Realistic Virtual Actors for Training in Counterterrorism

Dr Rick Evertsz; Mr Matteo Pedrotti; Mr William Glover
AOS Group
rick.evertsz; matteo.pedrotti; william.glover@aosgrp.com

Abstract. Terrorism has become a common international threat, and is particularly problematic to counter because it occurs in a civilian rather than military context. It is difficult to differentiate terrorists from normal civilians because their strategy is to merge into the crowd and thereby avoid detection. Providing security personnel with the necessary experience to spot terrorists in a crowd can be time consuming and resource intensive. Simulation can provide a valuable complement to current approaches, making them more efficient, cost effective and reducing the manpower and infrastructure required to support training. This paper describes how realistic virtual actors can be used to provide the counterterrorism trainee with experience in detecting abnormal behaviour and using tactics that increase the visibility of a terrorist concealed in a crowd. Our approach is to use the cognitive architecture, CoJACK™, to create virtual actors (civilians, terrorists) that react variably but realistically to the trainee’s and other virtual actors’ actions. These virtual actors are imbued with emotions, such as fear, anger and anxiety, which lead them to react plausibly to events in the scenario. The key point is that these reactions are not scripted but emerge from the interplay between the virtual actor’s cognitive and emotional faculties. This yields a flexible training environment that is responsive to what the trainee does. In many ways, it is more difficult to generate realistic behaviour in a civilian rather than military context - the actions and reactions can be quite subtle. This paper outlines a civilian counterterrorism scenario, set in London, in which there is a terrorist on reconnaissance, diverse types of civilian with varying goals and emotional states. The virtual environment, VBS2™ was used to represent the physical environment and the characters’ embodiments.

1. INTRODUCTION

A number of studies have concluded that terrorism has been on the increase since 9/11, e.g. (Conetta, 2006). Other studies that factor out the terrorism within Iraq and Afghanistan, e.g. (Mack & Nielsen, 2008), conclude that there has been no major increase. The term “terrorism” itself has been difficult to define, partly because of the political nature of the assessment – one side’s terrorist is the other side’s freedom fighter.

Regardless of which assessment one chooses to favour, the undeniable fact is that terrorism is a significant problem. A common terrorist strategy is to maximise casualties by targeting crowded public places. It is also easier for terrorists to conceal themselves in a crowd.

There are many approaches to countering terrorism. Some are strategic, for example detecting terrorists in a crowd by identifying anomalies in behaviour. All of these approaches are important and complementary. The detection of behavioural anomalies in members of crowds can be surprisingly informative. Anecdotally, the hostile intent of the 7/7 London bombers could have been detected from video surveillance alone1. Indeed, the UK Government’s INSTINCT² (INnovative Science and Technology IN Counter-Terrorism) programme was set up to investigate, amongst other things, how science and technology can help with detecting hostile intent in crowds.

Currently, training personnel to detect and deal with terrorists is manpower intensive, time consuming, expensive and relies on one or more fixed training sites. This approach can be made more efficient and cost effective by

---

1 Personal communication; UK counterterrorism expert.
2 http://security.homeoffice.gov.uk/science-and-technology/innovative-science-tech/
complementing it with a virtual training environment, greatly reducing the manpower and infrastructure required to support training. Such an environment could be used to simulate the contrasting behaviour of civilians and hostiles. This would enable the investigation of the impact of security forces’ actions on the behaviour of people, both hostile and benign. The simulation can also be used as a tool for the security forces, to reduce the costs involved in training.

Such a virtual training environment has the potential to supplement standard training in a number of ways, including:

• Providing experience in detecting behaviour that is indicative of someone either about to commit a terrorist act or on reconnaissance to plan a future attack.

• Training in how to prompt terrorists to reveal themselves. For example, having uniformed personnel walk past a suspect to “stimulate” them to behave suspiciously.

