CINEFORUM
ISSN: 0009-7039
Vol. 66. No. S9, 2026 (Special Issue S9)

From Rule-Based Scripts to Emergent Intelligence: Analyzing Autonomous Agent Architecture in "Free Guy

Mitkumar M Thakkar[1], Shreya Unmeshkumar Patel[2], Dr. Himanshu Parmar[3]

[1]Assistant Professor KN University, Gota, Ahmedabad
Email ID:- meet.work18@gmail.com
[2]Assistant Professor Shri Alpesh N. Patel Post Graduate Institute of Science & Research,Sardar Patel University
Email ID:- patelshreya101418@gmail.com
[3]Assistant Professor KN University, Gota, Ahmedabad
Email ID:- himanshu.rock.parmar@gmail.com

ABSTRACT
This article explores the representation of the evolution of video game NPCs from scripted simulations to intelligent emergent behavior by the film "Free Guy". We present the architecture of the autonomous agents found in the film by using narrative analysis and frame semantics with established concepts from AI research. Results show that the starter NPCs, including the protagonist, initially employ an extremely rule-based, scripted animation style (similar to simple reflex agents), and move toward learning when animated further in the story. We define terms like behavior trees, finite-state machines (FSMs) and learning agents to clarify this change, and compare the movie's architectural assumptions to real-world game AI research. We find that the movie represents an evolution of NPC autonomy from set in advance, scripted behaviors to a notion of learning and memory capacity. We conclude that "Free Guy" dramatizes NPC autonomy, that parallels research in AI research and agent architecture design. Possible directions for future work include further exploration of hybrid architectures, and the ethical implications of an autonomous NPC actor.
Keywords: Autonomous agents, Non-player character (NPC), Agent architecture, Rule-based system, Behavior tree, Finite-state machine (FSM), Emergent intelligence, Memory and learning Reinforcement learning, Hybrid AI

1.INTRODUCTION

Suppose you're playing an open-world game: everywhere, hundreds of NPCs have a set schedule, taking the same paths, speaking the same lines, performing the same actions, and repeating ad infinitum. In most video games, this set of background characters is heavily scripted, controlled by rule-based scripting or behavior trees. So you might think you'll never see a true AI in a game, and in some ways it's true. The film "Free Guy" upends the standard model by bringing an NPC in the game world to life: "what if one NPC got self-aware?" To play that role, "Free Guy" cast the first AI "alive" inside a game.

In this paper, we explore the film "Free Guy" through the system of AI agent architecture. How might the behaviors of Free Guy (the NPC) relate to the range of known architecture, from simple reactive reflex rules to complex emergent intelligence?

Our objectives in this paper are to

(1) set up the research problem of NPC architecture and autonomy in a fictitious universe,
(2) briefly review AI theories relevant to the film (behavior trees or reinforcement learning, for example),
(3) analyze examples from the film for clues as to the architecture of the NPC's brain,
(4) interpret these clues,
(5) draw conclusions about implications for game design or AI theory.

Why does this matter? Because "Free Guy" raises many of the issues faced by game designers and AI researchers, for example: how flexible or responsive can I make an NPC's code? For the rest of this paper, after the introductory paragraph, we review related work on game AI and agent design, describe our qualitative visual analysis approach, report on the results of analysis, then discuss and conclude.

2.LITERATURE REVIEW

Game AI has historically been implemented as "rule-based architectures". Most NPCs are implemented as decision trees, finite-state machines (FSMs) or hierarchy behavior trees. For instance, a game enemy could be implemented with a state machine: "patrol -> chase -> attack -> flee" given simple trigger. As IBM, "simple reflex agents" are the simplest type of AI: they always look at the current input state and return the correct output, with

