Unit IV Predicting Human Behavior and Privacy Issues.

Understanding and Predicting Human Behavior for social communities, User data Management, Inference and Distribution, Enabling new human experiences, Reality Mining, Context Awareness, Privacy in Online social Networks, Trust In online Environment, What is Neo4j, Nodes, Relationship, Properties.

4.1 Understanding and Predicting Human Behavior for Social Communities:-

⇒ Over the last years, with the express development in information technology, people take the advantages of the devices and services to remain "connected" and continuously enjoy the activity because they use for the sports, entertainment or work.

⇒ Ubiquitous computing environment will allow everyone permanent access to the Internet, anytime anywhere and anyhow.

⇒ Evolution of services, social aspects remain in the roots of every human behavior and activity.

⇒ Such phenomena used in Online social N/w

⇒ User Context help us to understand users in a better way.

* Quality of Experience (QoE)
  - ↳ User's Internal State (eg) Predispositions, Expectations, needs, motivation, mood)
* Characteristics of the designed System
  - ↳ (Eg Usuability, functionality, relevance ~~motivation, mood~~)
* Content (Environment)

⇒ Way to understand and predict human behaviour in future social-aware multimedia systems.

4-2 User Data Management - Inference and Distribution.

User data Management:-

It encompasses the collection, storage, processing and protection of information provided by users.

Inferences:- Conclusion reached on the basis of evidence and reasoning.

It involves using algorithms and analysis techniques to extract meaningful insights or predictions from this data, helping businesses make informed decisions or personalize user experiences.

Distribution

It refers to disseminating these insights or utilizing them across various platforms or channels, such as targeted advertising, recommendation systems, or product customization. It's essential to prioritize data privacy, security, and compliance with regulation to maintain trust and integrity in user data management.

⇒ Uses use the heterogeneity in devices, technologies so the need for the standardization of user related data and the architecture in user profile management systems.

Some standards.

* European Telecommunications Standards Institute (ETSI),
* Third Generation Partnership Project (3GPP),
* Open Mobile Alliance (OMA)

3GPP → as a framework for streamlining service independent user data and storing it under a single logical structure in order to avoid duplications and data inconsistency.

⇒ Logically centralized data storage mapped into physically distributed configurations & should allow data to be accessed in a standard format.

Several approaches for Interoperability.

1) Syntactic 2) Semantic 3) & Modelling approaches.

⇒ To improve the degree of service personalization, it is important to generate new information from the

existing ones. (5)

⇒ Social Network, User modelling, and reality Mining techniques can be used to study patterns, and predict future behaviors.

⇒ Sensitiveness of the information

↳ stored, accessed and distributed, preserving user's privacy, security and trust.

⇒ Inferring User's needs, desires or intentions, several research initiatives from different fields (eg eHealth, Marketing, Telecoms) are starting to become a reality.

⇒ Different methodologies and approaches, the user requirements and the user ~~requirements~~ and the technologies to address the ~~same~~ problems.

→ Commonly Social N/w analysis, Content-awarness used.

⇒ From the online environment various amount of data collected, In analysis of social n/w data, for Large dataset, cannot be handled by traditional methods.

So in this complex situations, the concept of

decomposition, Additional similar perspective is called Layered Reasoning.

In this,

First stage → Feature Extraction & grouping.

Second ~~st~~event → state & activity Recognition.

Last stage → Prediction and Inference of new knowledge.

Concern Social Networks.

⇒ Research Usually focuses on quantifying or qualifying the relationship between peers,

Centrality → Importance of a node in a n/w.

Clustering & classification → Applied to similarity computation

When user related data is associated with time and space ~~complent~~ dimensions,

↳ data mining techniques to find the hidden patterns.

Combining all these Ontology and semantic technologies. → We present a generic framework for managing user related data.

↳ It provides the way for Understanding and predicting future human behaviour with social communities.

4-3 Enabling New Human Experiences :-

↳ How the challenges related with user data management and new knowledge inference are dealt, it is important to understand what are the technologies behind it, how to link them and what can they achieve when combined in synergy.

