Universitat Rovira i Virgili

Master's Thesis

Masters of Artificial Intelligence and Computer Security

Behavior approximation using fuzzy-genetic systems

Mark Safronov

February 15, 2025

UNIVERSITAT ROVIRA i VIRGILI

Contents

1 Introduction **1**

1 Introduction

We will define "behavior" as a classical Reinforcement Learning problem.

Assuming we have a character described as a set of numeric characteristics

$$\mathbf{x} \in \mathbb{Z}^n \tag{1}$$

we have a set of possible actions

$$A = \{a_1, a_2, \ldots, a_m\} \tag{2}$$

which collectively form a transfer function

$$f(\mathbf{x}, a) = \mathbf{x}' \tag{3}$$

Now, we have a goal state

$$\mathbf{x}^* \in \mathbb{Z}^n \tag{4}$$

and a planning horizon

$$T \in \mathbb{Z} \tag{5}$$

We want to get an ordered actions sequence of length T which will lead $\mathbf{x}$ to $\mathbf{x}^*$:

$$\mathbf{a} \in A^T, x_o = \mathbf{x} : \bigodot_{i=1}^{T} f(x, a_i) = \mathbf{x}^* \tag{6}$$

where $\bigodot$ is a fold operator. The transfer function f is assumed to be completely determined, and the whole process being non-stochastic.

Normally, there is a number of classical approaches to this problem, including dynamic programming, reinforcement learning and state space searches using graph theory. However, in this work we'll focus specifically on the cases which lead to combinatorial explosion for classical solutions, that is, when we have sufficiently large amount of characteristics, actions to choose from and most importantly, very large planning horizon:

$$n > 50 \tag{7}$$
$$m > 40 \tag{8}$$
$$T > 2000 \tag{9}$$

As an additional restriction, we'll state that it's not enough to reach the desired state $\mathbf{x}^*$ sometime before T, instead, it is a must that we perform T actions and only then evaluate the final state $\mathbf{x}'$ of the character.

With these restrictions in place, a need in an heuristic arises to perform efficient search in the state space, as its size becomes unrealistically large.

In this work we evaluate one possible approach which is defined as follows.

Let's assume that we can segment the set of possible actions to clusters with the following particularities:

1. actions in the same cluster lead to “similar” changes in the character state $\mathbf{x}$.
2. the cluster as a whole can be described symbolically

In this case we can synthesize a set of numeric characteristics which we'll call “inclinations”:

$$\mathbf{I} \in \mathbb{Z}^q \tag{10}$$
$$q << n \tag{11}$$

From this, we can define a set of fuzzy rules mapping the inclinations to action choices:

1. if an inclination I_i has a fuzzy value V_I,
2. then P_a, the priority of an action a, is a fuzzy set V_A.

After the defuzzification of all the inferred fuzzy values P_a we select an action with the highest priority.

In here, the selection of fuzzy rules is paramount. In general, it's possible to construct the set of fuzzy rules mapping every $\mathbf{I}$ to every action's priority, or even further, every $\mathbf{x}$ to every action's priority. However, it will lead to the same combinatorial explosion as stated in the beginning.

The clusterization of actions described above was introduced to combat that effect. The assumption which we explore among others in this work is that it's possible to write a coherent set of fuzzy rules which will be clustered around the clusters of actions, and each inclination will tend to map to its own cluster of actions.

Now, using such a fuzzy controller $\xi(I, \mathbf{x})$ we can construct the goal function:

$$g(I, \mathbf{x}) = \bigodot_{i=1}^{T} f(x, \xi(I, x_i)) \tag{12}$$

the above formula being subject to improvements in expressiveness, the main point of which being the fuzzy controller $\xi(I, x_i)$ selecting the action to perform on the step i according to the inclinations and (ideally) the current state x_i.

For the g parameter $\mathbf{x}$ is essentially a constant, and given (11), we effectively performed dimensionality reduction on the original problem.

We can find $\arg\max(g)$ now using an appropriate optimization method. For this work, because of a strong biosocial analogies a genetic algorithm was chosen, with the vector of inclinations $\mathbf{I}$ as a chromosome.

The hypothesis explored in this work is that the combination of assumptions described above constructs an heuristic which allows solving the problem more efficiently than performing the reinforcement learning directly.

Historically, this work started as an attempt to write a formal solver for the Japanese computer game *Princess Maker 2* (1993) by the studio Gainax.