Evaluating a User-Elicited Gesture Set for Interactive Displays

Jens Heydekorn, Mathias Frisch, Raimund Dachselt

Otto-von-Guericke Universität Magdeburg
User Interface & Software Engineering Group

Abstract

Recently, many studies were conducted which focused on eliciting gestures from users in order to come up with gesture sets for surface computing. However, there are still many questions to clarify concerning the value of this method regarding to the usability of such gesture sets in real systems. In this work, we contribute a usability test of an implemented gesture set based on user suggested pen and hand gestures for node-link diagram editing on interactive displays. The results of the usability test gave valuable insight in how users interact spontaneously with such a gestural interface. In particular, we found that the methodology of eliciting gestures from users reveals *what kinds* of gestures users prefer but that it does not necessarily show *how* they are applied. Beyond that, we observed how participants differentiate between touch and pen within complex workflows.

1   Introduction

In recent years, several studies have been conducted applying the gesture eliciting approach (Nielsen et al. 2004) (Epps et al. 2006) (Wobbrock et al. 2009) (Micire et al. 2009). This means gestures are suggested by users. Thereby, the effect of a gesture is shown to participants and after that they are asked to perform the respective gesture. The general goal of this approach is to create more intuitive and natural gestural interfaces. However, users are not interaction designers and therefore they have no insight in new modalities and appropriate concepts (Hinckley et al. 2010). Moreover, during the eliciting process gestures are investigated separated from each other and usually no feedback is given to the participants. Because of that, there are several open research questions concerning user-elicited gestures. How useful are those gestures in a running application? Do users behave differently during the eliciting process and when they apply gestures during a real workflow with visual feedback? How do users differentiate between modalities in real workflows compared to the eliciting process?

As a contribution of this work, we present the results of a usability evaluation of an implemented gesture set which is based on the user eliciting approach. In particular, for the study we applied a gesture set presented in our previous work (Frisch et al. 2010). It was designed to support manual editing of node-link diagrams on interactive surfaces and consists of multi-touch and pen gestures. The results show that most of the gestures are spontaneously discovered and easily remembered. Furthermore, we found that many of the available gestures were performed in different ways, depending on the context. This behavior was never observed during the eliciting process. However, it should be considered when respective gestures shall be implemented in a real world application. Finally, we observed how touch and pen gestures are phrased together within complex workflows (Buxton 1995). Based on these findings, we reflect about the pros and cons of the user eliciting approach and could improve our prototype.

2   Related Work

2.1   Design of Gestures for Interactive Displays

In recent years various technical approaches for interactive multi-touch enabled displays have been introduced (e.g. (Dietz & Leigh 2001), (Rekimoto 2002), (Han 2005)). They allow the realization of gestural user interfaces and several one- and two-handed gestures were developed to manipulate digital content. The majority of these systems apply gesture sets defined by experts (e.g. (Wu & Balakrishnan 2003), (Wu et al. 2006), (Brandl et al. 2008)). To come up with natural and intuitive gestures, in some cases users are involved in the design process, usually by observing them "in the wild", e.g. while they are manipulating physical objects (Cao et al. 2008) (North et al. 2009) (Hinkley et al. 2010).

Another approach to come up with natural gestures is the user-elicitation method. Here, users propose gestures for concrete given tasks. The participants are aware of the respective input and output devices but do not get any feedback from the system. A general process for the user-elicitation method is presented by Nielsen et al. (Nielsen et al. 2004) and was applied in several works (Frisch et al. 2009) (Wobbrock 2009) (Micire et al. 2009). Recently, Morris et al. (Morris et al. 2010) compared user preferences of user-authored and expert gestures. The result of this comparison showed that participants mainly favored gestures suggested by users. Beyond that, further observations such as the preference for simple gestures or physical discomfort of multi-touch gestures are described. However, during these studies the given tasks and observed gestures are treated in an isolated way. Compound tasks and respective interactions are not considered. In contrast to that, the experiment presented in this paper evaluates a prototype application which implements user-suggested gestures. Amongst others, we investigated how users interact spontaneously with such a system and how gestures are linked together in workflows.