The Technologies :-

Social Networks :-

↳ Humans in all cultures at all times form complex social networks, The term social network here means ongoing relations among people that matter to those engaged in the group,

⇒ either for specific reasons or for more
⇒ general expressions.

↳ Likewise, social networks among individuals ~~may not be related can be validated~~ can be maintained by agreement on objects or even by choice of entertainment or self Interest Based.

↳ Usually, network members tend to trust with each other to provide information with other.

↳ Social networks are trusted because of shared experiences and the perception of shared values and shared needs. This phenomenon has recently created and converted existing online communities into complex online social networks.

Although the behavior of individuals in online networks can be slightly different from the same individuals interacting in a more ~~traid~~ traditional social networks.

4.4. Reality Mining :-

⇒ Reality Mining is the practice of gathering and analyzing data from everyday activities to understand patterns and behaviors.

⇒ It often involves collecting information from sources like mobile phones, social media, and sensors to gain insights into human behavior, social interactions and environmental factors.

⇒ It's used in various fields such as urban planning, healthcare and marketing to inform decision-making and improve services.

For eg A doctor could use data from a patient's fitness tracker to track their recovery after surgery or monitor their progress in managing a chronic condition like diabetes.

⇒ ~~These~~ This examples illustrate how reality mining can be applied in different domains to gain insights and improve decision-making processes.

4.5 Content Awarness :-

↳ Content Awarness is the ability of a system or device to perceive and understand its environment including aspects such as Location, time, user activity and surrounding conditions.

↳ This capability allows the system to adapt its behavior or provide relevant information based on the current content.

↳ Content is available, meaningful and carries rich information in such environment.

↳ User expectations, and user experience is directly related to content.

Decision making.

↳ Based on the interpreted content, the system makes decisions on hoco to respond or adapt. This could include adjusting settings, providing relevant information or initiating actions

Healthcare :-

Wearable devices and health apps use content awarness to monitor user activity, track fitness goals and provide timely reminders or

Security ;

⇒ In addition to social network security, Context awarness is crucial for detecting anomalies and potential security threats in various systems including computer n/w and IoT devices.

⇒ By the use of content awarness, systems can provide more relevant, timely and user friendly experience across a wide range of applications and industries.

4.6 Privacy In Online Social Network :-

⇒ Privacy In Online social networks is a critical concern. It involves protecting users personal Information from unauthorized access, ensuring control over what data is shared, and safeguarding against misuse of that data by third parties or the platform itself.

⇒ Users should empowered with privacy settings to manage their information and visibility to others Privacy In Online social networks, Important to consider various aspects.

1. Data Collection :-

⇒ Social networks collect vast amount of data about their users, including personal information, browsing habits, and interactions. Users should be aware of what data is being collected and have control over what they share. b.

2. Privacy Settings :-

⇒ Platforms should provide robust privacy settings that allow users to customize who can see their posts, photos and personal Information. This includes options to limit visibility to friends specific groups, or the post public.

3. Data Security :-

⇒ Social networks must Implement strong security measures to protect users data from unauthorized access,

hacking, and data breaches. This involves encryption, secure authentication methods, and regular security audits.

Transparancy :-

⇒ Social networks should be transparent about their data practices, including how they collect, use and share user's information. This helps users make informed decisions about their privacy.

Data Portability :-

⇒ Users should have the ability to easily export their data from social networks and move it to other platforms if they choose to do so. This promotes competition and gives users more control over their information.

Overall, privacy in online social network requires a balance between providing personalised experiences and respecting users' privacy rights.

Online Social Networks :-

⇒ Professor J.A. Barnes has introduced the term "social network" in 1967 to describe the associations of people by drawn together by family, work, hobby etc. for emotional, instrumental, appraisal, and information support.

⇒ Network may operate in many levels from family level to a level of nations and can play the important roles in communication among people, organizations and even nations as well as the way how problems are solved and how organizations may run in a better way.

⇒ In its simplest form, a social network is a map of the relevant ties between the individuals, organizations, nations etc.

⇒ By the evolution of Internet, provides a greater scope of implementing social networks online.

