that the coordination between the instruments and the screen is an important part of the piloting activity because it reflects the pilot's awareness of the plane movements. In addition, the use of the joystick and throttle were counted as an action if the pilot pushed or pulled the controls. As the copilot was responsible for activating the flaps and the landing gear, we tallied those moments as actions as well.

Results

Instrument use

How did the learner's instrument use change after participation and what did it look like in comparison to the expert's? Results from the gaze tracking (see Figure 5) revealed that during the pretest activity the learner looked twice as much at the instruments (294 sec, in blue) as at the screen (156 sec, in red). During the posttest activity, this proportion was approximately equal (200 sec at the instruments and 216 sec at the screen). During the expert-led activity, results revealed that the expert looked more at the screen (229 sec) than at the instruments (121 sec). Not only was this total proportion different between expert and apprentice, but also the average gaze length. The learner's average gaze length was 9.8 sec in the pretest activity, and decreased to 5.4 sec in the posttest activity. For the expert, the average gaze length was only 3.2 sec.

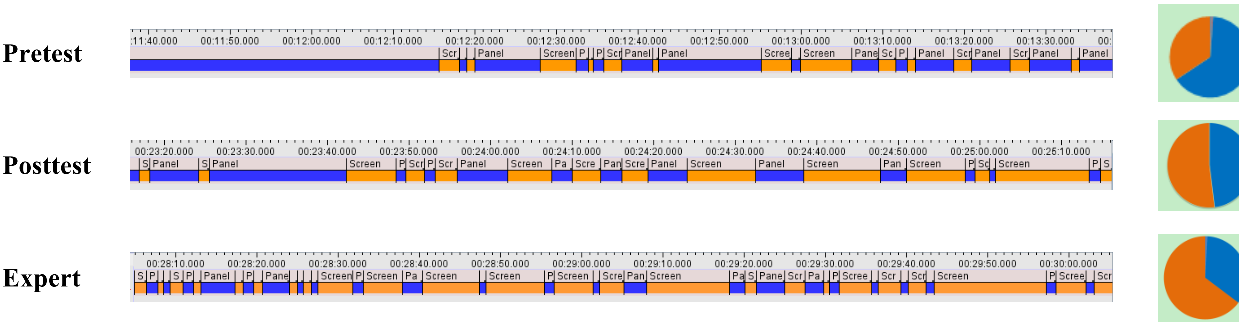


Figure 5. Gaze at instruments (blue) or at the screen (red)

These differences between pre and posttest and also between learner and expert are informative. During the pretest, the learner was trying hard to figure out how to read the *attitude* indicator, which is packed with information about the behavior of the plane. The apprentice had little time to observe how the "real" plane behaved (on-screen); instead, he seemed to be trying to respond to the information coming from the instruments and the directives from the copilot. Conversely, during the posttest activity, the learner seemed to better integrate these two sources of information, the instruments and the screen, into one blended space. His gaze moving back and forth more quickly from one source to the other supports this idea. On the other hand, it is apparent that the expert flies the plane looking primarily at the screen (the real behavior of the plane) and relies on the instrument information only to double check the correct performance of the plane (like dancing without looking at your feet). It seems that he does a rapid eye movement to glance at the instrument panel and then goes back to the screen, without losing sight of this latter for more than two or three seconds per time. The learner's posttest behavior seems half way between the pretest and the expert-like behavior. In the posttest, the apprentice did not rely on the instrument panel as much as in the pretest, but his gaze was not as quick as the expert's.

Similar changes occurred in the use of the controls. The expert seemed to have spent a lot of time moving the joystick and the throttle (71.8% and 63.5% of the total flight time, respectively). For the learner, we observed an increment in the use of these controls from the pretest to the posttest (from 34.8% to 71.2% for the joystick, and 18.4% to 46.9% for the throttle). We believe that these changes in the learner's behaviors are due to three factors: a) the development of a more fluid perceptual representation of the plane's behavior; b) the development of his motor coordination; and c) the development of an understanding of what the normative (or expert-like) practice should look like after his participation in the cockpit. We will elaborate on this latter point later.