OPEN ACCESS © CINEFORUM

no memory. These agents tend to be very predictable: they do the same thing every time, which is suitable for NPCs in early game levels who can be learned easily by players.
But the research community has been trying to achieve "more sophisticated behaviors" for a long time. Some games build in memory into their NPCs, such as "Shadow of Mordor", which employed the "Nemesis System": enemies had memories of battles and animosities towards the player. From a research perspective, we now have a kind of limited memory, and that means "model-based" or "learning agents". Xu "et al". Compile a taxonomy of agent types from the AI community: from reflex to learning agent; the latter changing their rules based on experience, via methods like neural networks or reinforcement learning.
The phrase "emergent intelligence" frequently is used to describe complex behavior that spontaneously arises from the interactions of seemingly simple rules. In AI terms, connectionist systems (neural networks) are often defined by the emergent pattern from the interaction of many simple units. Multi-agent studies the emergence of coordinated behavior between agents to produce unanticipated result. In gaming, emergent gameplay occurs when interactions between simple NPC AIs result in unexpected situations. There is no academic definition, but the general AI explanation in AI literature suggests that by introducing memory, planning, or learning, behavior that "feels" like intelligence can be produced.
For the specific "Free Guy" project, we only have a few of its own analyses. Science Focus reports that modern day self-aware NPCs are "basically impossible". Powley (Falmouth University) points out "creativity and self-awareness are decades away", partly because we have yet to understand human consciousness. Yet, other game developers push toward a more 'intelligent' NPC: UbiSoft's "Watch Dogs: Legion" applies a "Census" system, with NPCs' biographical details and connections to others. Yet, Powley warns that any NPC that becomes too unpredictable "breaks down game design". A Syfy Wire article explains that "Free Guy" actor asks us to see Guy's awareness as "not directly programmed", but rather as "a consequence of prior conditions" and "a reflection of [his] game experiences." An analysis of the film from the "Oxford Student" quotes the director that NPC is "the first known true artificial intelligence alive within the game", and also criticizes that NPCs never escape into the larger world (some stereotypes remain even after they are free).
Based on this literature we derive a gap: there is no formal model linking Free Guy's story to concrete agent architectures. Our hypothesis is: 'every NPC behaviour scene can be mapped to recognised agent classes' (e.g. simple reflex, model-based, learning). We compare the film's representation to actual AI technology (e.g. ML, hybrid systems).

3.METHODOLOGY

We performed a qualitative analysis of the film content combined with AI theory. We identified scene points in "Free Guy" that are examples of NPC actions and transitions, then segmented the film script and video clips and categorized each of Guy's actions (e.g. morning login, see odd anomaly, rebel leader again) and each deviation from the baseline as "data points" and an example of the algorithm to be discovered. We then mapped these actions against categories of agents found in AI literature, e.g. Was Guy's behavior in the first encounter a simple reflex agent and behavior tree? Is his behavior in the last part evidence that he has memory and planning capabilities? We relied on found articles, expert commentary & AI theory books for definitions and comparisons e.g. IBM's article for reflex agents, text books for agent architectures.
This is not about drawing any numbers as our "results" are more descriptive. For clarity, we break down the different behaviors into tables and figures. We also support our findings with secondary sources; for example, all commentary on other similar NPCs (Watch Dogs, Shadow of Mordor) supports our explanations. Although it is not a user study, the core rationale is sound: events are mapped onto genuine (although fake) ideas of AI. We are aware of shortcomings: as a satire and entertainment product, the film is not intended to be a case study in technology. But our aim is to interpret a plausible architecture model from the story.

4.RESULTS

Our analysis reveals a distinct “transition from rigid scripting to emergent behavior” in “Free Guy”. We organize the findings by observing “phases” of NPC behavior and inferring the controlling mechanism:

OPEN ACCESS © CINEFORUM

Phase	Behavior Observed (Example)	Inferred Agent Mechanism
Baseline Routine	Guy's daily loop: wakes up, greets others, orders coffee.	Pre-defined rule/behavior tree scripts (reflex agent). Actions repeat identically each day.
Encounter with Change	Guy is shot in the face by a robber but resets and repeats.	Built-in reset (finite-state logic). Still purely scripted.
External Stimulus	Guy hears talking (Molotov Girl's voice), sees new object (source code).	Condition-triggered branching in script (e.g. "if object present then investigate"). This adds a simple decision layer, not true learning.
Learning from Experience	Guy saves Buddy, causes player confusion, hears laughter at a joke.	Limited-memory agent: he records events (e.g. player laughter) and adjusts (repeats a joke he "learned" should be funny). This resembles a model-based or learning agent.
Emergent Awareness	Guy questions reality, breaks routine, creative (writes messages, recruits others).	Learning agent with goals: Guy exhibits self-directed behavior (seeking Molotov Girl and the code), not explicitly scripted. This is akin to an adaptive agent developing new goals.
Collective NPC Action	Other NPCs rally behind Guy (file out of bank, chase players, celebrate at end).	Emergent multi-agent behavior: NPCs follow Guy's leadership. Could be seen as a form of influence network or multi-agent coordination (beyond any single-script programming).

These labels are supported for a number of reasons, though. In "Baseline Routine", things Guy does each loop is identical- he even tells us that it remains constant. This is an archetypal "simple reflex" or scripted FSM: "taking existing conditions and selecting from [pre-programmed] actions". Science Focus also says that many games feature NPCs limited by either pre-set choices or algorithms, which fits his initial nature. The IBM document also identifies simple reflex agents as following condition-action rules with no memory.

When "Encounter with Change" takes place (e.g. when he is shot), Guy reverts back to his starting location. This is done by game scripting (also a common trope: looping if interrupted). There is initially no sign of the event having learned anything.

