MALT: Machine ALE Learning Technology

Christopher Yang, Jay Salmonson, Chris Young, Joe Koning, Luc Peterson
Weapons and Complex Integration/Design Physics
HYDRA$^1$ Simulates ICF Hohlraums

Mesh Entanglement

GOOD

CRASH

Lagrangian

Eulerian

Relax

Super Relax

Mesh moves and deforms

Fluid flows across fixed mesh

Refine a highly deformed mesh

Relax

+ Eulerian
Reinforcement Learning (RL) Roadmap

Finding the best mesh management policy through exploration

Every time step, the RL agent:

a. chooses relaxation directives
b. observes the result of the next step
c. receives a reward based on mesh quality

Goal: Maximize cumulative reward

1 Search
Identify problems with k-means clustering:
Acceptable zones are not relaxed
✓ Completed

2 Reward
After \( n \) time steps, compare:
RL Agent Policy
Issue reward based on mesh improvement
No Action*
*predicted by neural net\(^{1,3}\)
✓ Nearing Completion

3 Train
Offline learning:
Merlin + Hydra generates database for off-policy actor-critic agent
Online learning:
On-policy actor-critic algorithm
☐ Future work
1. Metrics Evaluate Mesh Quality
   - Scaled Jacobian
   - Condition Number

2. Rewards Assess RL Policies
   - Moment $M < 0$ if, on average, metrics improve
     \[ \Delta = M_{\text{no action}} - M_{\text{RL}} \]
   - Reward = \( \sum \text{metrics} s \Delta - s' \) (#relaxed nodes)
     \( (s, s' \text{ are scale factors}) \)

3. Example: Scaled Jacobian

References & Code

Combine with other metrics to calculate rewards!
Disclaimer
This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.