Performance Engineering Tutorial
Scaling, Forecasting & Control

Exercise 1. Using Amdhal's law, determine the parallel speedup for the following programs:

- Program A: 40% serial execution, 60% parallel execution
- Program B: 20% serial execution, 80% parallel execution
- Program C: 0% serial execution, 100% parallel execution

Assume two scenarios:

- Vertical scaling from a single-core machine to one with $n = 4$ cores.
- Vertical scaling from a single-core machine to one with $n = 8$ cores.

Solution:
Recall that Amdhal's law may be written as

$$S_n = \frac{T_1}{T_n} = \frac{T_1}{T_1(1-p) + pT_1/n} = \frac{n}{n(1-p)+p}$$

where p is fraction of the program execution time that is executed serially.

For $n = 4$ we therefore have:

- $S_4^A = \dfrac{4}{4 \cdot 0.4 + 0.6} = \frac{4}{2.2} = 1.82 \;\; (+182\%)$
- $S_4^B = \dfrac{4}{4 \cdot 0.2 + 0.8} = \frac{4}{1.6} = 2.50 \;\; (+250\%)$
- $S_4^C = \dfrac{4}{4 \cdot 0 + 1.0} = 4.00 \;\; (+400\%)$

Conversely for $n = 8$ we have:

- $S_8^A = \dfrac{8}{8 \cdot 0.4 + 0.6} = \frac{8}{3.8} = 2.11 \;\; (+211\%)$
- $S_8^B = \dfrac{8}{8 \cdot 0.2 + 0.8} = \frac{8}{2.4} = 3.33 \;\; (+333\%)$
- $S_8^C = \dfrac{8}{8 \cdot 0 + 1.0} = 8.00 \;\; (+800\%)$

Therefore, moving from 4 to 8 cores can have widely different effects depending on the fraction of serial and parallel execution.

Exercise 2. An autoscaling controller predicts the number of request arrivals A_t in timeslot t using an autoregressive AR(1) process. Assume A_t to be stationary. Suppose that the monitoring system reports the following statistics for A_t: mean $E[A_t] = 1$, variance $Var[A_t] = 4$, and lag-1 auto-covariance $K_1 = 2$. Fit the AR(1) process parameters to the data.

Solution:
For an AR(1) process we can write

$$A_t = c + \phi A_{t-1} + \epsilon_t$$

and the parameters are c, ϕ, and the white-noise variance σ_ϵ^2. Since $K_1 = Var[A_t]\phi = 4\phi = 2$ it must be $\phi = 0.5$. From the measurements we have then the conditions

$$E[A_t] = \frac{c}{1-\phi} = 1$$

$$Var[A_t] = \frac{\sigma_\epsilon^2}{1-\phi^2} = 4$$

which imply $c = 0.5$ and $\sigma_\epsilon^2 = 4 \cdot 0.75 = 3.0$.

Exercise 3. For a stationary AR(1) process:
Question 3.1 Show that the variance of the time series is given by

$$Var[A_t] = \frac{\sigma_\epsilon^2}{1-\phi^2}, \qquad \forall t$$

Solution:
We note first that by the properties of the Variance

$$Var[A_t] = Var[c + \phi A_{t-1} + \epsilon_t] = 0 + \phi^2 Var[A_{t-1}] + \sigma_\epsilon^2$$

where we used that $Var[c] = 0$, since c is a constant and not a random variable, and $Var[\epsilon_t] = \sigma_\epsilon^2$.
Note that the last formula assumes that the covariance between A_{t-1} and ϵ_t is zero. This can be shown to be true by observing that A_{t-1} can be recursively expressed as a linear sum of white noises, and the covariance of two distinct white noises is zero since they are purely random processes, uncorrelated with each other.
Using stationarity, we then write $Var[A_t] = Var[A_{t_1}] = V$, $\forall t$, so that

$$V = \phi^2 V + \sigma_\epsilon^2 \Rightarrow V = \frac{\sigma_\epsilon^2}{1-\phi^2}$$

Question 3.2 Show that the variance of the one-step ahead prediction is given by

$$Var[A_{t+1}|A_t] = \sigma_\epsilon^2$$

Solution:

$$\begin{aligned} Var[A_{t+1}|A_t] &= Var[c|A_t] + \phi_1^2 Var[A_t|A_t] + Var[\epsilon_t|A_t] \\ &= Var[\epsilon_t|A_t] \\ &= \sigma_\epsilon^2 \end{aligned}$$

where we used that $Var[c|A_t] = 0$ and $Var[A_t|A_t] = 0$ since in both cases the argument is a constant (in the second case because by conditioning on A_t, then A_t is a known constant).

