Module 1

A **parallel algorithm** is an algorithm that can execute several instructions simultaneously on different processing devices and then combine all the individual outputs to produce the final result.

Parallel Algorithm − The problem is divided into sub-problems and are executed in parallel to get individual outputs. Later on, these individual outputs are combined together to get the final desired output. It is not easy to divide a large problem into **sub-problems**. Sub-problems may have data dependency among them. Therefore, the processors have to communicate with each other to solve the problem.

Model of Computation

Both sequential and parallel computers operate on a set (stream) of instructions called algorithms. These set of instructions (algorithm) instruct the computer about what it has to do in each step.

Depending on the instruction stream and data stream, computers can be classified into four categories −

- Single Instruction stream, Single Data stream (SISD) computers
- Single Instruction stream, Multiple Data stream (SIMD) computers
- Multiple Instruction stream, Single Data stream (MISD) computers
- Multiple Instruction stream, Multiple Data stream (MIMD) computers

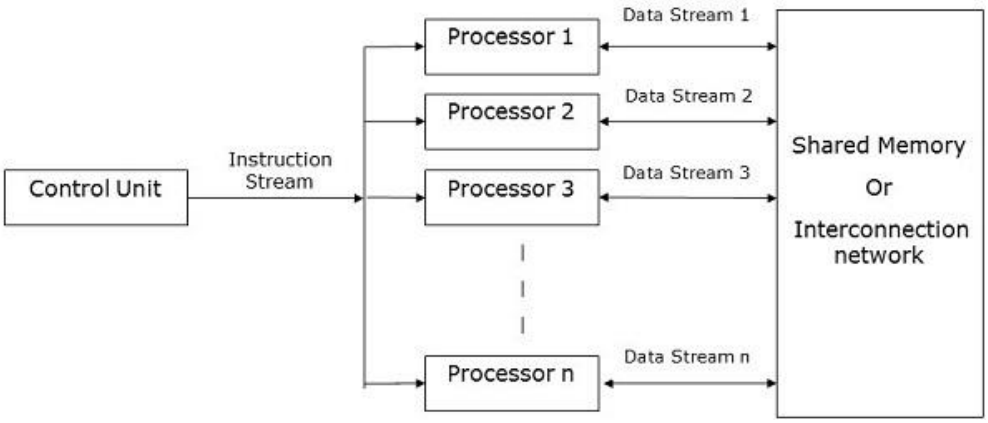
SISD Computers

SISD computers contain **one control unit, one processing unit,** and **one memory unit**.

In this type of computers, the processor receives a single stream of instructions from the control unit and operates on a single stream of data from the memory unit. During computation, at each step, the processor receives one instruction from the control unit and operates on a single data received from the memory unit.

SIMD Computers

SIMD computers contain **one control unit, multiple processing units,** and **shared memory or interconnection network**.



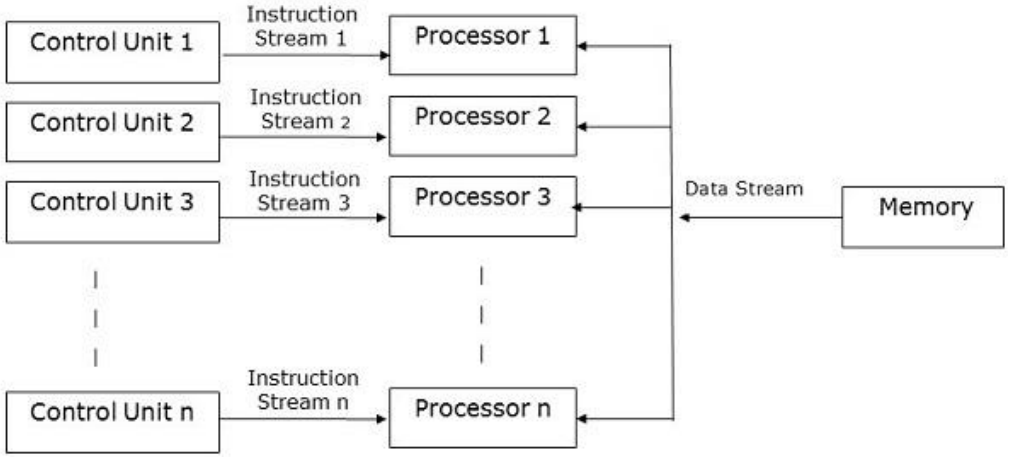
Here, one single control unit sends instructions to all processing units. During computation, at each step, all the processors receive a single set of instructions from the control unit and operate on different set of data from the memory unit.

Each of the processing units has its own local memory unit to store both data and instructions. In SIMD computers, processors need to communicate among themselves. This is done by **shared memory** or by **interconnection network**.