2.2   Combination of touch and pen

Several works investigate the combination of pen and multi-touch input on interactive displays. Brandl et al. (Brandl et al. 2008) suggest general design principles based on the theoretical framework of Guiard (Guiard 1987). They found that the combination of pen and touch is superior against other combinations in terms of speed, accuracy and user preference. Zeleznik et al. (Zeleznik et al. 2010) developed an algebra system, and Hinckley et al. (Hinckley et al. 2010) presented a digital drafting table application. Both systems support hybrid pen and touch input. In particular, Hinckley et al. introduced an interaction principle which they summarize with *pen writes, touch manipulates and pen + touch yields new tools*. Both systems are designed by experts and were evaluated by informal usability studies. The gesture set evaluated in this paper is based on the results of a user-elicitation study which is described in our previous work (Frisch et al. 2009). It also considers pen and touch input. However, we did not observe the spontaneous distinction of both modalities during the elicitation process. As a consequence, we treated both modalities equally in the resulting gesture set. The observations presented in this work clearly show in which way users spontaneously differentiate pen and touch. These results allowed us to consider combinations of touch and pen which are intuitively performed by users and to adopt the gesture set in the respective way.

3   Usability Evaluation

The usability evaluation presented in this paper is based on a collection of multi-touch and pen gestures suggested by users (Frisch et al. 2009). It covers nine basic diagram editing tasks. In (Frisch et al. 2010) we analyzed these gestures to create a consistent gesture set from this collection (see Table 1 in the Appendix)[1]. During the analysis, occurring ambiguities were resolved. To achieve that, a semi-transparent border region around every node was introduced. The border can be used for inserting edges in different ways (e.g., by sketching or tapping sequentially). In that way, slight adoptions were made but no further gestures were added. Instead, the most popular ones were reused for different tasks and users can still interact directly with the content (e.g., without invoking menus). Furthermore, the gesture set supports two prevalent mental models - *sketching* and *structural editing*.

For the usability evaluation, some visual feedback was added to the prototype. When a touch is recognized, a circle (slightly bigger than the finger or pen) appears. Beyond that, a node and an interactive border respectively is highlighted when its interior is touched by finger or pen. A group of nodes is visualized by an enclosing semi-transparent bounding rectangle, which can be used in the same way as single nodes. Furthermore, if single or grouped nodes

---

[1] an overview is available on *http://wwwisg.cs.ovgu.de/uise/Forschung/Projekte/DiagramInteraction/index.php.en*

are deleted by dragging them off screen, their color is set to gray as soon as they cross the border of the display. When the finger or pen is lifted the respective node is deleted and disappears.

3.1   Design of the Study

The study applied a within subject design consisting of two phases. In phase one, participants were asked to solve diagram editing tasks without explanation. Our goal was to see which gestures users would detect without help, how many attempts they would need for that and if they are able to solve given tasks without knowing the system. For the second phase, all gestures were explained to the participants and they were asked to train them. During the following tasks they could freely apply all available gestures. In that way, we were able to make thorough observations of their behavior during longer workflows. In both phases they could chose freely which modality (touch or pen) they want to apply.

Ten participants (2 female, all right-handed) took part in the study. Nine of them are employees of the computer science department and one is a student. All of them have knowledge about node-link diagram notations such as UML, but they are not everyday modelers. None of them is an HCI expert. Two of the participants are regular users of touch-enabled devices such as iPhones. Four of them are using devices with pen input such as TabletPCs. The study was conducted at a self-made FTIR multi-touch tabletop (resolution: 1280x800 pixels, size 100x80cm). The aim of the respective task was shown by a projection on the wall in front of the tabletop during the whole task. All activities were captured with a video camera from above the table. Furthermore, the participants stood in front of the interactive display.

3.2   Measurements & Procedure

We mainly analyzed the captured video material. The log files of recognized gestures (including modality, screen coordinates and timestamp) were only used in unclear situations. For the first phase we counted the attempts to accomplish the given task, considered used approaches – *sketching* and *structural editing* – and the applied modalities. For the second phase we counted how often a certain gesture was performed and the used modalities. Beyond that, we measured the completion time for both tasks. All counting was done by at least two observers.

At the beginning, every participant got a short demonstration of the general touch and pen handling at the tabletop. After that, participants practiced with a simple photo application. It allows translating, rotating and scaling images by simple touch or pen gestures such as dragging or pinching. After the participants became familiar with the surface and the usage of this application the initial training ended.

