

WELCOME

ডাটা স্ট্রাকচার অ্যান্ড অ্যালগরিদম

বিষয় কোড: ২৮৫৪২

অধ্যায়: ১

ডাটা স্ট্রাকচার

ডাটা এবং ইনফরমেশন এর সংজ্ঞা (DEFINE DATA & INFORMATION)

ডাটা (Data): ডাটা শব্দের অর্থ হলো উপাত্ত। সাধারণ অর্থে কোন বিষয় বা ঘটনার মানকে ডাটা (Data) বা উপাত্ত বলা হয়। একাধিক Data Item নিয়ে একত্রে গ্রুপ আইটেম (Group Item) বলে। Data শব্দ Latin শব্দ Datum শব্দ-এর বহুবচন। Data এর কোন নিজস্ব স্বকীয় সত্তা নেই। এই Data-কে process করে যে output পাওয়া যায় তাকে information বলে। যেমন- রাসেল এর বয়স ৯ বছর। এখানে আলাদাভাবে যদি রাসেল, ৯, এগুলো চিন্তা করা হয় তবে এর কোন অর্থ নেই, কিন্তু পুরো বাক্যটিকে চিন্তা করলে একটি অর্থ পাওয়া যায়। নিম্নে একটি ডাটা টেবিল দেখানো হলো-

Name	Age	Snemister	Sex	Subject
Nazmul	23	8	M	CSE
Asadul	19	5	M	CSE
Rakib	18	3.	M	EEE

ডাটা প্রক্রিয়াকরণ চক্র

(Data Processing Cycle)

ডাটাকে অর্থবোধক এবং তথ্যে পরিণত করার জন্য প্রক্রিয়াকরণ এর প্রয়োজন হয়। এ প্রক্রিয়াকরণ নির্দিষ্ট কিছু ধাপে সম্পন্ন হয়, যাকে ডাটা প্রসেসিং সাইকেলের ধাপ বলা হয়। নিম্নে ডাটা প্রসেসিং সাইকেলের ধাপসমূহ দেখানো হলো-

- (ক) ডাটা কালেকশন (Data Collection) : ডাটা প্রসেসিং সাইকেলের প্রথম এবং অত্যন্ত গুরুত্বপূর্ণ ধাপ হলো ডাটা কালেকশন। কারণ সঠিক ফলাফল অনেকটাই নির্ভর করে সঠিক ডাটা কালেক্ট করার উপর।
- (খ) প্রস্তুতকরণ (Preparation) : ডাটা কালেকশনের পর যে ডাটাগুলো ইনপুটে প্রদান করা হবে, সেই ডাটাগুলোকে এই প্রস্তুতকরণ ধাপে প্রস্তুত করা হয়।
- (গ) ইনপুট (Input) : যেসব ডাটা ইনপুট-এর জন্য প্রস্তুত করা হয় সেসব ডাটাকে এ ধাপে মেশিনের বোধগম্যতার জন্য মেশিন কোডে রূপান্তর করা হয়।
- (ঘ) প্রসেসিং (Processing) : ইনপুটকৃত ডাটার মাঝে বিভিন্ন প্রকার রিলেশন, কম্বিনেশন ইত্যাদি সম্পন্ন করে এদেরকে একটি অর্থবোধক তথ্যে রূপান্তর করা হয়।
- (ঙ) আউটপুট (Output) : এ ধাপে অর্থবোধক তথ্য বিভিন্ন প্রকার ব্যবহারকারীর নিকট হস্তান্তর করা হয়।
- (চ) জমা (Storage) : ডাটা, ইনফরমেশন ইত্যাদি ভবিষ্যতে ব্যবহারযোগ্য করে তোলার জন্য এদেরকে এই ধাপে জমা করে রাখা হয়।

স্ট্যান্ডার্ড ডাটা টাইপ উল্লেখকরণ (Mention Standard Data Types)

১.২ স্ট্যান্ডার্ড ডাটা টাইপ উল্লেখকরণ

(Mention Standard Data Types)

