Reinforcement Learning: Training Environment Simulator

John Morrow*

Abstract

This paper describes an oven simulator for accelerating the training of a reinforcement learning agent used to provide optimal oven control settings for an industrial reflow soldering process.

1 Introduction

In a previous project [Morrow, 2022][3], a reinforcement learning model is trained to find the optimal control settings for a reflow oven used for soldering electronic components to a circuit board (Figure 1 & Figure 2). The oven's moving belt transports the product (i.e., the circuit board) through multiple heating zones. This process heats the product according to a temperature-time target profile required to produce reliable solder connections Figure 3.

Figure 1: **Reflow oven**
(image licensed from Adobe)

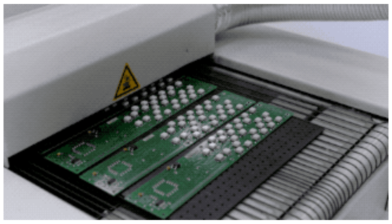

Figure 2: **Circuit boards on oven belt**
(image licensed from Adobe)

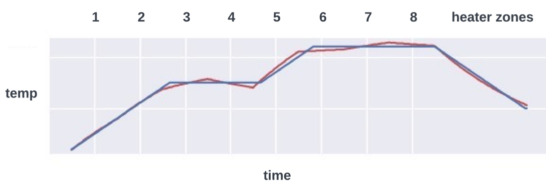

Figure 3: **Temperature-time profile** Blue trace: target profile. Red trace: actual product profile produced by oven.

*www.linkedin.com/in/johnmorrow1000

Since considerable time is required to stabilize an oven's temperature after changing the heater settings (up to 40 minutes) and passing the product through the oven (5 minutes), an oven simulator is used to speed up the process. The simulator emulates a single pass of the product through the oven in a few seconds compared to the minutes required by a physical oven.

The oven simulator has eight heating zones, each with a control for setting the temperature of the zone's heater (Figure 4). After each pass, the simulator provides the temperature readings of the product recorded as it traveled through the oven.

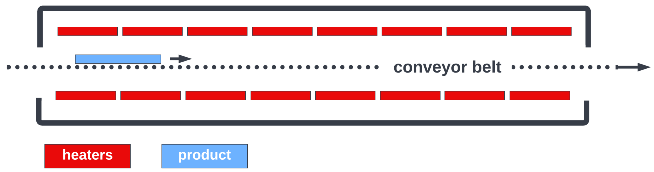

Figure 4: **Reflow oven schematic diagram**

2 Simulation model

The heating process is modeled using the finite-difference method. With this method, the product and the oven's heaters are modeled as many discrete elements as illustrated in Figure 5. The conductive and convective heat flow between the elements is illustrated in Figure 6.