Cockpit talk

How did the learner's talk change after participation and what did it look like in comparison to the expert's? Results show that during the pretest activity the learner did not talk much (only 18 out of 102 utterances of both participants combined), mostly assenting to the copilot's directives and asking questions (e.g., where to read the information from). The expert (copilot) often gave prompts about what, how, and when to do things. Conversely, during the posttest activity not only did the apprentice talk as much as the expert (42 out of 97 utterances), but also we observed a change in the type of utterances he made. In the posttest, the learner asked a lot of questions (16 out of 97 utterances) about the plane's position, altitude, and speed. And more importantly, many of his utterances were in the form of reflections (13 out of 97 utterances), pondering about how speed and altitude relate to each other in the behavior of the plane, and how these were to change if one applied power or flaps. On the other hand, when the expert flew as the pilot, an analysis of the talk revealed that the expert used lots of self-reflections (20 out of 91 utterances). Also, there were a good deal of inquiries back and forth between the expert (pilot) and the learner (copilot), as reflected in the significant proportion of questions and answers in this activity. Again, an analysis of the learner's posttest activity shows that it looks half way between the expert-like behavior and the pretest. This pattern is not quite like the expert yet because the expert (copilot) gave a great amount of prompts about what to do (18 out of 97 utterances), and the learner posed a lot of questions as well as. It also is important to note that when the expert flew the plane, he was trying to instruct, so would naturally comment more than when in an expert solo condition or with another expert, therefore, this "expert" pattern might not reflect these latter types of activities.

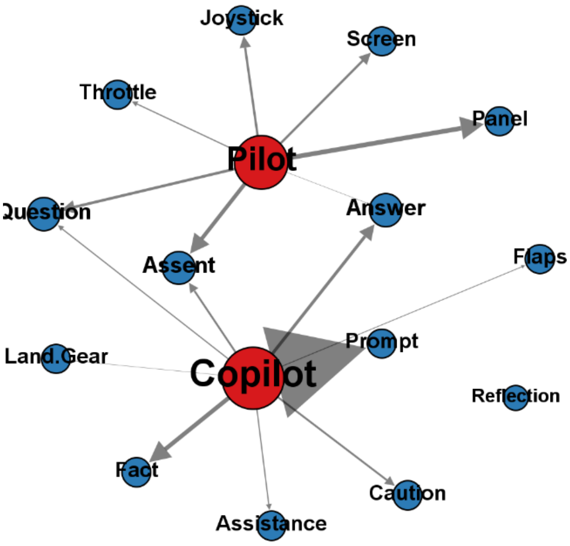
All in all, these trends of change make sense. During the first activity the learner was so overwhelmed by the ongoing demands of the piloting that he did not talk much and limited himself to follow directions on what to do and how to do it. In the posttest activity, the learner was a bit more experienced, was able to talk about the instruments with certain confidence, and was able to articulate his moves out loud. Certainly, the posttest talk was the product of previous interactions with the expert in that the meaning of the terms was situated in the use of the information coming from the instruments. We believe that not only was the learner appropriating the expert's discourse, but also his ways of reasoning about how the variables (e.g., speed, altitude, power) relate to each other. The following example illustrates the learner's emergent knowledge in practice. The excerpt lasts one minute and nine seconds. The copilot (the expert) is prompting the pilot (the learner) that he should reduce the speed because the airplane is going too fast. After this statement, the pilot ponders whether to use more flaps to reduce the altitude. The expert (the copilot) notes that the flaps do not reduce the altitude but the speed (because they create a drag force in the wings). However, the pilot (the learner) pointed out that a decrease in speed would also make the plane descend. And although the goal was not to descend yet (because they were not at the desired distance to the runway), this learner's reflection illustrates nicely why planes descend (not because the plane's nose points down but because of a reduction in its velocity).

Expert:	We are still far from the runway. Look at the GPS, it says we aren't close to the coast yet.
Learner:	Okay.
Expert:	You are going too fast, but try not to lose altitude.
Learner:	Can I use more F... Flaps?
Expert:	To lose speed?
Learner:	Yeah, to lose altitude.
Expert:	No, the Flaps reduce speed, not altitude.
Learner:	But if we lose speed then the plane would lose altitude, right?
Expert:	Yes, it's a consequence of that.
Learner:	Yeah. So… Yes, it would be good.
Expert:	Flaps 20 [*Flaps go down*]
Learner:	Ah, but I am gaining altitude again.
Expert:	Sure you do, because the flaps create resistance and lift, but the nose… but putting the nose down it will increase the speed. Try to keep the nose as horizontal as possible.

