Scanning Probe Microscopy Based On Reinforcement Learning

Eric Yue Ma

Department of Applied Physics, Stanford University, CA 94305

Motivation and Goal

- Scanning probe microscopy (SPM) is one of the most important tools in solid-state and molecular science and technology.
- Scanning a sharp tip across the sample surface while maintaining a constant tip-sample interaction, usually achieved via proportional-integral (PI) feedback.
- Use machine learning to improve this feedback:
  - Automatic PI parameters tuning via a neural network.
  - Reinforcement learning based feedback without explicit PI parameters.

SPM Simulator

- Input: vertical position of the sample surface ($z_s$) and tip ($z_t$).
- Output: measured tip-sample interaction signal $s$.
- In the simplest case, $s$ is the position of a laser beam deflected by a micro-cantilever, which is linearly proportional to the tip-sample force upon contact:

$$s(z_t - z_s) = \max[0, -(z_t - z_s)] + \eta$$

- Direct gradient descent impractical due to local minima.

PI Tuning via Neural Network

- PI feedback:
  - $a^{(i)}_{t+1} := a^{(i)}_t + P(s - s_0) + I \sum_{t} (s - s_0)dt$

- Performance of different PI parameter combinations:

- Fitting the error vs. $(P, I)$ function with a neural network.

- 400x400 "brutal force" exploration of $(P, I)$ space: 160,000 line scans.

- Typical performance with $N_s = N_A = 30$, $\gamma = 0.9$.

- Problems: stochastic and bottlenecked learning; inferior performance than well-tuned PI feedback; a continuous valued state space does not qualitatively improve performance.

- Speculative cause: the error due to the unpredictable $z_s$ is comparable to that due to unoptimized MDP/policy, thus there is no way to learn well with a MDP model.

Reinforcement Learning Based Feedback

- The MDP model:

<table>
<thead>
<tr>
<th>$S$, $A$, $P_s$, $\gamma$, $R$</th>
<th>$S$</th>
<th>$A$</th>
<th>$P_s$</th>
<th>$\gamma$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>discretized $\Delta z_s$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discretized $\Delta z_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Each line scan is one trial, after which $P_s$ is updated, and also the value function and optimal policy.

- High-level expectation: the model will learn about the tip-sample interaction and the feature of the sample ($z_s$), to achieve good performance (small error $s - s_0$).

Conclusion

- Sample line scans + neural network fitting is a robust way to automatically tune the PI parameters.
- MDP based reinforcement learning feedback is not particularly suitable for SPM applications.