7

Expert evaluation of the god-like interaction framework

This chapter discusses the real world applicability of the god-like interaction framework. The framework was designed to facilitate a number of experts coming together to discuss a problem and work towards a solution. It also facilitates experts communicating situational and navigational information to remote field workers. There are a number of real world scenarios that operate in this manner including: military operations, fire-fighting and search and rescue. The expert evaluation presented in this chapter was conducted with a participant from a military background. This chapter reports on the results of the evaluation which includes generalisations about the god-like interaction framework that can be applied to all domains, not just the military domain.

7.1 Evaluation

The Defence Science and Technology Organisation [1] (DSTO) is an Australian government defence agency that researches science and technology to improve Australian defence capabilities. This evaluation involved a participant from the DSTO's Land Operations Division. The participant had 15 years experience in the Army working on command and control and human-computer interaction research. The evaluation involved an introduction to the god-like interaction metaphor, a demonstration of the various god-like interaction techniques, an overview of the technol-

[1] http://www.dsto.defence.gov.au

ogy behind the system, and a discussion regarding the limitations of the techniques. The following captures observations made by the participant about the real world practicality of applying this technology to military operations.

7.1.1 General configuration

The military use 2D maps to visualise a geographical area where an operation is taking place. The maps are either displayed horizontally on a table surface or vertically on a wall depending on the task. Communication between personnel in the command center and field operatives is verbal through the use of radio. Experimentation with GPS tracking as well as visualisation of GPS tracked data is under-way, as described by the participant:

> Earlier this year we ran an experiment looking at soldiers running around Edinburgh airport with the ability to do what we call blue-force track. So each of the soldiers had a GPS, and the commander was on a laptop with a map. Similar to the satellite picture, he could see where everyone was. He was coordinating them dynamically doing certain tasks. So it is similar [to how the GPS data is represented on the HOG table] but it only needed a two dimensional representation, and an icon of where the people were. Because he could get that dynamic feedback as to where they were, whether they were off track or not. And he had some measure, through things like satellite photos of things like buildings. So he would say, "go to the building to the left" type things or "you need another 500m to go to this place".

7.1.2 Virtual Landmarks

Military manoeuvres are frequently conducted in areas of sparsely populated terrain. There may be no physical landmarks for people to refer to. Alternatively there may be landmarks visible to field operatives that are unknown to command center personnel. The participant indicated that the ability to create a virtual landmark is useful for communicating and coordinating people in a military context:

> This Vegemite jar certainly could work, because in terrain which is quite sparse, if you could insert an object to allow people to coordinate, that could be quite useful.

> Right now, it [navigation] is with respect to another object. If there aren't other objects to reference they are given a grid reference (a location on a map).

People, places and object are not currently referred to by a GPS coordinate. Where possible, an object's location is described relative to another object. If no other object is present then it is given with reference to a grid location. As described above, GPS information is being investigated.

7.1.3 Pointing

Whilst employing hand gestures for virtual landmark generation could be quite useful as discussed above, gestures for indicating an action or item of interest needs to be compared to other approaches. The participant was concerned that a pointing gesture was not the most efficient way to indicate the importance of something:

> I guess the god-like finger is not quite as useful for the military because they tend to be very focused on their environment and if you give them cues that they have to be looking for, like "that building" or "that feature", because they are already thinking about that building, or that feature. They don't have to translate mentally a big thumb. It's that extra cognitive effort in translating an icon, that they don't need to relate to their environment. So they will say, "that building" or, "that feature", or, "that road intersection" because they are already thinking about it. ... having the icon which is the hand with a finger probably gets in their way. Versus something that emphasis the actual object of interest (like the building flashes). Something like that would be better, because you get that immediacy.

This is a complex issue that deals with many human factors. Defining the best way to indicate an interesting item requires careful consideration of the specific problem as well as consideration of what is technically achievable. As such there are many ways to indicate targets. For military purposes, it is necessary to reduce the cognitive effort with respect to target identification. It is possible that methods for targeting other than highlighting with hand gestures are more efficient. This would certainly be the case if the hand gestures were not obvious or there was some ambiguity in the object being pointed to.

7.1.4 Persistent Information

I discussed Augmented Post-It notes (see Section 3.3.1.4) with the participant as a way to provided persistent information or to label objects in the world.