বিভিন্ন ধরনের কাজের উপর ভিত্তি করে ডাটার ধরন ও সংরক্ষিত মেমোরি পরিসরের দিক বিবেচনা করে ডাটার শ্রেণিবিভাগ করা হয়েছে। আবার প্রোগ্রামিং-এর ক্ষেত্রে বিভিন্ন ভাষায় বিভিন্ন ধরনের ডাটা ব্যবহৃত হয়। ডাটার প্রকৃতির উপর ভিত্তি করে ডাটা প্রধানত দুই প্রকারের। যথা-

(ক) আলফাবেটিক ডাটা (Alphabetic Data) ও

(খ) নিউমেরিক ডাটা (Numeric Data)। নিম্নে এ দুই ধরনের ডাটা টাইপের বর্ণনা দেয়া হলো

ক)

আলফাবেটিক ডাটা (Alphabetic Data): এক বা একাধিক অক্ষর বা বর্ণের সমন্বয়ে আলফাবেটিক ডাটা তৈরি হয়। যেমন : A, B, a, 2, #, *, %, "Computer", "Dhaka", "Bangladesh" ইত্যাদি।

আলফাবেটিক ডাটাসমূহকে আবার দুই শ্রেণিতে ভাগ করা হয়। যথা-

ক্যারেক্টার (Character): বিভিন্ন প্রতীক, বর্ণ বা ইংরেজি অক্ষরসমূহ এ ধরনের ডাটার অন্তর্গত। যেমন : A, B, a, Z, #, *, % ইত্যাদি স্ট্রিং (String): দুই বা ততোধিক প্রতীক, বর্ণ বা ইংরেজি অক্ষরের সমষ্টিকে স্ট্রিং বা অক্ষরমালা বলা হয়। যেমন :

"Computer", "Dhaka", "Bangladesh" ইত্যাদি। ৩.

নিউমেরিক ডাটা (Numeric Data): পূর্ণ বা দশমিক মানবিশিষ্ট বিভিন্ন ধরনের সাংখ্যিক মানসমূহ নিউমেরিক ডাটার অন্তর্গত। নিউমেরিক ডাটার উদাহরণ হলো : 10, 20, 152. 45, 3.5×10200 অর্থাৎ $3.5E - 200$, $3.5 \times$ (খ)

$10-203$ অর্থাৎ $3.5E - 203$ ইত্যাদি। প্রোগ্রামে এ ধরনের ডাটার ব্যবহারই অধিক। নিউমেরিক ডাটাসমূহকে আবার তিন শ্রেণিতে ভাগ করা হয়। যথা-

i. ইন্টিজার বা পূর্ণমান সংখ্যা।

ii. ফ্লোটিং বা দশমিক সংখ্যা।

iii. ডাবল বা এক্সপোনেনশিয়াল সংখ্যা।

ইন্টিজার ডাটা : পূর্ণমান সংখ্যাসমূহ

এ ধরনের ডাটার অন্তর্গত। যেমন : 100, 200, 300, 1000, 4890, 12345 ইত্যাদি।

ফ্লোটিং ডাটা : দশমিক চিহ্নবিশিষ্ট বিভিন্ন সংখ্যাসমূহ এ ধরনের ডাটার অন্তর্গত। যেমন : 100.0, 20.50, 23.00, 48.90, 123.45 ইত্যাদি।

মেমোরি লোকেশন, অ্যারে এবং লিস্ট (Define Memory Location, Array & List)

মেমোরি লোকেশন (Memory Location) : যে পদ্ধতিতে Computer এর Memory তে সংরক্ষিত উপাদানসমূহকে সহজে পর্যায়ক্রমিক Assess এর মাধ্যমে উপস্থাপন করা হয় তাকে মেমোরি লোকেশন (Memory Location) বলে। নিয়ে Memory Location কে চিত্রের মাধ্যমে দেখানো হলো-

