

MASTER THESIS PROJECT PROPOSAL

# **Autonomous Creation of Non-Playable Characters using Reinforcement Learning and Knowledge Graphs**

Antonio Mangoni di S. Stefano  
gusmangoan@student.gu.se

Supervisor at CSE:  
Kıvanç Tatar  
tatar@chalmers.se

Participating in the master thesis program at AI Sweden

February 12, 2024

# 1 Background / Introduction

Intro Context (1/3)

## 2 Aim

Context (1/3) Goals and Challenges

## 3 Problem Formulation

Problem Context (1/3)

## 4 Methodology

Approach

## 5 Risk Analysis and Ethical Considerations

stakeholder analysis

## 6 Time Plan

## 7 Introduction

This thesis investigates the application of Reinforcement Learning (RL) to enhance Non-Playable Characters (NPCs) in video games beyond conventional script-based methods. The focus is on developing NPCs that interact dynamically with the game environment while maintaining coherent behaviour, aiming to create more immersive and realistic gaming experiences.

Motivated by the limitations of pre-scripted NPC behaviours in current systemic game design [1], this research seeks to develop NPCs with responsive and context-aware actions that align with player-driven scenarios.

A key aspect of this thesis is the use of a knowledge graph (KG), a graph-structured database where nodes represent entities and their interrelationships as edges. This network will serve as a grounding method for the RL model, complementing the neural networks' representation of knowledge with a KG's symbolic structure.

The research would explore the dynamic creation of a KG integrated into the model architecture that would serve as a symbolic representation of data to ground the connectivist representation of data of the deep reinforcement learning artificial neural network.

Once a KG is established, it can serve as prior knowledge in subsequent training sessions, streamlining the development of new NPCs. The efficiency would be further enhanced by tailoring the KG to provide agents with only relevant, localised information. This would not only speed up the training of new NPCs but also aid in quickly adapting NPC behaviour in response to game updates.

My preliminary research question is: *How can dynamically created Knowledge Graphs be effectively integrated as symbolic grounding in model-based Reinforcement Learning for Non-Playable Characters in a complex game world environment, and what are the challenges and implications of this integration for the realism and adaptability of NPC behaviour?*

Additionally, the KG's network mapping the game world entities and their interactions offers insights into the game's dynamics, aiding in identifying and understanding possible emergent behaviours.

This work has broader implications beyond gaming, particularly in fields that require AI agent behaviours in complex adaptive systems, like simulations and robotics. Employing a KG can expedite the learning process, reduce undesirable actions, and enhance the reliability of AI agents, including addressing hallucination issues in large language models (LLMs).

To address challenges in NPC technology, this thesis proposes integrating established model architectures. Convolutional Neural Networks (CNNs) will process spatial data, Transformer-based networks will handle temporal analysis, and a world model will provide dynamic environmental understanding. These components will be linked to a KG, which translates sensory data and interactions into symbolic knowledge, enhancing NPCs’ environmental perception and contextual awareness.

## 8 Context

This thesis is situated within the broader scope of applying artificial intelligence in video game development, with a specific emphasis on the utilization of reinforcement learning. The foundational framework for this exploration is established by two studies: ‘Crafter’ ”Benchmarking the Spectrum of Agent Capabilities” [3], and ”A Survey of Deep Reinforcement Learning in Video Games” [5].

Crafter is an open-world survival game that serves as a benchmark for evaluating AI agents. The complexity of the game demands a range of agent capabilities, including generalization, exploration, and long-term strategic planning. These requirements underscore the substantial challenges present in the field of AI, particularly in the context of reinforcement learning applications in virtual environments.

Conversely, the survey provides a comprehensive overview of the evolution and current state of deep reinforcement learning within the realm of video games. This survey is the foundational knowledge that will be used during the project. This thesis endeavours to build upon this established foundation by delving into the development of NPCs capable of adaptation and response in complex game environments.

Other papers that are being reviewed: ”Towards Robust Knowledge Graph Embedding via Multi-Task Reinforcement Learning” [7] is being considered as a framework to use KGs effectively; ”Graph Embedding Priors for Multi-task Deep Reinforcement Learning” [4] is also being considered as using the KG for prior knowledge would give the agent a ’common sense’ starting point, helping accelerate the start of the training lesson.

## 9 Problem

The proposed methodology involves creating an adaptable KG that evolves as the agent interacts with various environments. The initial development will be in a custom, less challenging environment. Following this, it will be tested in the Crafter environment to assess its robustness. The model may need adjustments and will have to be either fine-tuned or retrained for the new setting.