The subsequent phase, "External Stimulus", begins when something new enters the game (Molotov Girl). Initially, Guy's code has a branch: if he hears Molotov Girl (the player character), he stops. This suggests a scripted conditional. Still, it also opens the door for "Learning from Experience". For example, in the scene where Guy learns a dirty joke, he is shown actively recording player reactions, as a film critic notes. Guy repeats a homophobic joke he overhears, thinking it must be funny because players responded with laughter. This is essentially unsupervised learning: Guy's AI adjusts its responses based on data (player laughter). Ruth Stalker-Firth's blog notes that Guy has done some unsupervised learning, and Milly later corrects him. This implies his agent has at least a feedback loop to update knowledge, moving beyond a pure reflex agent to a learning agent that adapts based on experience.

Finally, in "Emergent Awareness", Guy begins to act on his own. Rather than simply respond to commands, Guy begins to take his own actions (recovering the source code, protecting Molly). Based on the film, "the emergence of Guy as a fully-fledged awareness wasn't wholly directed. Instead, he blossomed from a set of prior conditions". This is precisely the definition of emergent behavior. This example would fit a hybrid agent with planning and memories for his architecture. He could be said to be developing a mental model of the world and selecting activities because of this. We can infer from this that what we have here is a hybrid reactive-deliberative architecture: his reactive aspect, by definition of our table, is low (he doesn't come up with clever plans), whereas his deliberative aspect is high (he wants to rescue Molly). Such a hybrid architecture of fast but naive reactions with slower and cleverer planning is also well known in AI.

Interesting emergent result from the final collective action (the NPCs siding with Guy). None of those NPCs were wired up for that role more exactly, but they do anyway, because they are following Guy. From an AI perspective that's a multi-agent emergent effect, one agent (Guy) going off-road triggers a whole chain of other agents (NPCs) doing the same. Basically an influence or contagion effect in an agent network. The film doesn't state how, but from a modelling perspective this suggests the implementation must have been that either Guy's independent actions changed how all NPCs experienced the environment, or some code in him could let his deviant actions

OPEN ACCESS © CINEFORUM

send a signal to the others (see next point). The review in the Oxford Student hinted that the NPCs remained aware of the system that was imprisoning them, which suggests he was more likely destroying their system (servers) which would kill them all. So again that ending is probably overselling this level of autonomy but it's consistent with a story-based agent network effect where a leader influences a following.
Bottom line, our findings outline a convincing spectrum mapping of the transformation of Free Guy's NPCs from a scripted reflex agent to a goal-oriented learning agent. Here are the main takeaway points:
"Strict rule-based conduct" dominates early (consistent with industry practice).
"Limited learning and memory" shows up in the middle of the story (as in "limited memory agents" from AI theory).
"Goal-driven, adaptive behavior" is coined as a later characteristic (like "learning agents" or hybrid systems).
The film because of this illustrates a shift from static architecture to an adaptive one.

5.DISCUSSION

While our reading of the Free Guy agent design differs from and complements previous understanding, the process of reconciling our reading with a literature we are less familiar with reveals some interesting parallels. As previous reviewers mention, "Free Guy" heightens the fantasy of NPC free will beyond what AI is capable of today. Our view concurred that a truly conscious NPC is "virtually zero" in real games, but the film was still informative: it was a vision of a "learning agent" in action, including world-model construction and goal setting.
Compare to real game production: today, NPCs are sometimes learning via unsupervised techniques (e.g. Legion's Census) but those are still mostly designer-driven systems. "Free Guy" proposes NPCs could learn unsupervised from what players do a dangerous gamble in the real world, as Ruth Stalker-Firth notes (bias on Twitter bots, etc.). The movie brings up that risk in a very quick offhand line: Milly "corrects" Guy's learning of a bad joke. In ML, you'd call that a human-in-the-loop that corrects an ML model. So you know the filmmakers know at least a little about machine learning pitfalls.
A limitation (or requirement) of our study (and the film) is that "Free Guy" does not have an explicit algorithm. We have to infer everything from the behaviour. For example, our descriptor of what happens to Guy's later phase is that it is a "learning agent" but the film does not show any training or reward signals just that the NPC "learns" through watching other NPCs. No doubt one expects reinforcement learning or neural nets in the real world both of which require vast amounts of training (AlphaStar is an example of learning in games). In the film, all this complexity is skipped over so it's a big oversimplification. The other major omission is the interaction between the player and the NPC playing the game is data for the NPC who would normally consider the players as invaders and not teachers. The contrivance of having Molly's voice speak to Guy is just a cheat to get him to learn.
Our results though also emphasize a dilemma described in the Science Focus story, too unpredictable, turns out to be the enemy of good game design. "Free Guy" allows NPCs to be a little unpredictable towards the end (characters break free of their roles) but perhaps, real-life games wouldn't let this happen (as Powley suggests). During play, part of the fun comes from discovering what non-player characters' systems are - if any NPC can do everything, it becomes less fun. This dilemma highlights that game designers may not always be on board with emergent intelligence.
We've identified some future research directions in our analysis. One would be the development of formal agent models inspired by the film; for instance, can we design a Unity simulation in which a basic NPC learns a reinforcement learning policy as triggers are hit? The "LLM-guided RL" paper alludes to hybrid models (language models guiding the policy) that might lead to more capable NPCs. Another is the question of ethics; if an NPC like Guy became sentient, what rights or duties would they hold? The film addresses this (deactivated server kills them all) but future consideration of virtual sentience might take cues from multi-agent AI ethics.
In the end though there are two lessons from this paper. From a scholarly point of view, it taps an oft-used pop culture example (AI agents) to provide concrete experiments on these otherwise-fluffy topics. From a practical point of view, it demonstrates just how large (or perhaps small) the steps are from the world of "Free Guy". Players may "feel" that their NPCs have agency, but reading the architectures, you'll realize how far away we are from that world.

