UNIVERSITAT ROVIRA I VIRGILI

MASTER'S THESIS

MASTERS OF ARTIFICIAL INTELLIGENCE AND COMPUTER SECURITY

Fuzzy-Genetic Approach to High-Dimensional Deterministic Planning via Behavioral Inclinations

MARK SAFRONOV

July 14, 2025

UNIVERSITAT
ROVIRA i VIRGILI

Contents

1 Introduction

Sequential decision-making problems with long planning horizons and high-dimensional state spaces are ubiquitous in real-world applications, from resource allocation and strategic planning to autonomous systems control. These problems become particularly challenging when the system must execute a fixed number of actions before evaluation, precluding early termination or idle states.

Classical approaches to such problems include dynamic programming, reinforcement learning [3], and graph-based search methods. However, these techniques face significant computational barriers when dealing with large state spaces, extensive action sets, and extended planning horizons. The combinatorial explosion of possible action sequences renders exhaustive search infeasible, while the curse of dimensionality challenges traditional learning algorithms.

This thesis investigates a novel approach that combines fuzzy logic controllers with genetic algorithms to address high-dimensional deterministic planning problems. The key insight is to perform dimensionality reduction through the concept of "inclinations"—latent behavioral parameters that capture strategic preferences—and use fuzzy rules to map these inclinations to concrete actions. This transformation reduces the optimization problem from searching over exponentially large action sequences to optimizing a much smaller space of behavioral parameters.

The motivation for this work stems from character development planning in role-playing games, where players must make sequences of decisions over extended periods to achieve desired character attributes. However, the proposed methodology is general and applicable to any deterministic planning problem with similar characteristics.

2 Problem Statement

We consider a deterministic sequential decision-making problem where an agent must transform a system state through a fixed sequence of actions to achieve a desired target state.

2.1 Formal Problem Definition

Let $\mathbf{x} \in \mathbb{Z}^n$ represent the system state, characterized by n integer-valued attributes. The agent has access to a finite set of m actions, each inducing a known deterministic state transformation. The system evolves according to:

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, a_t) \tag{1}$$

where $f : \mathbb{Z}^n \times \{1, 2, \ldots, m\} \to \mathbb{Z}^n$ is a deterministic transition function, and a_t is the action taken at time t.

The agent's objective is to find a sequence of actions $\mathbf{a} = (a_1, a_2, \ldots, a_T)$ that transforms the initial state $\mathbf{x}_0$ to a final state $\mathbf{x}_T$ that maximizes some reward function $R(\mathbf{x}_T)$. The complete trajectory can be expressed as:

$$\mathbf{x}_T = \bigodot_{t=1}^{T} f(\mathbf{x}_{t-1}, a_t) \tag{2}$$

where $\bigodot$ represents the sequential composition of state transformations.

2.2 Problem Complexity

This work focuses specifically on instances that lead to combinatorial explosion for classical solution methods. We consider problems with the following characteristics:

$$n > 50 \quad \text{(high-dimensional state space)} \tag{3}$$
$$m > 20 \quad \text{(large action space)} \tag{4}$$
$$T > 2000 \quad \text{(extended planning horizon)} \tag{5}$$

A critical constraint is that the agent must execute exactly T actions before evaluation. Unlike many planning problems where reaching the goal state early is acceptable, our formulation requires the full trajectory to be completed. This restriction reflects real-world scenarios where time progression is irreversible and every time step demands an action.

The total number of possible action sequences is m^T, which grows exponentially with the planning horizon. For the problem instances considered, this results in search spaces that are computationally intractable for exhaustive enumeration or traditional dynamic programming approaches.

2.3 Challenges for Classical Methods

The combination of high dimensionality, large action spaces, and extended planning horizons creates several challenges:

1. **Combinatorial explosion**: The action sequence space m^T becomes prohibitively large for exhaustive search methods.

2. **Curse of dimensionality**: Traditional reinforcement learning approaches struggle with the high-dimensional state space, requiring either impractical amounts of training data or sophisticated function approximation.

3. **Credit assignment**: The long planning horizon and delayed reward evaluation make it difficult to assign credit to individual actions, hampering learning efficiency.