অ্যারে (Array) : অ্যারে অর্থ হলো শ্রেণিবদ্ধ সজ্জা। একই টাইপের একাধিক ডাটাকে শ্রেণিবদ্ধভাবে সজ্জিত করাকে অ্যারে বলে। মেমোরি অ্যারের জন্য পরস্পর সংলগ্ন Byte Allocate করে।

লিস্ট (List) : লিস্ট অর্থ হলো তালিকা। সাধারণত একাধিক ডাটার সংগৃহীত সমাহারকে লিস্ট (List) বলে। এতে সহজেই ডাটা সংযোজন ও বিয়োজন করা যায়।

Temp



	1001
	1002
0	1003
	1004
	1005
	1006

ডাটা স্ট্রাকচার বর্ণনা (State Data Structure)

উপাত্ত সংগঠন বা ডাটা স্ট্রাকচার বলতে ডাটাকে কম্পিউটারের স্মৃতিতে রাখার একটি নির্দিষ্ট উপায়কে বোঝায়, যাতে প্রয়োজনে উপাত্ত/ডাটাকে দক্ষতার সাথে ব্যবহার করা যায়। যন্ত্রের সাথে বাছাই করা ডাটা স্ট্রাকচার, ডাটার উপর সবচেয়ে দক্ষ অ্যালগরিদমের ব্যবহার সম্ভব করে তোলে। একটি সুপরিকল্পিত উপাত্ত সংগঠন মেমোরি ও সময় যথাসম্ভব বাঁচিয়ে ডাটার উপর অনেকগুলো জরুরি অপারেশন প্রয়োগ করার সুযোগ দেয়। কোনো একটি প্রোগ্রামিং ভাষাতে প্রদত্ত ডাটা টাইপ রেফারেন্স ও অপারেশন অনুসারে ডাটা স্ট্রাকচার বাস্তবায়ন করা হয়।

বিভিন্ন ধরনের প্রোগ্রামের ডিজাইনে ডাটা স্ট্রাকচারের বাছাইকরণ একটি অন্যতম প্রধান বিষয়। বড় ধরনের সিস্টেম তৈরির অভিজ্ঞতা থেকে জানা যায় যে, চূড়ান্ত ফলাফলের মান ও সার্থকতা অনেকাংশে শ্রেষ্ঠ ডাটা স্ট্রাকচার বাছাই করার উপর নির্ভরশীল। ডাটা স্ট্রাকচার বাছাই করার পর, অ্যালগরিদমের ব্যবহার অবশ্যম্ভাবী হয়ে ওঠে। কখনো এর বিপরীত পদ্ধতি অবলম্বন করা হয় এবং সে ক্ষেত্রে কোনো নির্দিষ্ট অ্যালগরিদমের জন্য উপযুক্ত ডাটা স্ট্রাকচার বাছাই করা হয়। উভয় ক্ষেত্রেই প্রয়োজনীয় ডাটা স্ট্রাকচার বাছাই করা অত্যন্ত জরুরি। অন্যকথায়- বর্ণ, ফিল্ড, রেকর্ড ইত্যাদি সমন্বয়ে প্রক্রিয়াজাতকরণের উপযোগী উপাত্তকে ডাটা স্ট্রাকচার (Data Structure) বলে।

কোন নির্দিষ্ট ডাটা মডেল (Data Model) কেমন হবে তা দুটি বিষয়ের ওপর নির্ভর করে। যথা- (ক) স্ট্রাকচারের আওতাধীন ডাটাকে এমন বৈশিষ্ট্যের অধিকারী হতে হবে, যাতে বাস্তব জীবনে ডাটাসমূহের মধ্যে প্রকৃত সম্পর্কের প্রতিফলন ঘটে।

(খ) স্ট্রাকচার (Structure) অবশ্যই সাধারণ (Simple) হতে হবে যেন যে কেউ প্রয়োজনবোধে কার্যকরভাবে ডাটা Process করতে পারে।

ডাটা স্ট্রাকচারের প্রকারভেদ বর্ণনা (State Types of Data Structure)