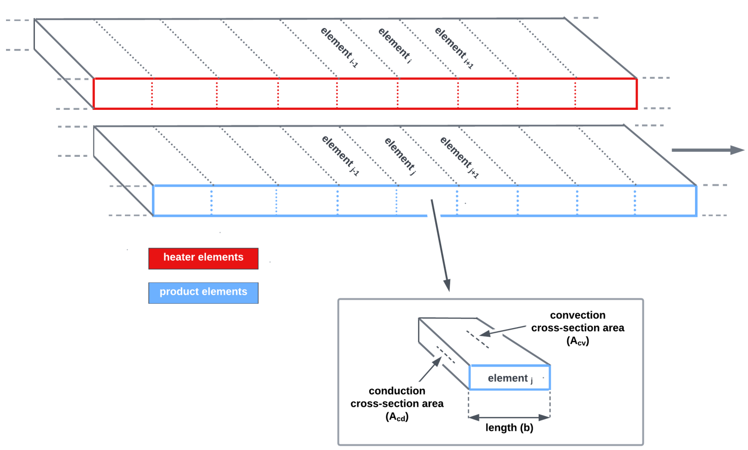

Figure 5: **Finite-difference discrete elements**

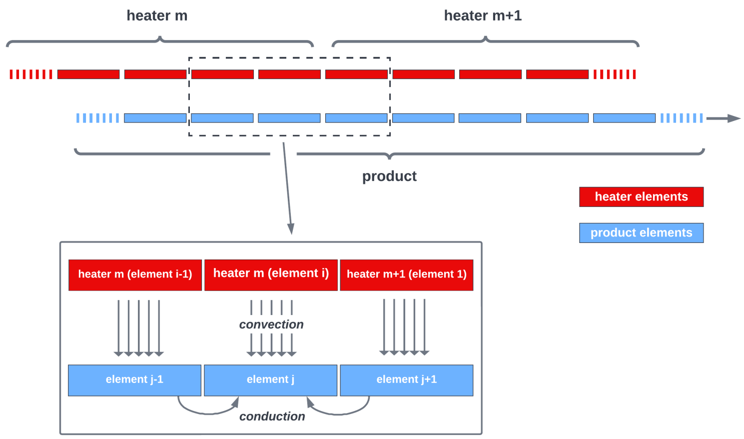

Figure 6: **Heat flow between element**

3 Simulation heat flow equations

The following equations define the conductive and convective heat flow between the elements: [Crank, 1975, pg.141][1] and [Lienhard, 2020, pp.13,22][2]

$$\Delta Q_{cd(l)} = \frac{k \cdot A_{cd}}{b} \cdot (T_{j-1}^{t} - T_{j}^{t}) \cdot \Delta t, \quad conduction: \ from \ element_{j-1} \tag{1}$$

$$\Delta Q_{cd(r)} = \frac{k \cdot A_{cd}}{b} \cdot (T_{j+1}^{t} - T_{j}^{t}) \cdot \Delta t, \quad conduction: \ from \ element_{j+1} \tag{2}$$

$$\Delta Q_{cv} = h \cdot A_{cv} \cdot (Th_i - T_{j}^{t}) \cdot \Delta t, \quad convection: \ from \ heater \ element_i \ (see \ note^1) \tag{3}$$

$$T_{j}^{t+1} = \frac{1}{c_p \cdot m} \cdot (\Delta Q_{cv} + \Delta Q_{cd(l)} + \Delta Q_{cd(r)}), \quad update \ for \ next \ time \ step \tag{4}$$

where,

$$\Delta Q_{cd(l)} : heat \ change \ in \ product \ element_j \ from \ element_{j-1}, \ J$$

$$\Delta Q_{cd(r)} : heat \ change \ in \ product \ element_j \ from \ element_{j+1}, \ J$$

$$\Delta Q_{cv} : heat \ change \ in \ product \ element_j \ from \ heater \ element_i, \ J$$

$$T_{j}^{t} : current \ temperature \ of \ product \ element_j \ at \ time \ step \ t, \ K$$

$$T_{j+1}^{t} : current \ temperature \ of \ product \ element_{j+1}, \ K$$

$$T_{j-1}^{t} : current \ temperature \ of \ product \ element_{j-1}, \ K$$

$$T_{j}^{t+1} : temperature \ of \ product \ element_j \ at \ next \ time \ step \ (t+1), \ K$$

$$Th_i : temperature \ of \ oven \ element_i, \ K$$

with constants,

$$A_{cd} : conduction \ cross-section \ area, \ m^2$$

$$A_{cv} : convection \ cross-section \ area, \ m^2$$

$$b : product \ element \ length \ along \ belt, \ m$$

$$k : thermal \ conductivity \ (product), \ W/(m \cdot K)$$

$$h : heat \ transfer \ coefficient \ (convection \ gas), \ W/(m^2 \cdot K)$$

$$c_p : specific \ heat \ (product), \ J/(kg \cdot K)$$

$$m : mass \ (product \ element), \ kg$$

$$\Delta t : simulation \ time \ step \ duration, \ s$$

[1] This applies for top heaters only. With top and bottom heaters active, the heat contribution from convection becomes $(2 \cdot \Delta Q_{cv})$.

