Reinforcement Learning Approach to Adaptive Maintenance Scheduling in Semiconductor Manufacturing

Mykhaylo Mykhaylov, BS in Computer Science

Under mentorship of Dr. Andi Wang

ABSTRACT

- **Objective**: develop a reinforcement learning (RL) model for maintenance scheduling in semiconductor manufacturing.
- **Approach**: Utilize future order information and analyze past and present maintenance records to generate optimized schedules
- **Performance**: Model outperforms conventional scheduling methods employed in the semiconductor industry.

PROBLEM FORMULATION

Inputs to the model from the environment come from future orders, machine current state and production, past maintenance records, and future scheduled maintenance. Two consequent states are 1 day apart.

- **Future orders:** model inputs are next 5 orders. Each order has the following:
  - Order quantity $N \sim \mathcal{N}(15000, 1000)$
  - Chip area $S \sim \mathcal{U}(1mm^2,\ 400mm^2)$
  - Price per chip $P = S \cdot \frac{50\$}{mm^2}$
  - Deadline in days is order quantity divided by expected daily production
  - Late penalty is 5 daily exp. profits/day

Machine state: model input is the state of the machine for the current day:

- Machine state: working, maintained, broken
- Yield follows exponential decay from the last day the machine was maintained and is adjusted to be smaller for larger chips
- Chips produced depend on yield and the chip size for the current order

- **Past maintenance records:** model inputs are days since the machine was maintained and/or broken
- **Future scheduled maintenance:** model input is next scheduled maintenance:
  - Binary value to indicate if maintenance was scheduled
  - Time in days until scheduled maintenance, 0 if today or not scheduled

Visually, the inputs to the model are represented as such:

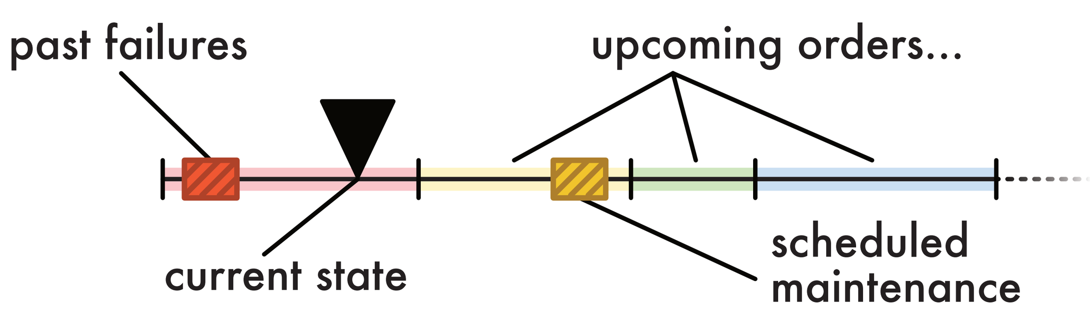

Fig 1.: Representation of semiconductor production state

Taking these inputs, model produces an **action**, which is one of the following:

- **Do nothing:** go to next day without changing maintenance schedule
- **Schedule maintenance:** add or replace scheduled maintenance 7/14/30/60/90 days ahead depending on the specific action taken
- **Deschedule maintenance:** remove scheduled maintenance

Model has a **reward** associated with each state, which is calculated as follows:

- **Daily production:** model is rewarded for satisfying part of current order with the chips produced on current day at order chip cost rate. 0 if no chips produced
- **Late penalty:** if the production is late on the current order, late penalty is subtracted from reward
- **Maintenance cost:** if machine is maintained, reward is negative cost of maintenance

METHODOLOGY

Model is based on **Proximal Policy Optimization** (PPO) proposed by (Schulman et al., 2017). This algorithm uses a two neural networks of same architecture, one to interact with the environment and the other to evaluate how good is it doing.

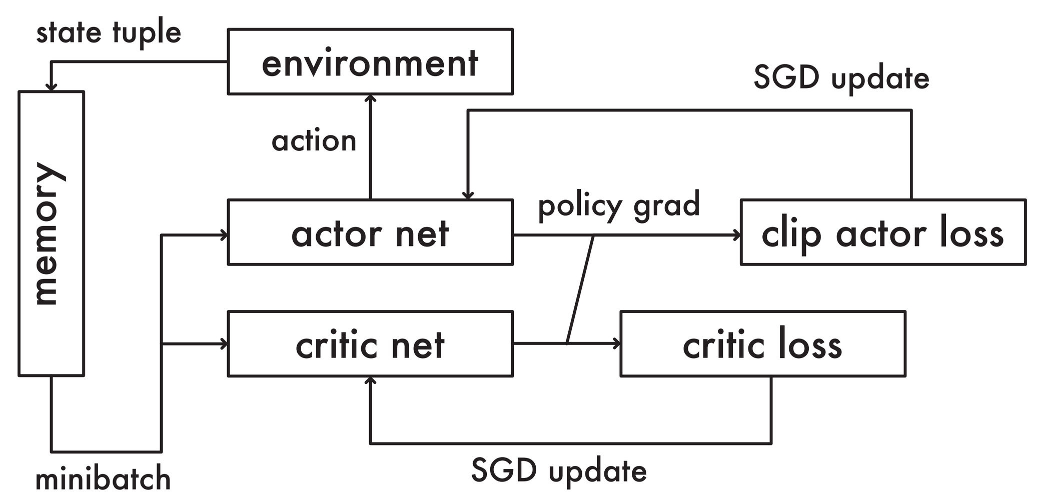

Fig 3.: PPO Architecture.