To make this methodology the KG needs to be constructed as the agent explores the environment, or at least have the possibility to add new nodes to the graph as new connections are discovered. This will allow it to be used in different environments, to test this we could start the project on our custom environment and then test it on the Crafter one. In this way we could both develop the methodology in a learning environment where the rewards are not so sparse but when it is working test it against state-of-the-art benchmarks.

One of the challenges in this thesis is the optimization of NPC learning algorithms, they need to enable realistic NPC behaviour within dynamic game environments, yet be efficient enough to avoid inference delays that hinder real-time interaction. Balancing complexity and computational efficiency is crucial in maintaining the game’s fluidity and responsiveness, ensuring player engagement is not disrupted by delays in NPC reactions.

In addressing these challenges, this thesis seeks to bridge the gap between the potential of dynamically evolving game worlds and the current capabilities of NPCs. The goal is to develop NPCs that are not only contextually aware and capable of adapting to complex environments but also efficient in their computational demands.

## 10 Goals and Challenges

The primary goal of this thesis is to develop an NPC using reinforcement learning that is capable of dynamic and coherent behaviour. This NPC should be able to learn and then use the acquired

knowledge to adapt in real-time to both the player’s actions and the changing environment in a video game.

**KG Creation:** Create a KG which can be used as a symbolic representation of knowledge for the artificial neural network.

**KG integration:** The KG needs to be integrated into the model architecture, being fed data by the CNN and interfacing with the world model and the transformer to help with grounding.

**RL Algorithms:** Using efficient learning algorithms that balance computational efficiency with effective learning.

**Real-Time Adaptation:** Ensuring the NPC can adapt its behaviour in real time without impacting the game’s performance.

**Behavioural Realism:** Ensuring the NPC can exhibit believable behaviours.

**Generalization:** Ensuring the NPC can generalize learned behaviours to new and unforeseen situations within the game.

## 11 Approach

### 11.1 Model Architecture

**Objective:** The objective is to integrate and apply neural network architectures, specifically Transformer-based networks and CNNs, within a reinforcement learning framework.

**Rationale:** Transformer-based networks are well-suited for capturing the sequential and temporal aspects of NPC behaviour, offering an understanding of the progression and consequences of various actions within the game.

On the other hand, CNNs are ideal for processing spatial information, crucial for interpreting the game’s terrain and the distribution of entities. They excel in identifying patterns and features in grid-like data, making them apt for analyzing the game’s environment structured as an x-y-2 3D array.

**World Model Integration:** To address the challenge of NPCs having a limited view of the environment, a world model will be integrated between the CNN and the Transformer network.

This model will use the spatial data processed by the CNN to construct an internal representation of the broader environment, enabling NPCs to make informed decisions and predictions about areas outside their immediate perception. The world model thus acts as a bridge, enhancing the spatial data before it is analyzed for temporal patterns by the Transformer.

How to integrate the KG into this architecture will have to be studied further.

**Implementation:** Two distinct approaches for processing environmental data will be explored:

- Option 1 – Separate Inputs: Utilize separate CNN input channels for terrain and entities, allowing the network to learn distinct features from each layer.
- Option 2 – Combined Tuple Input: Represent each coordinate as a tuple (terrain type and entity information), enabling the CNN to learn direct interactions between terrain and entities.

In both options, after processing the spatial information, the output will be fed into the Transformer network for temporal analysis.

## 11.2 Simulation Environment

**Development:** The simulation environment is developed using Pygame in Python. This environment is procedurally generated using Perlin noise to ensure that each instance of the game world is unique, preventing the agents from simply memorizing patterns.

**Environment Features:** The game world is structured as an x-y-2 3D array, where the first layer represents different terrain types, and the second layer indicates entities present in each cell. The terrain types are one-hot encoded for rendering, which facilitates their interpretation by the neural network. This environment includes various terrain types such as plains, hills, and water bodies, each with unique attributes affecting NPC movement and actions.

Additionally, resources like wood, rock, and food are distributed throughout the map, presenting opportunities for NPCs to interact with the environment. These features are critical for developing NPCs that can navigate and utilize the game world effectively, replicating realistic behaviours seen in actual gaming scenarios.

**Testing and Adaptation:** The procedurally generated environment serves as a testing ground for NPCs to adapt their behaviours. Through reinforcement learning, NPCs are trained to respond to the dynamic conditions of the game world, such as changing resource availability and terrain challenges.

The environment allows for the evaluation of NPC behaviours under various conditions, ensuring they can adapt to new challenges and maintain robust performance. This adaptability is tested by altering environmental factors like resource distribution or introducing new entities, providing a comprehensive assessment of NPC versatility and the effectiveness of their learned behaviours.