4 Model stability criteria

The simulation model is stable when the stability factor, $r \leq \frac{1}{2}$. Equation 5 defines the stability factor. [Crank, 1975, pp. 138, 145][1] and [Lienhard, 2020, p. 18][2]

$$r = \frac{k \cdot \Delta t}{c_p \cdot \rho \cdot b^2} \tag{5}$$

where,

$$k : thermal\ conductivity\ (product),\ W/(m \cdot K)$$

$$\Delta t : simulation\ time\ step,\ s$$

$$c_p : specific\ heat\ (product),\ J/(kg \cdot K)$$

$$\rho : density\ (product),\ kg/m^3$$

$$b : product\ element\ length\ along\ belt,\ m$$

5 Algorithm

Algorithm 1 oven simulation algorithm

algorithm parameters
$\boldsymbol{oven}$ array of oven elements
$\boldsymbol{ovenx}$ extended oven array
$\boldsymbol{prod}$ array of product elements [2]
$\boldsymbol{Tr}$ array of product temperature results
num_ov number of ovenx elements
num_pr number of product elements
Th_i temperature of ovenx element i
T_j^t temperature of product element j at time step t
Tr_t temperature result at time step t
T_amb ambient temperature
procedure simulation steps
  $\boldsymbol{product} \leftarrow T\_amb$ ▷ intialize product elements
  $\boldsymbol{Tr} \leftarrow T\_amb$ ▷ intialize result array
  $\boldsymbol{ovenx} \leftarrow \boldsymbol{product} + \boldsymbol{oven} + \boldsymbol{product}$ ▷ append product elements to oven elements[3]
  $steps = num\_ov - num\_pr$
  for $t \leftarrow 0, steps$ **do**
    $T_{j=0}^t \leftarrow T_{j=1}^t$ ▷ update left dummy element
    $T_{j=num\_pr-1}^t \leftarrow T_{j=num\_pr}^t$ ▷ update right dummy element
    for $j \leftarrow 1, num\_pr$ **do**
      $\Delta Q_{cd(l)}(T_{j-1}^t, T_j^t)$ ▷ $\Delta Q_{cd(l)}(\cdot)$ [Equation 1]
      $\Delta Q_{cd(r)}(T_{j+1}^t, T_j^t)$ ▷ $\Delta Q_{cd(r)}(\cdot)$ [Equation 2]
      $\Delta Q_{cv}(Th_{j+t}, T_j^t)$ ▷ $\Delta Q_{cd(l)}(\cdot)$ [Equation 3]
      $T_j^{t+1} \leftarrow \frac{1}{c_p \cdot m} \cdot (\Delta Q_{cd(l)} + \Delta Q_{cd(r)} + \Delta Q_{cv})$ ▷ update element temperature [Equation 4]
    end for
    $Tr_{t+1} \leftarrow T_{(product\ center\ element)}^{t+1}$
  end for
  return $\boldsymbol{Tr}$
end procedure

[2] The product array includes a dummy element on the left and the right to account for conduction end effects.
[3] The ovenx array facilitates the product starting and ending outside the oven as the product array moves across the ovenx array through a sequence of time steps.

References

[1] J. Crank, *The Mathematics Of Diffusion, 2nd Edition*. Oxford, England: Oxford University Press, 1975, also available as http://www-eng.lbl.gov/~shuman/NEXT/MATERIALS&COMPONENTS/Xe_damage/Crank-The-Mathematics-of-Diffusion.pdf.

[2] J. H. Lienhard, IV and J. H. Lienhard, V, *A Heat Transfer Textbook*, 5th ed. Cambridge, MA: Phlogiston Press, 2020, version 5.10. [Online]. Available: http://ahtt.mit.edu

[3] J. Morrow, "Reinforcement learning: A case study in model generalization," 2022. [Online]. Available: https://github.com/jmorrow1000/RL-generalize