While some of the processors execute a set of instructions, the remaining processors wait for their next set of instructions. Instructions from the control unit decides which processor will be **active** (execute instructions) or **inactive** (wait for next instruction).

MISD Computers

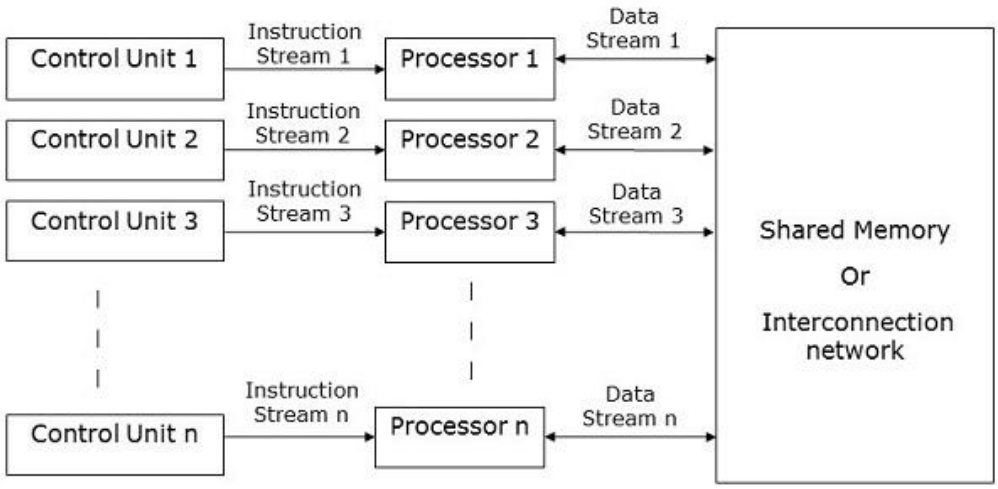
As the name suggests, MISD computers contain **multiple control units, multiple processing units,** and **one common memory unit**.



Here, each processor has its own control unit and they share a common memory unit. All the processors get instructions individually from their own control unit and they operate on a single stream of data as per the instructions they have received from their respective control units. This processor operates simultaneously.

MIMD Computers

MIMD computers have **multiple control units, multiple processing units,** and a **shared memory** or **interconnection network**.



Here, each processor has its own control unit, local memory unit, and arithmetic and logic unit. They receive different sets of instructions from their respective control units and operate on different sets of data.

Note

- An MIMD computer that shares a common memory is known as **multiprocessors,** while those that uses an interconnection network is known as **multicomputer**.
- Based on the physical distance of the processors, multicomputer are of two types −
  - **Multicomputer** − When all the processors are very close to one another (e.g., in the same room).
  - **Distributed system** − When all the processors are far away from one another (e.g.- in the different cities)

Analysis of an algorithm helps us determine whether the algorithm is useful or not. Generally, an algorithm is analysed based on its execution time **(Time Complexity)** and the amount of space **(Space Complexity)** it requires.

Since we have sophisticated memory devices available at reasonable cost, storage space is no longer an issue. Hence, space complexity is not given so much of importance.

Parallel algorithms are designed to improve the computation speed of a computer. For analysing a Parallel Algorithm, we normally consider the following parameters −

- Time complexity (Execution Time),
- Total number of processors used, and
- Total cost.

Time Complexity

The main reason behind developing parallel algorithms was to reduce the computation time of an algorithm. Thus, evaluating the execution time of an algorithm is extremely important in analysing its efficiency.

Execution time is measured on the basis of the time taken by the algorithm to solve a problem. The total execution time is calculated from the moment when the algorithm starts executing to the moment it stops. If all the processors do not start or end execution at the same time, then the total execution time of the algorithm is the moment when the first processor started its execution to the moment when the last processor stops its execution.

Time complexity of an algorithm can be classified into three categories−

- **Worst-case complexity** − When the amount of time required by an algorithm for a given input is **maximum**.
- **Average-case complexity** − When the amount of time required by an algorithm for a given input is **average**.
- **Best-case complexity** − When the amount of time required by an algorithm for a given input is **minimum**.

Asymptotic Analysis

The complexity or efficiency of an algorithm is the number of steps executed by the algorithm to get the desired output. Asymptotic analysis is done to calculate the complexity of an algorithm in its theoretical analysis. In asymptotic analysis, a large length of input is used to calculate the complexity function of the algorithm.

Note − **Asymptotic** is a condition where a line tends to meet a curve, but they do not intersect. Here the line and the curve are asymptotic to each other.