• Practice in crowd handling during an incident. How will a crowd respond to an attack? The popular notion of contagion spreading through a crowd, leading to mass panic, is largely false and only occurs in rare cases.

• Providing a low risk “sandbox” environment for evaluating new tactical approaches to dealing with terrorists in a civilian environment.

The success of such applications depends on having virtual actors that behave realistically and dynamically, responding to the unfolding situation much as real human beings would. Realistic behaviour in virtual actors has two main facets:

1. It is vital to accurately model human cognition and emotion and how they impact upon human action.

2. The virtual environment must be capable of expressing the actions that the cognition/emotion model generates.

The cognition/emotion models need only be accurate enough to generate convincing behaviour for the training scenarios being developed. Further accuracy is superfluous and counterproductive because it will likely degrade the performance of the virtual actors. Similarly, the virtual environment only needs to render the actors’ actions in sufficient detail to convince the trainee that the behaviour is realistic.

These two facets are subtle but important. If the cognition/emotion models are inaccurate, no amount of virtual environment richness will convince the trainee that the situation is real. Likewise if the virtual environment cannot express the subtlety of human action, for example facial expression, then the richness of the cognition/emotion models will be lost on the trainee.

This paper describes an initial prototype of such a training environment based upon the cognitive architecture, CoJACK™, and the virtual environment, VBS2™.

2. UNDERLYING TECHNOLOGY
We provide a brief overview of CoJACK and VBS2 in this section. More detail can be found in the context of a suicide bomber scenario in (Evertsz, Pedrotti, Busetta, Acar & Ritter, 2009).

2.1 VBS2™
VBS2 is the leading military 3D training environment. Developed by Bohemia Interactive, VBS2 offers both virtual and constructive interfaces onto high-fidelity worlds of synthetic human players, using motion-capture animations to model their movements realistically and efficiently. VBS2 is currently in use by military organisations worldwide, including the USMC, US Army, Australian Defence Forces (ADF) and UK MoD, for mission rehearsal, tactical training and simulated combined arms exercises.

2.2 CoJACK™
In the past decade, the computer games industry has developed truly impressive photorealistic 3D virtual environments. These are in widespread use in homes all around the world and have also been adapted to military applications. In the military context, photorealistic environments (e.g. VBS2) are
used for training and mission rehearsal, whereas tactics development, course of action analysis and hardware acquisition studies tend to be performed on more traditional simulation environments. Typically, whether for games or military applications, simulated human entities are an essential part of the scenario. However, in contrast to hardware such as aircraft, tanks and weapons systems, even highly-trained humans can vary significantly in their response to a given situation. Although the inherent variability of humans has been widely recognised, virtual environments have tended to neglect this phenomenon because it is very difficult to model the depth and breadth of human behaviour. CoJACK was developed to address this shortcoming.

CoJACK is a cognitive architecture, founded on the BDI (Beliefs, Desires, Intentions) paradigm. Virtual actors are defined in terms of their beliefs (what they know), their desires (what goals they are motivated to achieve) and their intentions (the goals they are currently committed to). Based on the work of (Bratman, 1987), the BDI paradigm is an intuitive abstraction that facilitates the implementation of virtual actors. Of particular interest is CoJACK’s high-level graphical plan representation (figure 1), which facilitates the definition of human reasoning and action, and also allows Subject Matter Experts to interact with the cognitive models.

CoJACK simulates the structural properties of the human cognitive system, that is, the information processing mechanisms that are fixed across tasks. For those unfamiliar with the concept of a cognitive architecture, the following analogy should help. Consider figure 2, showing a ball bearing that is released from the top of a ramp. The behaviour of interest is the time taken to reach the bottom. The length of the ramp and the maximum lean angle (90°) is fixed. These are structural properties of the system and are analogous to the invariant properties of human cognition. The angle of lean can vary within the 90° range and the surface can be covered with different coatings (e.g. glass vs. deep pile carpet). These are parameters of the system, and are analogous to the internal cognitive parameters of the virtual actor. By varying these parameters, we obtain different behaviour (the ball bearing takes differing amounts of time to reach the bottom). Taken together, these factors can be used to predict the behaviour of the ball bearing.

Like ACT-R (Anderson, 2007), CoJACK moderates the execution of its procedures by including sub-symbolic operations that define the time taken to reason, how memory is accessed and the errors that can occur. These sub-symbolic operations are controlled by parameters that perform a similar function to the parameters in the ball bearing example. By varying these parameters, the virtual actor’s behaviour changes, for example by making an error or taking too long to retrieve a belief from memory.
The simulation operator can manipulate these parameters at runtime, but more interesting than this is the ability to have the parameters controlled by moderators. These moderators can represent emotional and physiological factors such as fear and fatigue, influencing the course of action taken by the virtual actor (see Evertsz et al. (2009), section 4.6 for more detail on how moderators influence the perceived utility of a course of action).

4. CIVILIAN TERRORISM SCENARIO

A scenario was specified as part of the development of the counterterrorism prototype. The objective of the scenario was to provide a focus for understanding and modelling crowd behaviour in congested areas and to provide a case study where terrorist behaviour could be distinguished from that of civilians. Areas like shopping centres, stadiums and transport hubs are potential targets and so this scenario was set in a pedestrian precinct (figure 3).

![Figure 3: London pedestrian precinct](image)

The trainee controls the policeman. The scenario was populated with a mix of about 50 virtual actors:

- Workers (roughly 40%)
- Tourists (roughly 20%)
- Shoppers (roughly 40%)
- Children (roughly 1%)
- Mother/father (roughly 1%)
- Reconnaissance Terrorist (1 instance)
- Police dog (1 instance)

Each individual virtual actor is assigned a “personality” assembled from a palette of capabilities. A personality consists of the following main components:

- Beliefs – what the virtual actor knows (e.g. location of particular shops, where the bus stop is, who its family group is, etc.).
- Procedures – what the agent knows how to do (e.g. how to navigate to a given location, how to respond aggressively when obstructed, etc.).
- Goals – what the agent wants to achieve (e.g. go shopping, get to work on time, etc.).
- Temperament – the affective style of the virtual actor (e.g. naturally anxious). See (Davidson et al., 2003) page xiii.
- Mood – a diffuse affective state of relatively long duration, (e.g. being in a state of fear due to some event).
- Attitudes – predispositions to objects or situations (e.g. not liking police dogs, which types of event are considered dangerous, etc.).
- Cognitive Effectiveness – the baseline state of the underlying cognitive parameters, affecting capabilities such as ability to recall facts, hold intermediate results in working memory, and stay focused on a goal.

The virtual actor’s personality determines how it responds to events in the scenario and what actions it chooses to perform. For example, if a businessman’s goal is to get to work on time and he knows the way to work (beliefs) and knows how to navigate to that location (procedures), he will walk to work. However, if he is late (belief), he will run (procedures). If he is obstructed by a policeman, and is ill tempered (temperament) or in a bad mood, he may respond aggressively (procedures). If he is fatigued he may forget to look both ways when crossing a road (cognitive effectiveness).

The virtual actors can be assigned personalities manually, or automatically according to a predetermined probability distribution. Once this is done, the scenario
can be started and it plays out in a way that is
influenced by the interaction between the
trainee’s actions and the personalities of the
virtual actors. The behaviour of the virtual
actors emerges from these interactions. It is
not scripted to produce a particular outcome.
In this way, the trainee gets to experience a
more variable interaction than is possible with
typical scripted approaches to implementing
human behaviour.

4.1 Biasing the training effects
If desired, the trainer can bias the scenario in a
particular direction to focus the development
of the trainee’s skill set. An example would be
to turn up the baseline level of fear for all
civilians. This could be used, for example, to
simulate a situation in which there is a
heightened terror alert and the general
populous is a bit jumpy.
Moderators on individuals can be tweaked at
time to trigger behaviour that is apposite to
the desired training outcome. For example,
lowering the fearful level of the reconnaissance
terrorist so that he is less affected by the
trainee’s actions.