Exercise 4. An autoscaling controller predicts the number of request arrivals A_t in timeslot t using an autoregressive AR(1) process. Assume A_t to be stationary. Suppose that the monitoring system reports the following statistics for A_t: mean $E[A_t] = 1$, variance $Var[A_t] = 2$, and lag-1 auto-covariance $K_1 = 0.8$. Fit the AR(1) process parameters to the data.

Solution:
For an AR(1) process we can write

$$A_t = c + \phi A_{t-1} + \epsilon_t$$

and the parameters are c, ϕ, and the white-noise variance σ_ϵ^2.

Because

$$K_1 = Var[A_t]\phi = 2\phi = 0.8$$

it must be $\phi = 0.4$.

From the measurements we have then the conditions

$$E[A_t] = \frac{c}{1-\phi} = 1$$

$$Var[A_t] = \frac{\sigma_\epsilon^2}{1-\phi^2} = 2$$

which imply $c = 0.6$ and $\sigma_\epsilon^2 = 1.68$.

Exercise 5. A network element incurs at time t a packet loss rate y_t. The loss can be adjusted using the sliding window size parameter u_t. The temporal correlations between the two signals are captured by the input-output model

$$y_t = 0.4y_{t-1} + y_{t-2} - u_{t-1} - 2u_{t-2}$$

Let $Y(z) = \mathcal{Z}[y_t]$ and $U(z) = \mathcal{Z}[u_t]$ be the z-transforms for y_t and u_t, respectively. Determine the transfer function $H(z) = Y(z)/U(z)$ for the network protocol.

Solution:
Taking the z-transforms of both sides, due to linearity we can write

$$\mathcal{Z}[y_t] = 0.4\mathcal{Z}[y_{t-1}] + \mathcal{Z}[y_{t-2}] - \mathcal{Z}[u_{t-1}] - 2\mathcal{Z}[u_{t-2}]$$

Using the time delay property of the transform we get

$$\mathcal{Z}[y_t] = 0.4z^{-1}\mathcal{Z}[y_t] + z^{-2}\mathcal{Z}[y_t] - z^{-1}\mathcal{Z}[u_t] - 2z^{-2}\mathcal{Z}[u_t]$$

This may be rewritten as

$$Y(z) = 0.4z^{-1}Y(z) + z^{-2}Y(z) - z^{-1}U(z) - 2z^{-2}U(z)$$

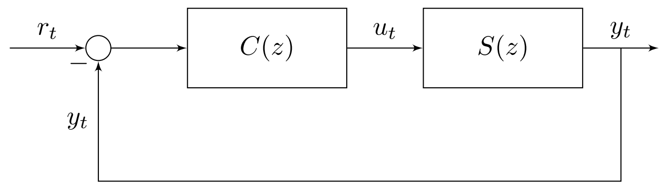
Grouping terms and dividing we find

$$H(z) = \frac{Y(z)}{U(z)} = \frac{-z^{-1} - 2z^{-2}}{1 - 0.4z^{-1} - z^{-2}} = \frac{z+2}{1 - z^2 + 0.4z}$$

Exercise 6. A closed-loop controller can periodically adjust a server configuration option u_t, at discrete time instants $t = 1, 2, \ldots$. This option allows to control the server queue-length. Let r_t be the target queue-length level at time t and let y_t be the monitored queue-length level in the server. Denote by $S(z)$ the transfer function of the server and by $C(z)$ the transfer function of the controller. Assume the following assignments

$$S(z) = \frac{1}{4z+1} \qquad\qquad C(z) = \frac{2z}{3z-1}$$

and a block diagram



where $e_t = r_t - y_t$ is an error signal.

Question 6.1 Determine the transfer function for the entire system.

Solution:
The transfer function for the entire system is

$$H(z) = \frac{C(z)S(z)}{1 + C(z)S(z)} = \frac{C(z)S(z)}{1 + C(z)S(z)} = \frac{2z}{12z^2 + z - 1}$$

Question 6.2 State if the system is stable and, if so, estimate its settling time.

Solution:
The roots of the polynomial $12z^2 + z - 1$ are

$$\lambda_1 = -1/3; \qquad \lambda_2 = 1/4;$$

therefore $\lambda = \max_{k=1,2} |\lambda_k| = 1/3$. Since $\lambda \leq 1$, the system is stable.

The settling time T will give us an idea of the time it takes on average for the system to converge to its normal regime following a change in input. This may be conservatively estimated as

$$T = -4/\log\lambda = 3.6410 \approx 4 \text{ time steps}$$

where the approximation takes into account that the number of steps needs to be discrete.

Question 6.3 Determine the long-term ratio between output signal y_t and input signal r_t.

Solution:
Let β denote the steady-state gain. This is by definition the long-term ratio between input signal r_t and output signal y_t. This can be obtained from the transfer function as

$$\beta = H(1) = \frac{1}{6}$$