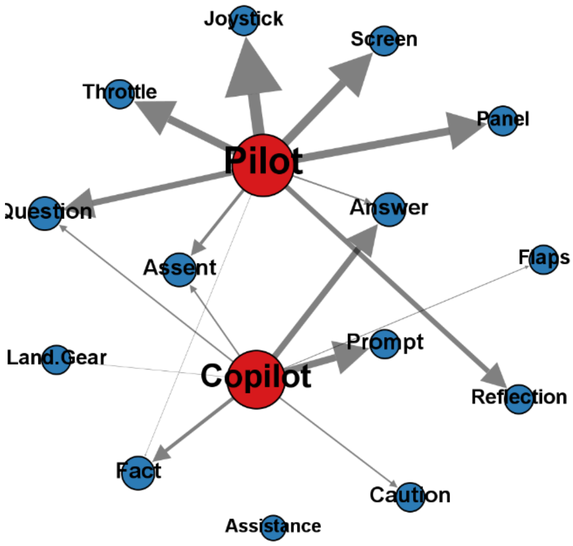
Situated-action networks

Thus far, we have been asking questions about changes in individual pieces of information for the learner's actions. Now, we ask how the overall functional model of actions has changed. According to our original goal, we want to trace the development of the learner's participation. Therefore, SAN graphs were constructed to represent the relations, densities, and centralities of both actors during each activity. Three networks, one for the pretest, one for the posttest, and one for the expert, along with a graph for the trajectory of the center of masses, are presented in Figure 6. The first network, the pretest, serves as the baseline representation of the learner as a newcomer; the second network, the posttest, contrasts this baseline to a more engaged practice; and the third network, the first expert pilot activity, serves as a baseline to understand how much the expert exerted control on the piloting activity.

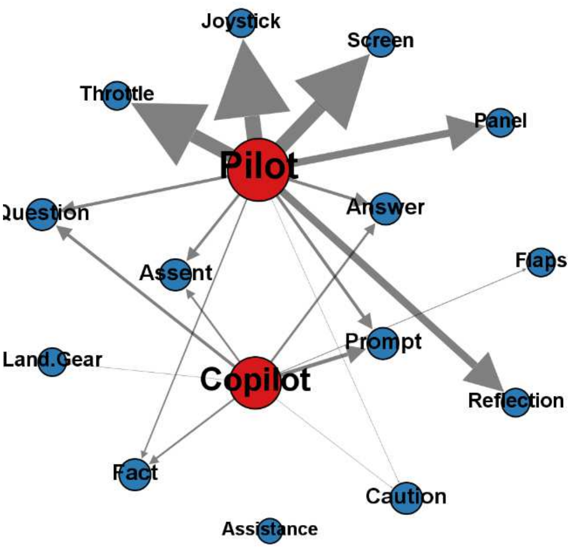
In the pretest network (see Figure 6.a), when the novice played the role of the pilot for the first time and the expert the copilot, the pilot's degree (i.e., the amount of the pilot's actions) was 36, and the network density (i.e., the average degree across all nodes) was 27. In the posttest network (b), which was after five landing activities in which the novice played the alternate role of the pilot or the copilot, it is apparent that his node played a more engaged role in the network. The pilot's degree for network (b) increased to 84, and the network density increased to 36.2. In addition, network (c) represents the activity when the expert played the role of the pilot. It shows a high level of engagement on the part of the pilot, with a high degree of 112 and high average network density of 40.6.