4.2 Moderators
For this scenario, fear, anxiety and anger were
implemented in a reservoir-based manner
similar to that used for the suicide bomber
scenario described in (Evertsz et al., 2009).

5. OVERVIEW OF THE PROTOTYPE
The scenario was developed for the
INSTINCT Technology Demonstrator Event,
held in London in November 2009. Although
VBS2 is targeted at military applications, we
selected it for the demonstration because we
had already integrated it with CoJACK, as
described in (Evertsz et al., 2009). VBS2
already has a broad collection of civilian
actors, so it was fairly straightforward to build
the scenario.
The behavioural range of the virtual actors is
too broad to cover within the confines of this
paper. To convey a flavour for the breadth of
behaviour, some of the virtual actor
capabilities are listed below:
• The dog follows the policeman, keeping in range because it becomes
anxious if its handler gets too far away. It barks when fearful, which can
be triggered by events such as a human gesticulating aggressively.
• The businessman becomes angry and
gesticulates aggressively if he is late for work and is hindered by the
policeman.
• The tourists wander around looking at
buildings. Their behaviour is hard to
distinguish from that of the
reconnaissance terrorist. However,
they react with less fear if confronted
by the policeman.
• Children follow their parents and will
play with the dog if it is nearby. If the
dog starts barking, the children become
frightened and run to their parents. The
announcement of a terror alert over the
public address system has no effect on
the children.
• The reconnaissance terrorist’s level of
fear increases when he sees the
policeman looking at him. His goal is
to find an area with a large number of
people and a place to conceal an
Improvised Explosive Device. If his
fear gets high enough, he will leave the
area.

6. DISCUSSION
Although the prototype successfully
demonstrated the viability of the approach,
there is room for improvement. The
outcome of the interplay of cognition and
emotion was convincing and led to subtly
different behaviour in each run of the scenario.
However, generating realistic behaviour in a
civilian context is challenging in a way that
typical military environments are not.: the
critical role of non-verbal communication.
Non-verbal signals are key to interpreting the
affective state of others, particularly if the
subject is attempting to conceal his intentions.
Non-verbal cues to intent can be picked up
from the subject’s facial expression, posture,
clothing (e.g. wearing a hood), direction of gaze, avoidance of eye contact with uniformed personnel and movement patterns (Argyle, 1988).

During face-to-face interaction, people pick up on each other’s non-verbal cues and use these cues to interpret a person’s verbalizations (Strongman, 1996). Facial expressions are considered to be the primary medium for interpreting emotional state (Knapp & Hall, 1978). Unfortunately, current virtual environments have a very limited repertoire of facial expressions available to virtual actors. But there is hope. Facial expression in computer-based avatars has already been studied (Fabri, Moore & Hobbs, 2002). In all but one emotional category, users were able to interpret the emotional significance of the avatar’s expressions. A widely accepted categorization groups facial expressions into six categories: Surprise, Anger, Fear, Happiness, Disgust/Contempt, and Sadness (Ekman, Friesen & Ellsworth, 1972). Indeed these categories are considered to be innate and therefore culturally universal. They correspond to distinct physiognomic arousal states (Argyle & Foss, 1994).

With the addition of an ability to render body language cues, the training system prototyped here will have much wider applicability. Although there is room for improvement in the graphical detail of current computer games, they are hindered more by the lack of realism in their virtual actors. This is likely to become an area of differentiation in the games industry, and improvements in these areas will inevitably flow on to more “serious games”, such as the counterterrorism application outlined in this paper.

ACKNOWLEDGEMENTS

We thank Bohemia Interactive for their support in creating the London street scene and UK-specific characters. We are also grateful to Detica for their support during the development of the demonstration.

REFERENCES