In the **first phase** of the usability test the participants were asked to work with the diagram editor without any help or explanation. Furthermore, they should speak aloud about their activities. The phase consisted of four sub-tasks: **"**Create a simple diagram consisting of three nodes, two undirected and one directed edge" (T1.1), **"**Select and copy two nodes" (T1.2), "Create a dashed directed edge" (T1.3) and "Delete the copied nodes" (T1.4). We set a time limit for each task to prevent frustration of the participants (6 minutes for T1.1. and 2 minutes for T1.2-T1.4) and interrupted them when the time limit was exceeded.

Before the second phase of the usability test, participants got an introduction to the gesture set. Each gesture was explained to them and had to be performed with touch and pen. In a second training phase, participants were asked to do a workflow consisting of five sub-tasks in a guided way. Thereby, they were allowed to ask for gestures. The final diagram for training contained 21 nodes, 10 undirected edges, 8 directed edges and 2 dashed edges. The whole gesture set was required to successfully create the final diagram.

The **second phase** of the test consisted of two similar tasks (T2.1 and T2.2). For each of them, participants were asked to recreate a given node-link diagram on their own. Again, the diagram was projected to the wall in front of the tabletop. The complexity of the diagrams was similar to the one of the training task (18 nodes, 8 undirected edges, 7 directed edges, 2 dashed edges). Five of the nodes differed from standard size and had to be scaled. At the end of the second phase, participants filled in a questionnaire with 14 questions. They rated the available gestures and the usability of the system on 5-point Likert scales (see Figure 1). A whole procedure had an average duration of 45 minutes.

4   Results

4.1 Observations from phase one

Mental models and spontaneously found gestures. All participants succeeded in creating the small diagram of T1.1. Thereby, spontaneously found gestures were mainly based on the *sketching approach*. Seven participants drew outlines of the nodes with respective sizes. However, five participants scaled nodes by spontaneously applying a pinch gesture. During T1.2 the grouping of nodes was also achieved by all the participants mainly by encircling the nodes using the pen. Only one participant created a group by tapping. Corresponding to our expectations on the rather abstract copy function, five participants gave up on the subsequent copying task. They mainly tried different variants of tapping or tried to invoke a menu. In T1.3 six participants gave up while trying to create a dashed edge. Four of them performed the sequential crossing gesture too fast, so that it was not recognized by our system. Only two participants performed the rake gesture. All participants were able to delete nodes (T1.4), in the majority of cases by dragging them off the screen. Only two participants performed wipe gestures. This functionality was often discovered before the actual task was

given to the users. For example, when participants dragged nodes created by accident aside to the edge of the display.

Interactive Border Region. All of the participants rather quickly understood the role of the interactive border region of a node. At least after four attempts all of them understood the concept and were able to sketch edges (exception was one user who needed 10 attempts). However, we observed that several users hit the border region when the view was zoomed out and they tried to quickly reposition tiny nodes. As a result, they started to draw a stroke instead of dragging the node.

Applied modalities. Participants started drawing nodes and edges mainly with the pen (see Figure 2 centre). This behavior applies also for grouping nodes. Two participants ignored touch interaction completely, even when it came to the copying task which could only be solved bimanually. In general, combinations of touch and pen were hardly used spontaneously. For copying nodes, only one participant applied a touch+pen combination.

4.2   Observations from phase two

Gestures and Mental Models. The gestures were easily remembered by the participants. They already applied them without further asking during the training tasks. An exception was the gesture for creating directed edges by holding the node and dragging a rubber band from the interactive border with finger or pen. Three participants asked again how to perform it and then applied it repeatedly. In all, this gesture was performed 43 times during Task 2.1 and 2.2. Some participants were unsure how to perform the proper gesture for creating edges by tapping sequentially on the border of the source and target node. However, they did not request help but needed some attempts to perform the gesture properly. All participants applied *sketching and structural editing* interchangeably. For example, in some cases standard-sized nodes were created by tapping and nodes of different size were sketched and then copied. Other users sketched nodes and scaled them afterwards to the proper size. Furthermore, some created undirected edges by tapping and changed them to directed edges by drawing an arrow head.