4. **No early termination**: The requirement to execute exactly T actions eliminates the possibility of early stopping strategies that could reduce computational complexity.

These challenges motivate the need for alternative approaches that can effectively navigate the high-dimensional planning space while maintaining computational tractability.

3 Methodology

Classical reinforcement learning methods become intractable in this domain due to the high dimensionality of the state space, large action set, and long planning horizon. Moreover, the inability to halt or take neutral actions further exacerbates the combinatorial explosion of the trajectory space.

To address these challenges, we propose a novel approach that combines fuzzy logic controllers with genetic algorithms through the concept of behavioral inclinations. Our methodology consists of three key components: dimensionality reduction through inclinations, fuzzy rule-based action selection, and genetic optimization.

3.1 Inclinations: Behavioral Parameter Space

The core insight of our approach is to perform dimensionality reduction by introducing the concept of "inclinations"—latent behavioral parameters that capture strategic preferences and decision-making tendencies.

We assume that the set of possible actions can be segmented into clusters with the following properties:

1. Actions in the same cluster lead to similar changes in the system state $\mathbf{x}$.

2. Each cluster can be described symbolically and corresponds to a coherent behavioral strategy.

Based on this clustering assumption, we synthesize a set of numeric characteristics called inclinations:

$$\mathbf{I} \in \mathbb{Z}^q \tag{6}$$

$$q \ll n \tag{7}$$

where q is significantly smaller than the original state dimension n, achieving substantial dimensionality reduction.

3.2   Fuzzy Controller Design

We model the behavior policy as a fuzzy controller [2] $\xi(\mathbf{I}, \mathbf{x})$ that maps the current state and inclinations to concrete actions. The fuzzy rule base follows the general form:

1. **If** an inclination I_i has a fuzzy value V_I,

2. **Then** P_a, the priority of action a, is assigned a fuzzy set V_A.

After fuzzy inference, we defuzzify all inferred priority values P_a and select the action with the highest priority.

The selection and design of fuzzy rules is a critical aspect of this approach. In this thesis, the fuzzy rule base is constructed manually, leveraging domain knowledge to define the mapping from inclinations to action priorities. The rules are designed to cluster around action groups, with each inclination tending to favor its corresponding action cluster.

While it is theoretically possible to define fuzzy rules that map every possible inclination vector $\mathbf{I}$ or state $\mathbf{x}$ to action priorities, such exhaustive rule sets would quickly become infeasible due to combinatorial growth. This reinforces the importance of dimensionality reduction and symbolic clustering in making the fuzzy-genetic approach tractable.

3.3   Genetic Optimization

Using the fuzzy controller $\xi(\mathbf{I}, \mathbf{x})$, we can construct the objective function:

$$g(\mathbf{I}, \mathbf{x}_0) = R\left(\bigodot_{i=1}^{T} f(\mathbf{x}_{i-1}, \xi(\mathbf{I}, \mathbf{x}_{i-1}))\right) \tag{8}$$

where $R(\cdot)$ is the reward function evaluating the final state after T steps, and $\mathbf{x}_0$ is the initial state.

Since the initial state $\mathbf{x}_0$ is fixed for a given problem instance, the optimization reduces to finding:

$$\mathbf{I}^* = \arg\max_{\mathbf{I}} g(\mathbf{I}, \mathbf{x}_0) \tag{9}$$

Given the relationship $q \ll n$ from equation (7), we have effectively performed dimensionality reduction on the original problem, transforming it from optimizing over m^T action sequences to optimizing over a much smaller space of inclination vectors.

We employ a genetic algorithm [1] to solve this optimization problem, treating the vector of inclinations $\mathbf{I}$ as a chromosome. The genetic algorithm is particularly well-suited for this task due to:

1. The discrete nature of the inclination space

2. The black-box nature of the objective function $g(\mathbf{I}, \mathbf{x}_0)$

3. The ability to handle multimodal fitness landscapes

4. Strong biological analogies with behavioral evolution

3.4 Algorithmic Framework

The complete algorithm proceeds as follows:

1. Initialize a population of inclination vectors $\{\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_{\mathrm{pop}}\}$

2. For each individual $\mathbf{I}_j$:

   (a) Simulate the trajectory using the fuzzy controller $\xi(\mathbf{I}_j, \mathbf{x}_t)$