Below is a single iteration of the **training loop**:

1. Model runs current policy in environment for 1000 steps, collects (state, action, reward) tuples into memory
2. Calculate the advantage of the taken actions using the critic network
3. Loop through minibatches, use advantage to calculate actor network loss
4. Backpropagate loss and update bother networks using Adam optimizer

Network architecture is shared between the actor and critic. The only difference is in the out layer, where the actor has log-probabilities of 7 actions while critic has 1 "advantage" output.

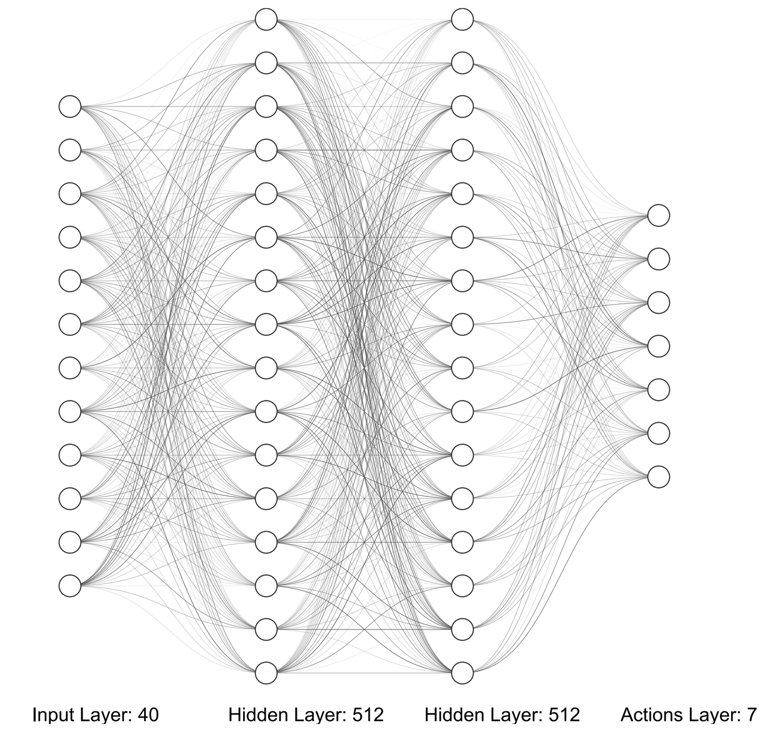

Fig 4.: Actor NN Architecture.

- Input is all model inputs which totals to 40 values
- Inner layers are 512 neurons, use ReLU, outputs use Softmax
- Actor output is 7 actions, critic is 1 "advantage" value

RESULTS

Model is compared to standard industry techniques for maintenance scheduling:

- **Ad-hoc:** no maintenance scheduling in advance, only fix when machine breaks
- **Bi-weekly:** schedule maintenance to be performed every 2 weeks
- **Monthly:** schedule maintenance to be performed every 4 weeks

There are two evaluation criteria:

- **Production loss due to failure:** how many chips were not produced due to machine breaking
- **Production loss due to maintenance:** how many chips were not produced while machine was on maintenance.

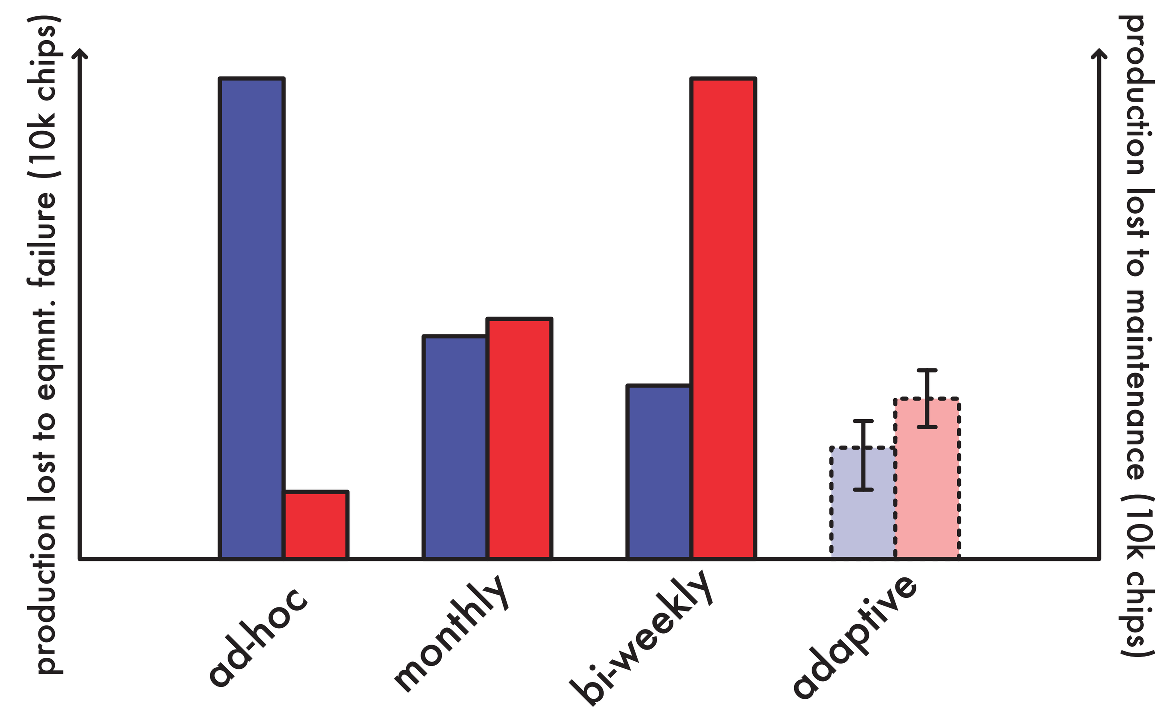

Fig 5.: Projected improvements at poster submission time