The RED Network

Push Forward the use of DEVS Modeling and Simulation Through Research and Engineering

DEVS MODELING AND SIMULATION BASED ON MARKOV DECISION PROCESS OF FINANCIAL LEVERAGE EFFECT IN THE EU DEVELOPMENT PROGRAMS

E. Barbieri*, L. Capocchi*, J.F. Santucci*
*University of Corsica - SPE UMR CNRS 6134
{barbieri_e,capocchi_l,santucci_j}@univ-corse.fr
Outline

1. Introduction & Context
2. Objectives
3. Machine Learning
4. DEVS-based Markovian Decision Process
5. Case study: Stock Market indices’s leverage effect on IF
6. Conclusion
Introduction

- Need to enhance the stakeholder IFC during the EU regional development program instruction phase.
- Need to push forward:
  - Markovian Decision Process (MDP) associating a Reinforcement Learning algorithms and
  - DEVS formalism.
- Build a collaborative generic framework to deal with decision making in a disruptive system.
Context

- DEVS formalism and environnement.


- Regional economy growth: The optimisation of the EU development program leverage effect a strong regional economics underlying asset.
Objectives

● The disruptive proposed approach associate DEVS and Reinforcement Learning algorithm in order to help optimal decision making in discrete-event model.
● Optimize the expected leverage effect of the EU development program.
● Model and simulate a Reinforcement Learning (RL) DEVS model to build a managing and monitoring volatility leverage effect policy decision tool.
Machine Learning

1. Supervised learning: classification, regression
2. Unsupervised learning: clustering
3. Reinforcement learning:
   - more general than supervised/unsupervised learning
   - learn from interaction between agent and environment to maximize an expected long term reward
   - solving an RL problem
     - Dynamic Programming
     - Monte Carlo methods
     - Simplest Temporal-Difference learning (Q-learning)
Machine Learning

Q-learning (from Peter Bodík RAD Lab, UC Berkeley)

- based on a Q function (state-action map)

\[ Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[ r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right] \]

- Q directly approximates Q* (Bellman optimality eqn)

- Optimal policy

\[ \pi^*(s) = \arg \max_a Q^*(s, a) \]
1. Initialize Q-values arbitrarily

2. Until learning is stopped…
   a. Choose an action (a) in the current world state (s) based on current Q-value estimates.
   b. Take the action (a) and observe the new state (S_{t+1}) and reward (r_{t+1}).
   c. Update \( Q(s,a) = Q(s,a) + \alpha [r + \gamma \max_{a'} Q(s',a') - Q(s,a)] \)
DEVS-based Markovian Decision Process

- DEVS formalism involves the use of experimental frames, which would permit to integrate the loops required by Q-Learning to find the optimal policy.
- DEVS allows an event-based implementation of the Q-Learning algorithm (improves the control of the algorithm)
- Two distinct atomic models:
  - Agent: generate the actions and proceed to the Q function update
  - Environnement: generate the new state and reward in response to an action from the agent.
DEVS-based Markovian Decision Process

- MDP with DEVSimPy

```python
def UpdateQLearning(self, s0, s1, a0, r):
    """
    1. use any policy to estimate Q
    Q(s0,a0)=Q(s0,a0)+alpha[r + gamma.maxQ(s1)-Q(s0,a0)]
    2. Q directly approximates Q^\pi (Bellman optimality eqn)
    3. independent of the policy being followed
    4. only requirement: keep updating each (s,a) pair
    """
    ga = np.max(self.Q[s1,:])
    td = r + self.y*ga - self.Q[s0,a0]
    # Q update
    self.Q[s0,a0] += self.lr*td
    # Q mean update
    self.Q_old_mean = self.Q_new_mean
    self.Q_new_mean = self.Q.mean()
```

E. Barbieri, JDF Les journées DEVS francophones : Théorie et Applications - April 30- May 4 2018
Case study: Stock market indices’s leverage on IF

- EU development programs are leverage effect programs.
- Stop freezing the economic growth up to 24 months during the pre-trial phase.
- Use the market leverage effect to reduce the expected economic growth increase time.
- Minimize the risk of economic growth loses due to project abandonment.
Case study: Stock market indices’s leverage on IF