Applied modalities. Sketching nodes and grouping nodes by encircling was mainly performed with the pen (in 86% and 61 % of the cases, respectively). In contrast to that, tasks based on *structural editing* were mainly done by touch. This could be observed for dragging nodes in order to reposition or delete them (92% of the cases), creating edges by tapping (81% of the cases) and selecting nodes by tapping (75% of the cases). In these cases, users often tucked the pen between the fingers or temporally swapped the pen to the non-dominant hand. The same behavior could be observed for bimanual gestures such as pinch gestures or holding and dragging for copying nodes. The copying gesture, for example, was applied 82 times, but only in one case it was performed with a combination of touch and pen. However, no clear preference for a particular modality was found for deleting elements by means of a wipe gesture and creating dashed edges by sequential crossing.

Further Observations. The majority of participants created sub-graphs of about 6 to 10 nodes. They created the nodes first at the proper positions and then connected them. After that, they continued to create another part of the diagram in the same way. We also observed that participants zoomed out rather quickly: four of them did that right at the beginning of the task. However, participants did then not zoom in again. They continued working with rather tiny nodes, although it was hard for them to hit the border of a node accurately. However, participants did not switch to the pen to get a higher precision (see Figure 2 right).

Completion Times. To get an insight of the learning effects, we compared the completion times for both tasks in the second phase of the test. The mean completion time for the first task was 5:25 minutes and for the second task 4:00 minutes, which corresponds to a reduction of 26%. However, a Wilcoxon signed rank test did not reveal a significant effect (z=-1.599, p>.05). In the best case the completion time halved on the second try (three participants). In contrast to the mostly better completion times, no clear change in the prevailing mental model of the used gestures was observed.

User Satisfaction. The participants were asked to rate the system by a questionnaire at the end of the test (see Figure 1). It is clearly recognizable that users were not satisfied regarding the feedback of the system, what we also observed during the study. However, the pen got a better rating than touch, due to the "fat finger" problem and because users often worked with small nodes by applying fingers. Although not all resulting errors are due to the used technology, users did not realize their imprecise touches at all. All in all, the participants perceived the combination of pen and touch as acceptable and beneficial.

5 Discussion

Variations of Gestures. As we observed during the usability test, the same gesture can occur in different ways. In most cases, variations took place in terms of *speed* and *precision*. For example, positioning nodes was usually achieved by rather carefully performed dragging gestures. In contrast to that, deleting nodes by moving them off screen was often done by flicking away respective nodes. When users quickly wanted to get rid of a node created by accident, they acted too hastily for the system or they hit the interactive border of the node. As a result, participants had to drag more carefully and slower again. We see the reason for these "inaccurate" gestures in the circumstance that during a particular workflow some
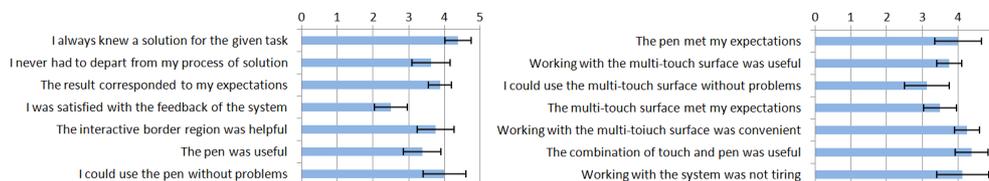
Figure 1: Results of the questionnaire.

gestures have less valuable contribution than others for achieving the goal of the task. Similar situations were observed for changing a solid edge to a dashed one by sequentially crossing the edge. Again, many users performed rather quick and short movements which were not correctly recognized. In this case, users simplified the gesture in order to be faster. Beyond that, repeating the same gesture several times (such as crossing the edge) leads to imprecision due to fatigue effects. Therefore, simplified gestures should be considered in the implementation. For example, deleting nodes should also be possible by a flick gesture. Flicking a node should result in moving it towards the edge of the display, where it stays visible for a few seconds until it is deleted. In this way, deletion by accident could be prevented. Concerning the change of a solid edge to a dashed one we also suggest sequential flick or even tap gestures. They should be used in addition to the rather slow gestures currently implemented.