⇒ Online social networks have broader and easier coverage of members to share information and resources

⇒ The first online social Networks called Usenet Newsgroup (www.usenet.com) designed and built by Duke University graduate students Tom Truscott and Jim Ellis in 1979. Since then the Online social networks have a continuous growth in size and Numbers

⇒ In February 2010, Online social Network giant Facebook cross the massive 370 million registered

monthly active Users. (14)

shows a Brief Timeline of the History of online social networking.

1971 Ray Tomlinson invents email.

1973 First group chat program.

1975 First mailing list, called Msggroup.
First Computer Conferencing system.

1978 First Multi-user Dungeon (MUD) for Multi-user gaming.

1979 USENET Newsgroup Created.

1984 Birth of the Fido Network of Bulletin Board Systems (BBSes)

1985 Whole Earth Letronic Link (WELL) community begin

1988 Internet Relay chat (IRC) Invented.

1991 TimeBerners-Lee posts "World-Wide Web: Executive Summary" to USENET group

1992 Berners Lee created the What's New? page, arguable the first blog.

. . .

2005 Skype hits million downloads.

. . .

2009 Facebook hits 350 million registered Users.

A January 2009 Compete.com study ranked Facebook as the most used social Network by worldwide monthly active Users, followed by Myspace.

Rank	Site	Monthly Visit
1	facebook.com	1,191,373,339
2	myspace.com	810,153,536
3	twitter.com	54,218,781
4	flixter.com	53,389,974
86	tagged.com	39,630,927
5	linked.com	42,744,438
7	classmates.com	35,219,210
8	myyearbook.com	33,121,821
9	livejournal.com	25,221,854
10	imeen.com	22,993,608

Top ten mostly visited social Networks in Jan 2009 - Based on MAU.

Country	1/1/2010	1/2/2010	change	change (%)
U.S	102,681,240	108,062,900	5,381,666	5
Indonesia	15,301,280	17,301,760	2,000,480	13
Turkey	16,961,140	18,556,840	1,595,700	9
U.K	23,076,700	24,342,820	1266,120	5
France	14,301,020	15,498,220	1,197,200	8
Mexico	6,671,560	7,624,120	952,560	14
Germany	5,796,940	6,674,740	877,800	15
India	5,658,080	6342,800	684,720	12
Philippines	8,806,300	9317,180	510,880	6
Brazil	2,373,520	2869,920	496,400	21

Country wise Monthy growth of Facebook Users - as on Feb 2010

Once Facebook began offering the service in multiple Languages (its available in more than 70 of them as of today), it started blowing up in many countries like, Canada, Iceland, Norway, South Africa, Chile etc.

No	Network Name	Web Link	Reg User.
1	Facebook	www.facebook.com	350,000,000
2	Qzone (Chinese	http://qzone.qq.com	200,000,000
3.	Myspace	www.myspace.com	130,000,000
4.	Windows Live Spaces	http://home.spaces.live.com	120,000,000
5.	Habbo	www.habbo.com.au	117.000,000
6.	Orkut	www.orkut.com	100,000,000
7.	Friendstar	www.friendster.com	90,000,000
8.	Hi5	www.hi5.com	80,000,000
9.	Flixster	www.flixster	63,000,000
10	Netlog	www.netlog.com	59,000,000.

Top Ten Largest Social networks in Feb'10, Based on Registered Users.

Growth Comparison MAU of top ten countries between January and February 2010.

Trust In online Environment

4.1

⇒ Trust has become important topic of research in many fields including sociology, psychology, philosophy, economics, business, Law and IT. It is not a new topic to discuss.

⇒ Over a Long period of time, "trust" word was discussed in scholarly articles and several books.

⇒ Trust is a complex word with multiple dimensions.

⇒ Dozens of proposed definitions are available in the literature, a complete formal unambiguous definition of trust is rare.

Two generalized definitions of trust is defined by which they called

* Reliability Trust (Evaluation Trust)
* Decision Trust

Reliability Trust (Evaluation Trust):-

↳ It means reliability of something or somebody. It can be defined as the subjective probability by which an individual, A,

expects that another individual, B performs a given action on which its welfare depends. On the other hand,

2) decision Trust

↳ It is defined as which one party is willing to depend on some thing or somebody in a given situation with a feeling of relative security, even though negative consequences are possible.