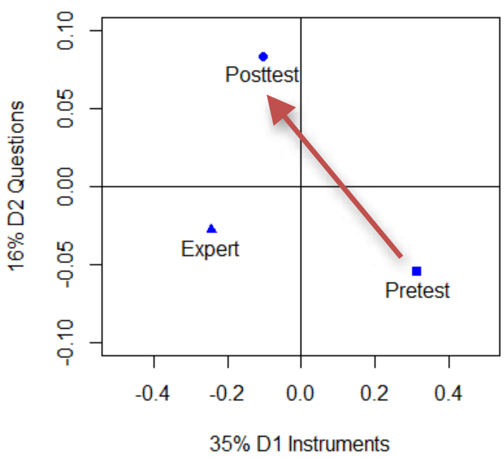
(a) Pretest Network – Novice Pilot



(b) Posttest Network – Novice Pilot



(c) Expert Network – Expert Pilot



(d) Learner's Progression

Figure 6. Situated-Action Netwoks for the novice pilot and the expert pilot

Finally, the participation progression can be seen in Figure 6.d, where center of masses are plotted along the first two principal components. The first component represents the amount of gaze at the instruments (and represents 35% of the variance) whereas the second component represents the amount of questioning (and represents 16% of the variance). It is apparent that the apprentice's dependency on the instruments decreased from the pre- to the post-test activity as revealed by a lower posttest center of mass on the first dimension, that is, the center moved to the left. However, the apprentice also asked a higher amount of questions during the posttest compared to the pretest, as revealed by his higher center of mass on the second dimension, that is, it moved up in relation to the pretest. On the other hand, the expert's network has a low center of mass on both instrument and questioning dimensions. What stands out as really powerful from these network representations is that the researcher can track the learner's progression from several pieces of evidence (i.e., talk and instrument use) simultaneously. These multivariate analysis supports the visualization of complex, distributed activities in ways simple causation models cannot do. An examination of these networks can pick up subtle relationships between participants and practices that may not be possible from an analysis of disconnected pre- and posttest difference scores.

Discussion and conclusion

In this paper I have outlined a new visualization tool called Situated-Action Networks. With these visualizations researchers can gain insights into the learning process as a situated activity in which the learner's progression from peripheral types of participation are developed into more engaged forms. Because this tool supports the integration of actors, speech acts, and artifacts into one mode of representation, i.e., the network graphs, it allowed us to integrate the evidence of individual pieces of information to create together a complex visualization of the learning process. In the future, we will try to capture these interactions with the use of automated techniques. In some of our latest analyses we are exploring the use of point-of-view cameras (i.e., egocentric head-mounted cameras), digital pens, and infrared cameras that can capture anthropometric features. However, I have shown how this visualization technique can be produced from more traditional sources of video data, and I believe it can be useful in computer-based learning environments as well. In addition, we are thinking about how to incorporate relevant interactional information that the networks still do not represent, such as the moment-by-moment interactions between participants. For instance, showing how the pilot's use of the controls vary in response to the copilot's directions. Finally, I envision a possible use of situated-action networks to construct generalizable learning claims about participation levels in order to assess the learner's abilities in open-ended, ill-structured scenarios. A limitation of SAN is that, in order to provide the necessary points of comparison, it requires that several iterations of relatively similar learning activities, or of the same learning activity, are implemented.

References

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32.

Cole, M. (1996). *Cultural psychology: A once and future discipline*: Harvard University Press.

Engeström, Y. (1987). *Learning by expanding. An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsulti Oy.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American psychologist, 53*(1), 5.

Hutchins, E. (1995). *Cognition in the Wild* (Vol. 262082314): MIT press Cambridge, MA.

Hutchins, E., & Klausen, T. (1996). Distributed cognition in an airline cockpit. *Cognition and communication at work*, 15-34.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*: Cambridge Univ Pr.

Leont'ev, A. N. (1974). The problem of activity in psychology. *Soviet psychology, 13*(2), 4-33.

Roth, W.-M. (2007). On Mediation Toward a Cultural-Historical Understanding. *Theory & Psychology, 17*(5), 655-680.

Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on language and social interaction, 26*(1), 99-128.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*(4), 289-327.

Shaffer, D. W., & Clinton, K. A. (2006). Toolforthoughts: Reexamining thinking in the digital age. *Mind, Culture, and Activity, 13*(4), 283-300.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., . . . Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning.

Vygotsky, L. S. (1978). Mind in society (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.): Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language (A. Kozulin, trans.)*. Cambridge, Massachusetts: MIT Press.

Wasserman, S. (1994). *Social network analysis: Methods and applications* (Vol. 8): Cambridge university press.