6.CONCLUSION

In this paper we combined film analysis and AI theory to conceptualize a pathway from rule-based NPCs to emergent intelligence, using "Free Guy" as a hypothetical illustration. We established that at first, the NPC has a traditional game AI (scripts/behavior trees), then transitions into a learning autonomous agent (memory/goal-seeking), with the emergence of a self-awareness feature presented as a surprise. Ultimately, "Free Guy" may

OPEN ACCESS © CINEFORUM

have introduced a scenario of autonomous agent architecture conceptual development, which was greater than our present-day capabilities, but theoretically the trajectory was there.

The main findings indicate that emergent intelligence, although fictional in this context, corresponds with the concept of hybrid agent architectures, like combining a reactive loop with a deliberative planner, and learning agents that adapt based on player interactions. The importance is in connecting theoretical ideas with popular imagination, demonstrating to readers that even in a comedic setting, genuine AI concepts are present. The contributions include clarifying relationships between various AI techniques by linking narrative elements to formal agent types, and highlighting practical difficulties that would arise if such a system were to be implemented.

For further research, one might operationalize the NPCs of Free Guy in simulation. For example, take an NPC and make them a behavior-tree agent in a virtual world; then gather information from them (like repetition of game chat conversations) and feed it into the learning module, calculating the time or trigger requirements. Or, analyze other game-based movies and fiction to see if there are any coherent themes in AI portrayals. Or, an interdisciplinary project might consider the social and ethical issues of Free Guy: how would gaming change if we ever successfully created sentient NPCs? These are only some directions to extend the ideas in this film, and concretely explore the ideas it puts forth

REFERENCES

1. Ward, Cassidy. “How Close Are We to Free Guy’s Digital Awareness? The Science Behind the Fiction.” “Syfy”, 25 Aug. 2021, syfy.com/syfy-wire/free-guy-ai-digital-awareness-real-technology.
2. Kelly, Stephen. “Free Guy: Why Video Game AI Will Always Be Stupid.” “BBC Science Focus”, 8 Oct.
3. 2021, sciencefocus.com/future-technology/free-guy-artificial-intelligence-npc.
4. Evans, Amy. “Analysing ‘Free Guy’.” “The Oxford Student”, 29 Nov. 2021,
5. theoxfordstudent.com/2021/11/29/analysing-free-guy/.
6. Stalker-Firth, Ruth. “Free Guy: Our Favourite AI.” “RuthStalkerFirth.com”, 15 Dec. 2021,
7. ruthstalkerfirth.com/blog/2021/12/15/free-guy-our-favourite-ai/.
8. Xu, Bin, et al. “AI Agents: Evolution, Architecture, and Real-World Applications.” “arXiv”, 12 Mar. 2025,
9. arxiv.org/abs/2503.12687.
10. Nair, Hrithika Deepu, and Kayvan Karim. “LLM-Guided Reinforcement Learning for Adaptive NPC
11. Behavior in Multi-Agent Combat Games.” “arXiv”, 27 Aug. 2026, arxiv.org/abs/2609.02931.
12. Stryker, Cole. “Types of AI Agents.” “IBM Think”, IBM Corporation, n.d., ibm.com/think/topics/ai-agent-types.
13. Wells, Jackson. “AI Agent Architecture Patterns, Challenges, and Governance for Production Systems.”
14. “Galileo.ai Blog”, 7 Apr. 2026, galileo.ai/blog/ai-agent-architecture

OPEN ACCESS © CINEFORUM