Modelling and Simulation for Infantry using Weapon Target Assignment

Tumelo R. A. Uoane
Landward Sciences
DPSS
18 May 2015
Outline

- Introduction
- Weapon Target Assignment (WTA)
- Related Work
- Problem Formulation
- Markov Decision Process (MDP)
- Reinforcement Learning (RL)
- Implementation
- Future Work
Introduction

- Military application areas.

- Developing a simulation tool.
- WTA departure point.
Weapon Target Assignment

The problem can be formulated in the following form:

\[
\text{minimize} \quad \sum_{j=1}^{n} V_j \left( \prod_{i=1}^{m} q_{ij}^{x_{ij}} \right) \tag{1}
\]

subject to

\[
\sum_{j=1}^{n} x_{ij} \leq W_i \quad \text{for all} \quad i = 1, 2, \ldots, m \tag{2}
\]

\[
x_{ij} \geq 0 \quad \text{and integer, for all} \quad i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, n \tag{3}
\]

where

- \( q_{ij} \) denotes the probability of the survival of target \( j \) if a single weapon of type \( i \) is assigned to it
- \( V_j \) denotes the value of the target \( j \)
Related Work

WTA well studied.

1 Around 1990 by Hosein and Anthans proposed a dynamic WTA model (Command & Control (C2)) .
   • Processes designed supports human operator in decision-making process.

2 Karasakal presented two integer linear programming models.
   • To determine the probability of shooting down all incoming targets.
   • The challenge of allocating the air defense missiles to incoming air targets was addressed.
Problem Formulation

- Military base camp (Where planning and strategies are being formed).
Problem Formulation

- Battlefield positioning.
Problem Formulation

- Upon target detection weapons are assigned.
- Firing period begins.
WTA Model as an Markov Decision Process

- Modelling the problem as an MDP (Certain properties).
- Actions and states.
- States evolve according to certain dynamics (Markov Property).
- Advantages of MDPs (Lots of solution techniques)
- What is an MDP?
- What is an MDP solution?
- What is an optimal MDP solution?
A Markov decision process is a tuple $< S, A, P, R >$ where:

- $S$: is a set of states
- $A$: is a set of actions
- $P$: is a transition distribution $P : S \times A \times S \rightarrow [0, 1]$
- $R$: is a reward function $R : S \times A \times S \rightarrow R$
MDP: Solution

- Choose an action in a state (policy $\pi$)

- What does a policy look like?
  - Policy is just a distribution over actions for a state.
  - This is the random policy: $\pi(S) = 1/|A|$ for all $S$ (where $|A|$ is the number of actions)

- MDP solution is a probabilistic state-dependent
  $\pi : S \times A \mapsto [0, 1]$

- Optimal MDP solution ($\pi^*$).
  - Choose the action that gives the best long term reward in every situation.
  - i.e. $\pi^*(S) \geq \pi(S)$ for all $\pi$ and $S$
Reinforcement Learning (RL)

- A simulation-based learning paradigm useful on large-scale and complex MDPs
- This will help us in determining the optimal policy in troops deployment.
- The RL agent interacts with the environment.
Implementation

Agent that make decisions from the distribution of actions for a state. At each time the agent decides whether

- take cover/hide ($C$)
- Advance ($A$)
- shoot ($S$)

The agent makes a decision based on the position of the target.

- cover ($c$),
- range ($r$),
- target know your position ($k$) and
- there is a target (enemy) ($e$).
Implementation

The state $S = \{crke\} \in S$ where each feature is a binary $\{1, 0\}$
Future Work

• Partial information of the battlefield. (POMPD)
• Scaling up the problem. (Abstraction techniques)
• Translating the POMDP to a usable interface.
Kea Leboha!!!! Thank YOU ALL!!!!