> In the military context, they always want to do that sort of thing. They will come across something (things in the environment - objects of interest) and want to give it a name. The issue is, how easy is it to facilitate the writing of the name. To be able to attach a label to an object is very handy.

From the discussion, the Augmented Post-It notes appear to be a convenient way to label objects of interest in a military context. As discussed in Section 7.1.8 there is also a need to capture this information in a format understandable by a computer.

7.1.5 Photo-realism

The god-like interaction framework was designed to facilitate quick and intuitive communication. It accomplishes this by capturing an interaction and displaying it for human interpretation. Therefore, by their very nature these interaction are very realistic and visually match the actual interaction. Props used on the table will also appear as photo-realistic representations of their physical original. As discussed by the participant this could pose a problem in a military context:

> So my question is, "What is the purpose of having a photo-realistic depiction of an object, and why do you have to object there?" If I'm only labelling things, if I'm only placing objects in to geo-spatially locate a label ... the photo-reality is not that important. And so the purpose is the essential thing. And it could be that the photo-realism gets in the way, by confusing people because of the accuracy. And the reason why you are putting it there may have nothing to do with having a really high quality image. You need to decide, what is the essence of why you are putting an object there and deciding what is good enough for that. So you know, having an arrow might be adequate.
>
> Because if it becomes very intuitive that the hand of god has told you to go there, just with a finger. But if there is a coke can in the environment, maybe there was a coke can there anyway and so it might not jump out as much as the hand. So there does have to be a trade off. You need to

> think through about which is the most efficient way, to do what I want to do.

The god-like interaction framework addresses the problem "How do you supply that arrow?" in a meaningful or intuitive way. It is easier just to point to something than it is to select it with a mouse cursor. The god-like interaction framework shows a field worker a hand because without doing any processing that is what you have already... a hand. But it is also important to consider how that imagery is depicted outside. For example, is the imagery appropriate for the current situation? Is it confusing? Showing the hand might actually detract from the task. Is it better to show an arrow or to show a hand? Answering these questions is outside the scope of this dissertation, but the god-like interaction framework creates the possibility to answer them.

7.1.6 Shoulder-to-shoulder collaboration

This system is designed to facilitate bringing many experts together physically to discuss a problem and plan a solution. Giving each expert the ability to interact standing shoulder-to-shoulder, meaning they do not have to tell someone else to tell the field workers where something is, they can just tell them or show them directly. The idea of bringing domain experts together is an important component of military operations, as discussed by the participant:

> The whole point of coming together in a team is that it is very efficient. Say that there are 10 people in a team, only two need to agree that this is where the power station is... everyone now realises where the power station is and so it is a very efficient way of doing it. And there is a lot of that interaction going on. I guess the issue is "what does the artefact look like" and in your case it is this blue covered cylinder thing [the HOG table]. And it is a question of working out what the most appropriate artefact that people can use to come together to discuss and to build their shared understanding.

The participant was concerned that the blue screen surrounding the HOG table may restrict its usefulness. Currently, the size of the table is bound by the height of the blue screen (as discussed in Section 4.1.2.1). The participants concern is not restricted to the military domain, any domain where the size of the display needs to be bigger will share this same concern because at some point the height of the

blue screen will be too high to practically work with. Using an alternate mechanism to extract the 3D information from above the table surface could remove this requirement and therefore the system would scale well.

7.1.7 Symbology

Props and Augmented Post-It notes can be quickly and easily used for representing and labelling objects in the real world. In the military context, symbols are almost exclusively used to represent all things known about and discovered in the terrain. The participant discussed their role:

> Right now the military use a 2D map which has a lot of symbology. You know like a lot of icons with a lot of meaning on them. So they are very careful about how they draw their symbols. Because every little stroke of their symbol means something. And so they all come to understand those complex identification of the icons. The 2D map allows them to do quite a lot of complex interactions. Because there is a long history of how they built up this symbology of that map. They don't need a 3D representation because the things like the contours give them that representation. They just build that mental representation by just visualising the contours. And so that is an issue that is sort of maybe a bit different with yours. You require that physical representation for people's understanding. The military have built that understanding of the symbols that allows that interaction.