- Enhancement of the stakeholder borrowing capacity.
- Reducing the risk of financial leverage by investing on world stock indices (underlying).
- Use the leverage effect: buying WSI for 2 up to 10 the total value of the IF.
Case study: Stock market indices’s leverage on IF

- Consider a share or a stock market not individually but as the global market of stock market indices.
- WSIs represents the global environment.
- Minimize the IF loses by a strong monitoring of the underlying.
Case study: Stock market indices’s leverage on IF

- **DEVSimPy modeling**
  - *Index1, Index2, Index3*: generators of indices in “real time”.
  - *EnvQLearning* and *Agent* atomic models for MDP modeling.
  - *ViewState* for display the finite state automata.
  - *Qmean*: mean of the Q matrix (for convergence)
Case study: Stock market indices’s leverage on IF

Simulation results exemple:
environment change + 3k cash

init state

end state
Case study: Stock market indices’s leverage on IF

Simulation results

Q-Learning convergence

Optimal policy

Q-Vars: 16
Start state (*): 2SBUX 2INTC
End state (**): 2SBUX 2WDC 2INTC
Policy:
1INTC: nothing for SBUX / nothing for WDC / buy 1 INTC
2INTC: buy 2 SBUX / nothing for WDC / nothing for INTC
1WDC: nothing for SBUX / nothing for WDC / buy 2 INTC
1WDC 1INTC: buy 2 SBUX / nothing for WDC / nothing for INTC
1WDC 2INTC: buy 2 SBUX / nothing for WDC / nothing for INTC
2WDC: buy 2 SBUX / nothing for WDC / nothing for INTC
2WDC 1INTC: buy 2 SBUX / nothing for WDC / nothing for INTC
2WDC 2INTC: buy 2 SBUX / nothing for WDC / nothing for INTC
1SBUX: nothing for SBUX / nothing for INTC / buy 2 WDC
1SBUX 1INTC: buy 1 SBUX / nothing for INTC / nothing for WDC
1SBUX 2INTC: buy 1 SBUX / nothing for INTC / nothing for WDC
1SBUX 1WDC: nothing for SBUX / nothing for WDC / buy 2 INTC
1SBUX 1WDC 1INTC: nothing for SBUX / nothing for WDC / buy 1 INTC
1SBUX 1WDC 2INTC: buy 1 SBUX / nothing for WDC / nothing for INTC
1SBUX 2WDC: buy 1 SBUX / nothing for WDC / nothing for INTC
1SBUX 2WDC 1INTC: buy 1 SBUX / nothing for WDC / nothing for INTC
1SBUX 2WDC 2INTC: buy 1 SBUX / nothing for WDC / nothing for INTC
2SBUX: nothing for SBUX / nothing for WDC / buy 2 INTC
2SBUX 1INTC: nothing for SBUX / buy 2 WDC / nothing for INTC
2SBUX 1WDC: nothing for SBUX / nothing for WDC / buy 2 INTC
2SBUX 1WDC 1INTC: nothing for SBUX / nothing for WDC / buy 1 INTC
2SBUX 1WDC 2INTC: buy 1 SBUX / nothing for WDC / nothing for INTC
2SBUX 2WDC: nothing for SBUX / nothing for WDC / buy 2 INTC
2SBUX 2WDC 1INTC: nothing for SBUX / nothing for WDC / buy 1 INTC
2SBUX 2WDC 2INTC (**): wait

E. Barbieri, JDF Les journées DEVS francophones : Théorie et Applications - April 30- May 4 2018
Case study: Stock market indices’s leverage on IF
Conclusion

● DEVS allows the modeling and simulation of MDP into an experimental frame.
● DEVS allows the control of the Q-Learning algorithm and will allow to improve its management.
● Adaptive environment changes management both Internal and External (+/-IF, market correction…)
● Future work:
  ○ Introduce the optimal decision time into the optimal policy (optimal action a* can be taken at time t for the state s)
  ○ Securisation the risk of IF loses by a strong monitoring of the market indices volatility on an multi-agent time calibrated decision making process.