The user-elicitation methodology is certainly valuable to find *what kinds* of gestures are preferred. It can be applied to get a feeling for gestures in a new unexplored domain (such as diagram editing). However, as we have seen from our study, eliciting gestures from users does not necessarily reveal the *different ways* a gesture is performed (e.g., concerning *speed* and *precision*). This depends on particular situations within compound tasks. Observing gestures isolated from each other – as it was done in recent work on user-elicited gestures – may not be sufficient to create easy to use applications.

Separation of modalities. In contrast to the eliciting process we observed how users spontaneously differentiate between pen and touch interaction. The pen served for *sketching* diagram elements or for encircling nodes. Touch interaction was applied for activities based on *structural editing,* such as dragging and copying nodes and creating edges by tapping. The latter is done by holding the node with one finger and dragging the copy off from it with a finger of the preferred hand (according to Guiard (Guiard 1987)) (see Figure 2 left). This stands in contrast to the concept of Hinckley et al. (Hinckley et al. 2010) who suggest copying with a combination of touch and pen. We think that there is one more ergonomic reason against this method. If dragging off a copy is allowed only with the pen held in the preferred hand, then copies can only be dragged to one side in a comfortable way (e.g., to the right for right-handed users). Otherwise, users would have to cross over their arms, which is inconvenient. Repositioning objects for copying them is unfavorable for diagram editing as it could destruct the layout. For deleting diagram elements by wiping and creating dashed edges, we did not observe any preference of modalities. We assume that the reason for this
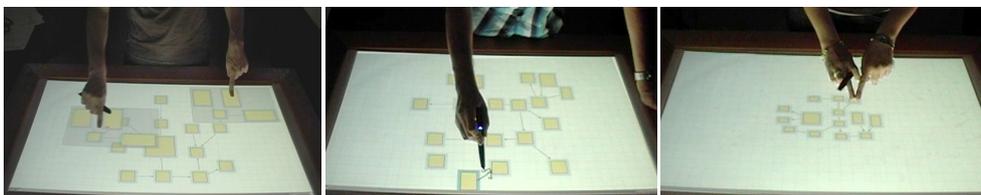
Figure 2: Pictures from the study. Left: User bimanually copies a group of nodes. Center: Participant draws an edge with the pen. Right: Participant pinches on small node with the pen tucked between the fingers.

lies in the similarity to physical interaction. Sketches on paper can be scratched out with the pen and on whiteboards they can be erased by wiping with fingers.

The observed distinction of modalities can be easily integrated in our prototype. This allows users to sketch edges in a more ad hoc way, as they do not have to care if they hit the interactive border of a node. However, the border is still necessary to distinguish dragging nodes from creating edges by sequentially tapping with fingers.

6 Conclusion and Future Work

We contributed an evaluation of user-suggested gestures. The results showed that users are able to work with such a system spontaneously. In our opinion this is a strong indication for the intuitiveness of systems following the user-elicitation method. Nearly the whole gesture set was utilized after a short training to achieve complex editing tasks. Nevertheless, we also found correlations of commands, mental models and modalities which were not observed during the elicitation process. One example is the distinction of modalities: the pen is applied for *sketching*, whereas *structural editing* is done by touch. Moreover, we found that users can suggest *what kind of* gestures they would prefer. However, a working system is necessary to observe *in which way* these gestures are applied during compound tasks. This should be considered to implement easy to use applications. For future work, we will adopt the prototype according to our findings. More comprehensive usability studies will be conducted concerning user acceptance and experience.

References

Brandl, P., Forlines, C., Wigdor, D., Haller, M., and Shen, C. (2008). Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In: Proc. AVI '08, ACM, pp. 154-161.

Buxton, W. A. (1995). Chunking and phrasing and the design of human-computer dialogues. In Human-Computer interaction: Toward the Year 2000, Eds. Morgan Kaufmann Publishers, pp. 494-499.

Cao, X., Wilson, A.D., Balakrishnan, R., Hinckley, K., Hudson S. E. (2008). ShapeTouch: Leveraging Contact Shape on Interactive Surfaces. In: Proc. of TABLETOP '08, IEEE, pp. 139-146.

Dietz, P. and Leigh, D. (2001). DiamondTouch: a multi-user touch technology. In: Proc. of UIST 2001, ACM, pp. 219-226.