বিভিন্ন ধরনের অ্যাপ্লিকেশনের উপর ভিত্তি করে ডাটা স্ট্রাকচারের শ্রেণিবিভাগ করা হয়েছে। ডাটা স্ট্রাকচারকে প্রধানত দু

ভাগে ভাগ করা যায়, যথা-

- (ক) প্রিমিটিভ ডাটা স্ট্রাকচার (Primitive Data Structure)
 - (খ) নন-প্রিমিটিভ ডাটা স্ট্রাকচার (Non-Primitive Data Structure)
- প্রিমিটিভ ডাটা স্ট্রাকচার আবার চার প্রকার, যথা-

- (i) ইন্টিজার (Integer)
- (ii) ফ্লোট (Float)
- (iii) ক্যারেক্টার (Character)
- (iv) বুলিয়ান (Boolean)।

নন-প্রিমিটিভ ডাটা স্ট্রাকচার আবার দুই ভাগে বিভক্ত, যথা-

- (i) লিনিয়ার ডাটা স্ট্রাকচার (Linear Data Structure)
- (ii) নন-লিনিয়ার ডাটা স্ট্রাকচার (Non-Linear Data Structure)।

লিনিয়ার ডাটা স্ট্রাকচার আবার চার ভাগে বিভক্ত, যথা-

- (i) অ্যারে (Array)
- (ii) লিংক লিস্ট (Link List)
- (iii) স্ট্যাক (Stack)
- (iv) কিউ (Queue)।

নন লিনিয়ার ডাটা স্ট্রাকচার আবার দুই ভাগে বিভক্ত। যথা-

- (i) ট্রি (Tree)
- (ii) গ্রাফ (Graph) ইত্যাদি।

লিনিয়ার ডাটা স্ট্রাকচার (Linear Data Structure)

এ ধরনের ডাটা স্ট্রাকচারের উপাদানসমূহ নির্দিষ্ট নিয়ম নীতি মেনে ক্রমানুসারে সাজানো থাকে। যেখানে ডাটাসমূহ একটি আরেকটির সাথে কানেক্টেড থাকে। খুব দ্রুত ডাটাসমূহ ট্রাভার্স বা অ্যাক্সেস করা যায়। লিনিয়ার ডাটা স্ট্রাকচার আবার ৪ (চার) প্রকার। যথা-

- (i) অ্যারে (Array)
- (ii) লিংক লিস্ট (Link List)
- (iii) স্ট্যাক (Stack)
- (iv) কিউ (Queue)

অ্যারে (Array): একই ডাটা টাইপের একাধিক ভেরিয়েবল এর সমষ্টি বা সংগ্রহকে অ্যারে বলে। অ্যারে যে কয়টি Element নিয়ে গঠিত তাকে অ্যারের দৈর্ঘ্য বা সাইজ বলা হয়। অ্যারে যদি n সংখ্যা Element নিয়ে গঠিত হয় তবে 1, 2, 3..... সাবস্ক্রিপ্ট দ্বারা আলাদা আলাদা element কে প্রকাশ করা হয়। মনে করি, X অ্যারের n সংখ্যক element আছে, তবে এদের উপাদান (element) গুলোকে X , X_2 X_n

বা, $X[1], X[2], \dots, X[n]$ দ্বারা প্রকাশ করা হয়।

লিংক (Link): লিংক (Link) অর্থ হলো তালিকা। তবে লি লিস্ট এ কি যার ডাটা উপাদানগুলোর মধ্যে এক বা একাধিক সম্পর্ক তৈরি করা যায়। লি লি অ্যাবেতে উপাদানের সন্নিবেশ। জারের এক একটি উপাদানকে বলা হয় নোড।

স্ট্যাক (Stack): স্ট্যাক হলো একটি বিশেষ ধরনের ডাটা স্ট্রাকচার, যাতে ডাটা উপাদানগুলো একটি তর আকারে স হয়। স্ট্যাকে নতুন ডাটা সংরক্ষণ অথবা সংরক্ষিত ডাটা অপসারণ শুধুমাত্র উপর থেকে করা সম্ভব। এজন্য স্ট্যাককে LIFO (Last-in First-out) ধরণের ডাটা স্ট্রাকচার বলা হয়।