The military do not currently have a need for 3D symbols as the 2D symbols has been used successfully for many years. The military would need to see a clear benefit for using 3D representations. However, it is not clear whether symbols and 3D representations have been compared in a military context. Perhaps in the future, when the field workers begin to use HMDs to overlay virtual content, there will be a desire for 3D representations.

7.1.8 Object recognition

The god-like interaction framework was designed to facilitate communication directly between people. It was not designed so that the system itself could interpret human actions. The participant wondered if it would be possible to integrate some intelligence into the system to facilitate information capture:

> What about using objects that are cued? So an object might be blue or shiny red. Because sometimes you know you might only need to have a block and everyone including the system knows that this block represents a street sign. That might get rid of some of the processing that you are doing. So you might have 15 different objects, one of them is sign, one of them is a tree, one of them is a Vegemite jar and that might be all you need. Because often you only need to work with a small number of artefacts. Because all you really want to do is read a cue such that this object is going to be identified as whatever in the virtual Tinmith view.
>
> There are certain tasks that you need to do. And so you will need to know that there are certain objects that headquarters require to interact with. In a military context you want to know the enemy, you want to know that there is a bomb there. There are a limited number of objects that you want to know about and so you don't have to have this flexibility of seeing a pictorial representation of a sign. There maybe only 10 or so possibilities relating to the task. So you just pick the best objects for the task. So the important thing is the interaction, god knows about big things. And so he can tell the individuals that something is going to come and bite them... Because the guy on the ground might not be aware. And you need to do fine tuning. What does god know that the individual needs to know?

The need to capture information depends on the task. In a military context it would be beneficial if the system could understand gestures, understand what props had been placed where, and interpret messages written on Augmented Post-It notes. This benefit would also be seen in other domains where a database of information captured over time assisted with some task. Capturing of this information also opens the possibility of changing the representation the field worker sees, it would no longer have to be the actual interaction that took place.

7.1.9 Unconstrained interaction

The real-time nature of the god-like interaction framework means that movement is captured and conveyed. The participant was critical of this ability for a military context:

> I'm not sure that there is enough moving to warrant a pictorial representation. Because there is only a limited number of possibilities. But

obviously there are situations where you might need this flexibility. But it would be a task that is not so well defined so that you can get away with having this flexibility.

In a rigid military context, free flowing movement is not necessary to communicate. But, whilst it may not be suitable for highly structured military operations, it is useful for communicating non-rigid ideas. Perhaps even in structured domains for times when an unforeseen event occurs for which there is no predefined operation, the god-like interaction framework could be useful.

7.2 Discussion

A number of aspects of the god-like interaction framework were appealing to the participant:

- The ability to create virtual landmarks in sparsely populated areas.
- Labelling of objects with Augmented Post-It notes.
- Shoulder-to-shoulder collaboration.

However, there were a larger number of aspects of the system that were not deemed as useful for a military context:

- Flexibility is not essential as typical military operations are highly structured.
- Suitability and accuracy of pointing for highlighting targets.
- Photo-realism.
- Unconstrained interaction.

There are two areas in which the system could be extended potentially improving the usefulness across many domains:

1. Object recognition: while the primary goal of the god-like interaction framework is to directly convey interactions, it would be possible to integrate some technology to monitor events happening on or above the table surface. In this way hand gestures could be recognised, props could be identified and writing could be interpreted. All of this information could be understood by a system. This information could then be used improve the capabilities of the system.

Such an ability would also mean that if the physical prop used on or above the table was not the ideal representation for the field worker, it could be replaced by a more appropriate one.

2. Scalability: the size of the HOG table is currently limited by the height of the blue screen which facilitates background subtraction. Using an alternate mechanism to extract the 3D information from above the table surface could remove this requirement and therefore the system would scale well.

7.3 Conclusion

The military have a methodical approach to their operations which has worked well for many years. Any additional technology needs to have a clear benefit for them to adopt it. In the military, structure and detail are given precedence over ease of use and flexibility. However, the military are constantly evaluating new technology. They are always looking to integrate technology that works well and provides a significant improvement over existing approaches.

In more general terms, the god-like interaction framework is better suited for people who are perhaps not as well trained, or are new to the job. For example, country fire fighters who are volunteers and unfamiliar with the environment they are working in, might benefit more. These people could benefit greatly from visualisation aids that help them to understand spatial relationships better without the need to train for many weeks or even months.