Epps, J.; Lichman, S. & Wu, M. (2006). A study of hand shape use in tabletop gesture interaction, In: Proc. of CHI 2006 Ext. Abstracts , ACM, pp. 748-753.

Frisch, M., Heydekorn, J., Dachselt, R. (2009). Investigating Multi Touch and Pen Gestures for Diagram Editing on Interactive Surfaces. In: Proc. of IST '09, ACM, pp. 167-174.

Frisch, M., Heydekorn, J., Dachselt, R. (2010). Diagram Editing on Interactive Displays Using Multi-Touch and Pen Gestures. In Proc. of Diagrams 2010, Springer, pp. 182-196.

Guiard, Y. (1987). Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model, The Journal of Motor Behavior, 19(4), pp. 486-517.

Han, J. Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. In: Proc. of UIST'05, ACM. pp. 115-118.

Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H., and Buxton, B. (2010). Manual deskterity: an exploration of simultaneous pen + touch direct input. In: Proc. of CHI EA '10. ACM, pp. 2793-2802.

Micire, M., Desai, M., Courtemanche, A., Tsui, K., Yanco, H. (2009). Analysis of Natural Gestures for Controlling Robot Teams on Multi-touch Tabletop Surfaces. In: Proc. of IST '09, ACM, pp. 41-48.

Morris, M.R., Wobbrock, J.O. and Wilson, A.D. (2010) Understanding users' preferences for surface gestures. In Proc. Graphics Interface 2010. Canadian Information Processing Society, pp. 261-268.

Nielsen, M., Störring, M., Moeslund, T.B. and Granum, E. (2004). A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In: Int'l Gesture Workshop, LNCS vol. 2915, Springer, pp. 409-420.

North, C., Dwyer, T., Lee, B., Fisher, D., Isenberg, P., Robertson, G., and Inkpen, K. (2009). Understanding Multi-touch Manipulation for Surface Computing. In: Proc. of IFIP TC 13, Part II. LNCS, vol. 5727. Springer, pp. 236-249.

Rekimoto, J. (2002). SmartSkin an infrastructure for freehand manipulation on interactive surfaces. In Proc. CHI '02, ACM, pp. 113-120.

Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In: Proc. of CHI '09, ACM, pp. 1083-1092.

Wu, M. & Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In: Proc. of UIST'03, ACM, pp. 193-202.

Wu, M.; Shen, C.; Ryall, K.; Forlines, C. & Balakrishnan, R. (2006). Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces, In: Proc. of TABLETOP '06, IEEE, pp. 185-192.

Zeleznik, R., Bragdon, A., Adeputra, F., Ko, H. (2010). Hands-on math: a page-based multi-touch and pen desktop for technical work and problem solving. In: Proc. of UIST'10, ACM p. 17–26.

Contact

Jens Heydekorn
Otto-von-Guericke-Universität Magdeburg
Universitätsplatz 2, 39106 Magdeburg

7   Appendix

Table 1: Condensed overview of the gesture set for diagram editing. The interactive border region of a node can be used to create edges (see task 2).
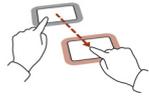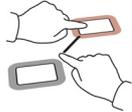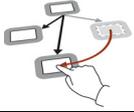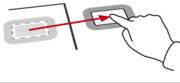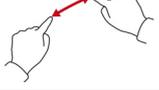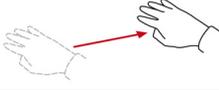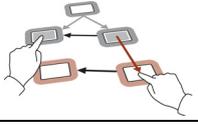
Task	Gestures
1. Create Node	a) tapping b) sketching c) copying by hold + drag
2. Create Edge	a) dragging (rubber band) b) sequential tapping or hold + tap c) sketching
3. Select Node(s)	a) tapping b) encircling
4. Move Node(s)	Dragging
5. Delete Node or Edge	a) wiping b) dragging to off screen
6. Resize Node	pinch gesture on node for non-uniform scaling
7. Zoom & Pan Diagram	a) Zoom: pinch gesture on background b) Pan: drag with three or more fingers on background
8. Copy Sub-graph	copying by hold + drag
9. Change Edge from Solid to Dashed	a) "rake" gesture b) sequential crossing