A Black-Box Approach to Sim-to-Real Transfer

Siddharth Desai, Haresh Karnan, Ishan Durugkar, Josiah Hanna, Garrett Warnell, Peter Stone
1. The Sim-To-Real Problem
2. Grounded Action Transformation (GAT)
3. End-to-End Action Transformers (RGAT)
4. GAT in Stochastic Domains (SGAT)
5. Questions?
The Sim-to-Real Problem

Real world training is

- Expensive
- Time Consuming
- Potentially unsafe
Our approach learns policies that are more robust to stochastic environments.
Sim-to-Real Paradigms

Learn a robust policy by training in multiple simulators
- Deploy directly in real world
- Tradeoff between generalization and performance

Train partially in simulation
- Finish training in real world
- Still prone to overfitting in simulator

Use a little real world data to ground simulator
- Iteratively improve simulator until desired performance is reached
Grounded Simulation Learning (GSL)

How?
Can we modify the simulator?

1. The Sim-To-Real Problem
2. **Grounded Action Transformation (GAT)**
3. End-to-End Action Transformers (RGAT)
4. GAT in Stochastic Domains (SGAT)
5. Questions?
Reinforcement Learning

Grounded Action Transformation (GAT)

GAT Example

(a) Simulator

(b) Real World
GAT Example

(a) Simulator

\[ s_0 \xrightarrow{a_1} s_1 \quad +1 \]
\[ s_0 \xrightarrow{a_2} s_2 \quad -1 \]

(b) Real World

\[ s_0 \xrightarrow{a_1} s_1 \quad +1 \]
\[ s_0 \xrightarrow{a_2} s_2 \quad -1 \]

(c) GAT Grounded Simulator

Forward Model
GAT Example

(a) Simulator

(b) Real World

(c) GAT Grounded Simulator

Inverse Model
GAT Example

(a) Simulator

(b) Real World

(c) GAT Grounded Simulator
1. The Sim-To-Real Problem
2. Grounded Action Transformation (GAT)
3. **End-to-End Action Transformers (RGAT)**
4. GAT in Stochastic Domains (SGAT)
5. Questions?
Reinforced Grounded Action Transformation (RGAT)

\[
\begin{align*}
    & s_{t+1}, r_{t+1} 
    \quad \xrightarrow{\text{Agent}} \quad a_t \\
    & f_{\text{real}}(s_t, a_t) \\
    \quad \xrightarrow{\sim} \quad \tilde{s}_{t+1} \\
    \quad \xrightarrow{f_{\text{sim}}^{-1}} \quad \tilde{a}_t \\
    \quad \xrightarrow{\text{Simulated Environment}} \quad \tilde{a}_t \\
    \quad \xrightarrow{\text{Action Transformer}} \quad \tilde{s}_{t+1}
\end{align*}
\]
Reinforced Grounded Action Transformation (RGAT)

Learning the action transformation function end-to-end makes it easier to learn small changes.
Reinforced Grounded Action Transformation (RGAT)
Reinforced Grounded Action Transformation (RGAT)

Treating the grounding step as a reinforcement learning problem matches the whole trajectory instead of individual transitions.
Reinforced Grounded Action Transformation (RGAT)

Agent → Simulated Environment → Action Transformer → f_{real}(s_t, a_t) → s_{t+1}

Use forward model to generate reward signal

Agent

Simulated Environment

Environment

s_{t+1} → \hat{a}_t → \hat{s}_{t+1}
RGAT Policies Deployed on the “Real” World

Grounding Step 1

Our method learns a good policy within two grounding steps.
Transformed actions when the “real” pendulum is...

- Heavier
- Lighter
1. The Sim-To-Real Problem
2. Grounded Action Transformation (GAT)
3. End-to-End Action Transformers (RGAT)
4. **GAT in Stochastic Domains (SGAT)**
5. Questions?
Stochastic Grounded Action Transformation (SGAT)

(a) Simulator

(b) Real World
Stochastic Grounded Action Transformation (SGAT)

(a) Simulator

(b) Real World

(c) GAT Grounded Simulator
Stochastic Grounded Action Transformation (SGAT)

(a) Simulator

(b) Real World
<table>
<thead>
<tr>
<th></th>
<th>Grounding Step 1</th>
<th></th>
<th>Grounding Step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speed (cm/s)</td>
<td>Falls</td>
<td>Speed (cm/s)</td>
</tr>
<tr>
<td>GAT</td>
<td>15.7 ± 2.98</td>
<td>6/10</td>
<td>18.5 ± 3.63</td>
</tr>
<tr>
<td>SGAT</td>
<td>16.9 ± 0.678</td>
<td>0/10</td>
<td>18.0 ± 2.15</td>
</tr>
</tbody>
</table>
Thank You!
1. The Sim-To-Real Problem
2. Grounded Action Transformation (GAT)
3. End-to-End Action Transformers (RGAT)
4. GAT in Stochastic Domains (SGAT)
5. Questions?