Asymptotic notation is the easiest way to describe the fastest and slowest possible execution time for an algorithm using high bounds and low bounds on speed. For this, we use the following notations −

- Big O notation
- Omega notation
- Theta notation

Big O notation

In mathematics, Big O notation is used to represent the asymptotic characteristics of functions. It represents the behaviour of a function for large inputs in a simple and accurate method. It is a method of representing the upper bound of an algorithm's execution time. It represents the longest amount of time that the algorithm could take to complete its execution. The function −

$$f(n) = O(g(n))$$

iff there exists positive constants **c** and $\mathbf{n_0}$ such that $\mathbf{f(n) \leq c * g(n)}$ for all **n** where $\mathbf{n \geq n_0}$.

Omega notation

Omega notation is a method of representing the lower bound of an algorithm's execution time. The function −

$$f(n) = \Omega\ (g(n))$$

iff there exists positive constants **c** and $\mathbf{n_0}$ such that $\mathbf{f(n) \geq c * g(n)}$ for all **n** where $\mathbf{n \geq n_0}$.

Theta Notation

Theta notation is a method of representing both the lower bound and the upper bound of an algorithm's execution time. The function −

$$f(n) = \theta(g(n))$$

iff there exists positive constants $\mathbf{c_1}$, $\mathbf{c_2}$, and $\mathbf{n_0}$ such that $c1 * g(n) \leq f(n) \leq c2 * g(n)$ for all **n** where $\mathbf{n \geq n_0}$.

Speedup of an Algorithm

The performance of a parallel algorithm is determined by calculating its **speedup**. Speedup is defined as the ratio of the worst-case execution time of the fastest known sequential algorithm for a particular problem to the worst-case execution time of the parallel algorithm.

$$\text{speedup} = \frac{\text{Worst case execution time of the fastest known sequential for a particular problem}}{\text{Worst case execution time of the parallel algorithm}}$$

Number of Processors Used

The number of processors used is an important factor in analysing the efficiency of a parallel algorithm. The cost to buy, maintain, and run the computers are calculated. Larger the number of processors used by an algorithm to solve a problem, more costly becomes the obtained result.

Total Cost

Total cost of a parallel algorithm is the product of time complexity and the number of processors used in that particular algorithm.

$$\text{Total Cost} = \text{Time complexity} \times \text{Number of processors used}$$

Therefore, the **efficiency** of a parallel algorithm is −

$$\text{Efficiency} = \frac{\text{Worst case execution time of sequential algorithm}}{\text{Worst case execution time of the parallel algorithm}}$$

The 15-puzzle problem is invented by sam loyd in 1878.

- In this problem there are 15 tiles, which are numbered from 0 – 15.
- The objective of this problem is to transform the arrangement of tiles from initial arrangement to a goal arrangement.
- The initial and goal arrangement is shown by following figure.

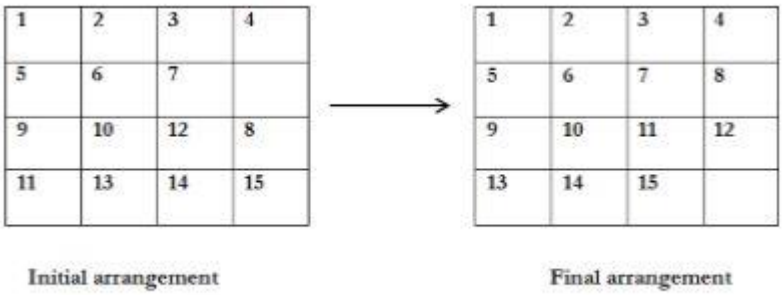


Figure 12

- There is always an empty slot in the initial arrangement.
- The legal moves are the moves in which the tiles adjacent to ES are moved to either left, right, up or down.
- Each move creates a new arrangement in a tile.
- These arrangements are called as states of the puzzle.
- The initial arrangement is called as initial state and goal arrangement is called as goal state.
- The state space tree for 15-puzzle is very large because there can be 16! Different arrangements.
- A partial state space tree can be shown in figure.
- In state space tree, the nodes are numbered as per the level.
- Each next move is generated based on empty slot positions.
- Edges are labelled according to the direction in which the empty space moves.
- The root node becomes the E – node.
- The child node 2, 3, 4 and 5 of this E – node get generated.
- Out of which node 4 becomes an E – node. For this node the live nodes 10, 11, 12 gets generated.
- Then the node 10 becomes the E – node for which the child nodes 22 and 23 gets generated.
- Finally, we get a goal state at node 23.
- We can decide which node to become an E – node based on estimation formula.

Estimation formula:

C (x) = f (x) + G (x)

Where, f (x) is length of path from root to node x.

G (x) is number of non-blank tiles which are not in their goal position for node x.

C (x) denotes the lower bound cost of node x.

State-Space Tree: