Broadcast in the α-β-Model

The time taken to send a message of size s from one process to another is $T(s) = \alpha + s\beta$. If a process sends a message of size s at the time t it cannot send another message before $t + T(s)$.

In the lecture we have seen the analysis of a broadcast over a binary and a binomial tree. However, we can also define a k-ary as well as a k-nomial tree broadcast. In a k-ary tree brodcast every node forwards the received message to k children. A k-nomial tree is produced by forwarding the message to $k - 1$ children every round, until all processes are reached.

1. What is the runtime of a k-ary tree broadcast in the $\alpha\beta$ model if we assume small messages, i.e., $s = 1$?
2. What is the runtime of a k-nomial tree broadcast in the $\alpha\beta$ model if we assume small messages, i.e., $s = 1$?

Communication Cost Models

1. What are the differences between the $\alpha\beta$ model, the LogP and the LogGP model?
2. Can you think of useful additions to those models? Why could they be inaccurate in practice?

1 Strassen Matrix-Multiplication

In the lecture, we discussed recursive matrix multiplication

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \cdot \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

computed as

$$\begin{aligned} C_{11} &= A_{11} \cdot B_{11} + A_{12} \cdot B_{21} \\ C_{21} &= A_{21} \cdot B_{11} + A_{22} \cdot B_{21} \\ C_{12} &= A_{11} \cdot B_{12} + A_{12} \cdot B_{22} \\ C_{22} &= A_{21} \cdot B_{12} + A_{22} \cdot B_{22} \end{aligned}$$

which has the following communication cost (see lecture for derivation)

$$T_{\mathrm{MM}}(n, p) = O\left(\frac{n^2}{p^{2/3}} \cdot \beta\right) + O(\log(p) \cdot \alpha).$$

Strassen's algorithm reorganizes 2-by-2 block matrix multiplication and achieves a lower asymptotic computation cost complexity ($O(n^{log_2(7)})$ instead of $O(n^3)$):

$$
\begin{aligned}
M_1 &= (A_{11} + A_{22}) \cdot (B_{11} + B_{22}) \\
M_2 &= (A_{21} + A_{22}) \cdot B_{11} \\
M_3 &= A_{11} \cdot (B_{12} - B_{22}) \\
M_4 &= A_{22} \cdot (B_{21} - B_{11}) \\
M_5 &= (A_{11} + A_{12}) \cdot B_{22} \\
M_6 &= (A_{21} - A_{11}) \cdot (B_{11} + B_{12}) \\
M_7 &= (A_{12} - A_{22}) \cdot (B_{21} + B_{22}) \\
C_{11} &= M_1 + M_4 - M_5 + M_7 \\
C_{21} &= M_2 + M_4 \\
C_{12} &= M_3 + M_5 \\
C_{22} &= M_1 - M_2 + M_3 + M_6
\end{aligned}
$$

Derive the asymptotic communication cost of Strassen's algorithm on p processors using the α-β model. Does Strassen's algorithm have a higher flop/byte (computation per communication) ratio than regular standard matrix multiplication?