   (b) Evaluate the fitness $g(\mathbf{I}_j, \mathbf{x}_0)$

3. Apply genetic operators (selection, crossover, mutation) to evolve the population

4. Repeat until convergence or maximum generations reached

5. Return the best inclination vector $\mathbf{I}^*$

3.5 Hypothesis

The central hypothesis of this work is that the proposed fuzzy-genetic approach can solve high-dimensional deterministic planning problems more efficiently than classical reinforcement learning methods when:

1. The action space can be meaningfully clustered into behaviorally coherent groups

2. The problem admits a compact representation in terms of behavioral inclinations ($q \ll n$)

3. The fuzzy rule base can capture the essential decision-making logic with a manageable number of rules

Specifically, we hypothesize that:

- **Computational efficiency**: The dimensionality reduction achieved through inclinations will result in faster convergence compared to RL methods that must learn over the full action sequence space.

- **Sample efficiency**: The structured search in inclination space will require fewer evaluations than the exploration required by RL algorithms.

- **Scalability**: The approach will maintain reasonable performance as the problem dimensions (n, m, T) increase, whereas classical RL methods will suffer from exponential scaling issues.

- **Interpretability**: The resulting solutions will be more interpretable than neural network policies, as they are expressed in terms of meaningful behavioral parameters.

These hypotheses will be tested through experimental evaluation on problems of increasing complexity, comparing the proposed approach against baseline reinforcement learning methods in terms of solution quality, computational time, and convergence properties.

4 Implementation

TBD

5   Experiments and Results

TBD

5.1   Trivial case

2 characteristics, 4 mutually exclusive actions, 3 steps.

This scenario represents a trivial case, with only 4^3 possible action sequences—a total of 64. The small state space allows for exhaustive enumeration and manual verification of results. This case serves to validate the correctness of the implementation and the fuzzy controller, as the system's behavior can be easily traced and analyzed by hand.

5.2   Base control case

4 characteristics, 12 actions, 100 steps.

This case is the base case, as it introduces enough complexity to test the proposed approach and at the same time compare it with classical approaches.

A decision tree of the size 12^{100} is already too large to be completely enumerated.

However, with 4 characteristics and 12 actions, the problem is still well within the range where Reinforcement Learning methods—especially those using function approximation—can be applied efficiently. The planning horizon of 100 steps is long enough to be non-trivial, but does not pose significant challenges for standard RL algorithms.

5.3   Large-scale case

The original problem which historically started this work is a problem with 50 numeric characteristics and 25 actions to choose from, with a planning horizon of 2500 steps.

This case is an attempt to directly solve this problem using the proposed fuzzy-genetic approach. It will be used as a benchmark for the proposed solution on a real-world problem.

6   Discussion

(to be filled based on the results of the experiments)

7   Conclusions and Future Work

Despite the context of the problem being a computer game, the problem itself is a general one, and the proposed approach can be applied to any high-dimensional deterministic planning problem. Which constitutes the core value of this work.

In the span of this work, only three distinct cases were explored, and a more thorough exploration of the parameter space is left for a dissertation-level research. Theoretically this approach can be applied without any changes to more intrinsically complex problems, such as those with stochastic transitions, non-deterministic actions, or even partially observable states. This flexibility makes the proposed method a promising candidate for a wide range of applications beyond the initial scope.

The main limitation of the current implementation is the manual construction of the fuzzy rule base. While this is sufficient for the cases explored, it does not scale well to larger problems or more complex action spaces. Future work should focus on automating the rule generation process, potentially using machine learning techniques to learn the fuzzy rules from data. This

would allow the approach to adapt to a wider variety of problems without requiring extensive domain knowledge.

References

[1]   Melanie Mitchell. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 1999. ISBN: 9780262631853.

[2]   S. Kumar Ray. *Soft Computing and Its Applications, Volume II*. Boca Raton, FL: CRC Press, 2014. ISBN: 978-1-4822-5793-9.

[3]   Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. 2nd ed. Available online. MIT Press, 2018. ISBN: 978-0-262-03924-6. URL: http://incompleteideas. net/book/the-book-2nd.html.