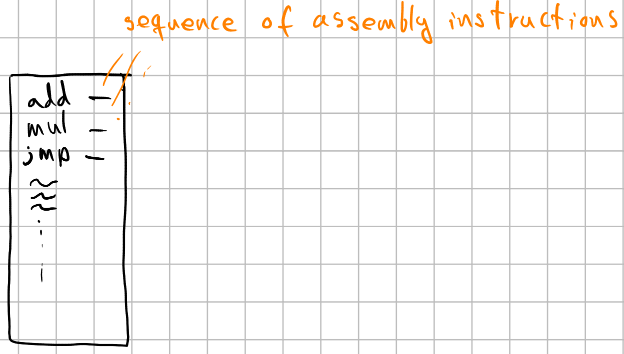
Amdahl's law: (1967)

Given a program,

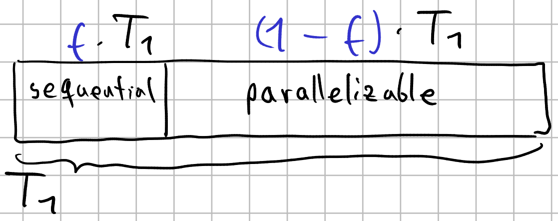
how much speedup can we

achieve by using multiple processors?



Let T_1 the total execution time on 1 processor.

Let f the portion of the execution time (or "work") that is spent on sequential (= non-parallelizable) work.



At best, it will take

$$f\cdot T_1+\frac{(1-f)\cdot T_1}{p}=T_1\cdot\left(f+\frac{1-f}{p}\right)=T_p$$

to execute this on p processors.

$$\Rightarrow \quad S_p\le\frac{T_1}{T_p}=\frac{T_1}{T_1\cdot\left(f+\frac{1-f}{p}\right)}=\frac{1}{f+\frac{1-f}{p}}\xrightarrow{p\to\infty}\frac{1}{f}$$

speedup on p processors

e.g. let $f=\frac{1}{n}$, $T_1=1$. With $p\to\infty$ many processors, the dur. with mult. processors is:

$$\frac{1}{n}+\underbrace{\frac{1-\frac{1}{n}}{p}}_{\to 0}=\frac{1}{n}$$

⇒ parallel program is $n=1/f$ times faster.

→ With e.g. $f=\frac{1}{4}$ (25% of the execution time is spent on doing non-parallelizable work, e.g. incrementing a counter, creating threads, ...)

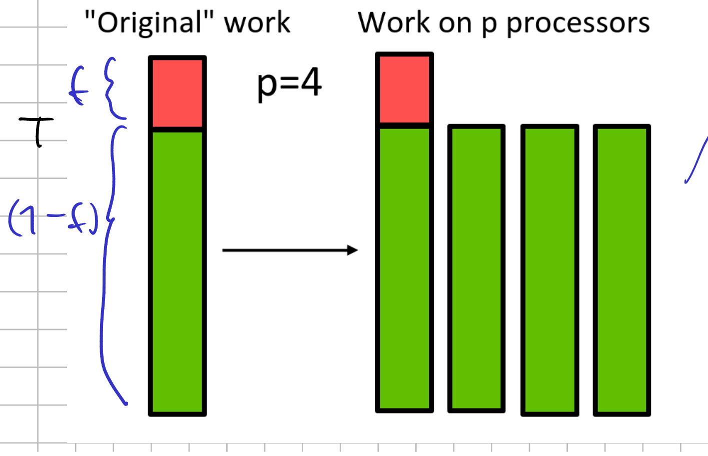
then $\underline{S_p\le 4}$!

Gustafson's Law: (1988)

Above we assumed that the amount of work is fixed. Things could be different in reality:

e.g. when a user replaces an old laptop, he/she is likely to want to run more programs at the same time since it was too slow before, instead of just running the exact same programs as before and expecting them to be faster.

Gustafson assumes that the parallel work in the program scales linearly with the number of processors:



T:= time for original work sequentially

Time for new work:

$$T_1=f\cdot T+p\,(1-f)\cdot T$$

$$T_p=f\cdot T+\frac{p\cdot(1-f)\cdot T}{p}$$

$$=T$$

$$\Rightarrow \quad S_p=\frac{T_1}{T_p}=f+(1-f)\cdot p$$

An alternative way to arrive at the derivations:

We can measure "speed" by work done over time: $\frac{W}{T}$.

How fast with ↗ 1 processor? → $\frac{W_1}{T_1}$

How fast with ↘ p processors? → $\frac{W_p}{T_p}$

⇒ How much faster with p proc.? → $S_p = \frac{W_p/T_p}{W_1/T_1}$

(parallel speed > sequential speed ⇒ $S_p > 1$)

$$S_p = \frac{W_p/T_p}{W_1/T_1} \xrightarrow[W_1 = W_p]{\text{Amdahl}} S_p = \frac{T_1}{T_p}$$

$$S_p = \frac{W_p/T_p}{W_1/T_1} \xrightarrow[T_1 = T_p]{\text{Gustafson}} S_p = \frac{W_p}{W_1} = \frac{f \cdot \cancel{W_1} + (1-f) \cdot \cancel{W_1} \cdot \underline{p}}{\cancel{W_1}}$$

Gustafson assumes that parallel work scales with p. With p proc., p times more parallel work can be done in the same time.

$$= f + (1-f) \cdot p$$