প্লেটের জন্ম, আলমারিতে রাখা কাপড়ের সূপ, ইটের ভাটায় রাখা রাখা ইটের ওম ইত্যাদি স্ট্যাক এর উদাহরণ।

নন-লিনিয়ার ডাটা স্ট্রাকচার (Non-Linear Data Structure)

এ ধরনের ডাটা স্ট্রাকচারের উপাদানগুলো নির্দিষ্ট ক্রমানুসারে না থেকে এলোমেলোভাবে সজ্জিত থাকে।

পিনিয়নের ডাটা স্ট্রাকচার আবার দুই ধরনের

(+) (Tree) এবং

(খ) (Graph)।

S (Tree) Tree এক ধরনের Non Linear Data Structure অনেক সময় ডাটার বিভিন্ন উপাদানের মধ্যে উর্ধ্বক্রম

বা Hierarchical relationship বিদ্যমান থাকে। এখানে Hierarchical বলতে একটি ডাটার অধীনে সম্পর্কযুক্ত

একাধিক ডাটা উপাদানকে বুঝানো হয়েছে। যে Data Structure-এর মধ্যে এ ধরনের Hierarchical সম্পর্কের সুস্পষ্ট প্রতিফলন ঘটে, তাকে Tree বলে। Tree একটি Non Linear ধরনের Data Structure: একটি গাছের যেমন মূল বা কাজ থেকে একাধিক শাখা বের হয়, আবার প্রতিটি শাখা থেকে একাধিক প্রশাখা বের হয় এবং এভাবেই একটি গাছ বিস্তার করে তেমনই Tree Data Structure-এ একটি মূল (Root) থাকে। একটি মূলের (Root) কয়েকটি Group থাকে। প্রতিটি Group-এর আবার কয়েকটি sub Group থাকে এবং এভাবেই Tree Structure

Bird-কে আমরা দুই প্রকারে ভাগ করতে পারি। উভয় প্রকারের Bird-এর অনেক উদাহরণ রয়েছে। Bird-এর Classification -কে আমরা Tree-এর মাধ্যমে প্রকাশ করতে পারি। কোন কিছুকে Tree আকারে প্রকাশ করলে বুঝতে সুবিধা হয় এবং প্রয়োজনীয় data নিয়ে কাজ করতে বা Process করতে

সুবিধা হয়।

স্ট্যাটিক এবং ডাইনামিক মেমোরি অ্যালোকেশন বর্ণনা (State Static and Dynamic Memory Allocation)

Language-এর ওপর ভিত্তি করে Array-এর দুইভাবে মেমোরি Space নির্দিষ্ট (Allocate) হয়।

(i) Static Memory Allocation.

(ii) Dynamic Memory Allocation

স্ট্যাটিক মেমোরি অ্যালোকেশন (Static Memory Allocation) : কোন কোন Language যেমন Pascal এবং Fortran-4 Compile করার সময় Statically মেমোরি Allocate হয়। এ কারণে Program নির্বাহের সময় অ্যারের সাইজ অপরিবর্তিত (Fixed) থাকে। এই পদ্ধতিকে স্ট্যাটিক মেমোরি অ্যালোকেশন (Static Memory Allocation) বলা হয়।

ডাইনামিক মেমোরি অ্যালোকেশন (Dynamic Memory Allocation) : কিছু কিছু Language যেমন- C-তে মোট

কয়টি ডাটা সংরক্ষণ করতে হবে তা পূর্বেই জানিয়ে দেয়ার বাধ্যবাধকতা নেই। প্রোগ্রাম চলাকালীন সময় ডাটা সংরক্ষণ ও ব্যবহারের জন্য নিজেদের প্রয়োজন অনুযায়ী মেমোরি Allocate হয়। এই পদ্ধতিকে ডাইনামিক মেমোরি অ্যালোকেশন (Dynamic Memory Allocation) বলে।



END