## 11.3 Dynamic Knowledge Graph Integration

**KG Expansion:** This project will explore the possibility of agents contributing to the expansion and updating of the KG. This process, operating alongside the reward system, would not be directly linked to rewards but is considered essential for the agent’s learning and exploration.

**KG Utilization:** As part of its exploration, the agent might identify new entities, objects, and their interactions within the game environment. These discoveries could then be added to the KG, potentially enhancing the agent’s understanding of the game world.

**Potential Adaptive Exploration Strategy:** One aspect under consideration is adjusting the agent’s exploration strategy based on the development of the KG. For instance, as the KG becomes more comprehensive, the agent could reduce its exploration rate (epsilon value in epsilon-greedy exploration), focusing more on exploiting its existing knowledge.

**Potential Impact on Agent Behavior:** A dynamically updated KG could provide NPCs with a deeper understanding of the game environment. This, in turn, could improve their decision-making and adaptability, enabling them to more effectively respond to new challenges and scenarios in the game.

## 11.4 Reward System Design

**Formulation:** The reward system will be designed to enhance the adaptability and decision-making capabilities of NPCs. The system will focus on encouraging NPCs to interact effectively with the game environment and respond dynamically to player actions.

**Adaptive Learning:** NPCs will be incentivized to learn from their environment, with rewards structured to promote exploration and intelligent decision-making. This approach aims to develop NPCs that can understand and adapt to complex scenarios in the game.

**Impact on Learning and Adaptability:** The goal is to balance between achieving immediate objectives and developing strategies for long-term gameplay challenges. This will be achieved by using Feudal and Hierarchical RL [6].

## 11.5 Performance Optimization

**Optimization Strategies:** The "...Survey of Deep Reinforcement Learning in Video Games" paper will be used as a starting point for the decision of which RL algorithm to use.

**Efficiency Considerations:** Efficiency considerations in this project are centred around monitoring processing speed and resource utilization to ensure NPCs can infer effectively in real time within the game environment. This monitoring will involve measuring inference time and identifying any bottlenecks within the observation-inference-action pipeline.

## 11.6 Research Methodology for NPC Behaviour Analysis

### 11.6.1 Quantitative Methodologies

**Learning Curve Assessment:** In this project, learning curves will serve as a primary tool for monitoring the NPCs' skill development over time. [3] These curves will provide insights into the NPCs' learning efficiency, their adaptability to the dynamic game environment, and their interactive capabilities. This will also be used during the development phase and analysis of the model.

**Benchmarking Using the Crafter Framework:** To evaluate our model, we will use the benchmarking system provided by the Crafter environment. This will involve testing our NPCs within the Crafter setup and comparing our models' performance against its established benchmarks. [3]

### 11.6.2 Qualitative Methodologies

**Behavioural Realism Assessment:** The plan is to implement a modified version of the Turing test to gauge the behavioural realism of the NPCs. In this test, observers will be challenged to differentiate between the NPC and human players based on their in-game behaviour. [2]

**Comparative Evaluation:** These metrics used to analyse NPC behaviour will include reaction time, the complexity of decision-making processes, and their overall adaptability in different game scenarios. This will also be used during the development phase and analysis of the model.

## References

- [1] Systemic games. Accessed: 2023-12-02.
- [2] Sam Devlin, Raluca Georgescu, Ida Momennejad, Jaroslaw Rzepecki, Evelyn Zuniga, Gavin Costello, Guy Leroy, Ali Shaw, and Katja Hofmann. Navigation turing test (ntt): Learning to evaluate human-like navigation, 2021.
- [3] Danijar Hafner. Benchmarking the spectrum of agent capabilities. *ArXiv*, abs/2109.06780, 2021.
- [4] Neev Parikh, Zachary Horvitz, Naveen Srinivasan, Aansh Shah, and George Dimitri Konidaris. Graph embedding priors for multi-task deep reinforcement learning. 2020.
- [5] Kun Shao, Zhentao Tang, Yuanheng Zhu, Nannan Li, and Dongbin Zhao. A survey of deep reinforcement learning in video games. *ArXiv*, abs/1912.10944, 2019.
- [6] Alexander Sasha Vezhnevets, Simon Osindero, Tom Schaul, Nicolas Heess, Max Jaderberg, David Silver, and Koray Kavukcuoglu. Feudal networks for hierarchical reinforcement learning. In *International Conference on Machine Learning*, pages 3540–3549. PMLR, 2017.
- [7] Zhao Zhang, Fuzhen Zhuang, Hengshu Zhu, Chao Li, Hui Xiong, Qing He, and Yongjun Xu. Towards robust knowledge graph embedding via multi-task reinforcement learning. *IEEE Transactions on Knowledge and Data Engineering*, 2021.