4.8 Neo4j

Neo4j is a NoSQL database. (→ non relational, distributed, flexible & scalable.) It is highly scalable and schema-free.

It is a world most popular graph database management system.

Neo4j was developed by Neo Technology and called an ACID complaint transactional database with native graph storage and Processing.

Neo4j Working :-

Neo4j stores and displays data in the form of graph. In neo4j, data is represented by nodes and relationship between those nodes.

Neo4j databases → Graph database doesn't use the tables, rows or columns to store or present data. But in Relational databases use (Tables, rows, colum to store datas →

Neo4j is best for storing data that has many interconnecting relationship. So it is better when dealing with relational databases.

It is doesn't require Predefined Schema.

Advantages of Neo4j :-

1. Flexible data Model :-

Neo4j provides a flexible simple and yet powerful data model, which can be easily changed according to the applications and industries.

Real - time Insights :-

Neo4j provides results based on real - time data.

High availability :-

Neo4j is highly available for large enterprise real - time applications with transactional guarantees.

Connected and semi - structures data :-

Using Neo4j, you can easily represent connected and semi - structured data.

Easy Retrieval :-

Using Neo4j, you can not only represent but also easily retrieves (traverse / navigate) connected data faster when compared to other databases.

Cypher query Language :-

→ Declaratives query language Using ascii - art syntax.

The commands of this language are in human readable format are very easy to Learn.

No joins :-

Using neo4j, it does NOT require complex joins to retrieve connected / related data as it is very easy to retrieve its adjacent node or relationship details without joins or indexes.

Features of Neo4j : (21)

1. Data Model : (Flexible Schema) :-

Neo4j follows a data model named native property graph model (Nodes, Relationship, Properties)
entities
Key value pairs.

In neo4j no need to follow a fixed schema. You can add or remove properties as per requirement. It also provides schema constraints

2. ACID Properties : -

Neo4j supports full ACID (Atomicity, Consistency, Isolation, and Durability) rules.

3. Scalability and Reliability : -

Neo4j provides increasing the numbes of reads/writes
Neo4j alsop provides support for replication for data safety and reliability.

4. Cypher Query Language : -

Neo4j also provides powerful declarative query Language known as cypher. It uses ASCII - art for depicting graphs. Cypher is easy to learn and can be used to create and retriere relations between data without using the complex ~~qu~~ queries like joins.

Built In Web applications :

Neo4j provides a built-in Neo4j browser web applications. Using this you can create and query your graph data

Drivers:-

Neo4j can work with,

* REST API to work with programming Languages such as Java, spring, scala etc.
* Java script to work with UI MVC such as Node JS
* It support two kinds of Java API.
  1) cypher API
  2) Native Java API

  } to develop Java developers.

Indexing:-

Neo4j supports Indexes by using Apache Lucence

What is Graph?

⇒ A graph is a pictorial representation of objects which are connected by some pair of lines.

⇒ A Graph contains two elements : 1) Nodes (vertices) and 2) Relationship (Edges)

Graph Database :- (It is simple compared to other databases)

⇒ A graph database is a database which is used to model the data in the form of graph It store any kind of data using,

* Nodes
* Relationship
* Properties.

Nodes :-

⇒ Nodes are the records / data in graph databases. Data is stored as properties and properties are simple name / value pairs.

⇒ Nodes can be grouped together by applying a Label to each member. A node can have zero or more Labels.

⇒ Labels do not any properties. Storing data in Neo4j is similar to add more records in other databases.

Relationship :-

It is used to connect nodes. It specifies how the nodes are related

* Relationship always have direction
* Relationship always have a type.
* Relationship form patterns of data.

Properties:-

Properties are named data values.

Popular Graph Databases:-

Neo4j is the most popular Graph Database.

Other Graph Databases are,

* Oracle NoSQL Database.
* OrientDB
* HyperGraphDB
* GraphBase
* Infinite Graph
* Allegrograph etc.

4.9 Neo4j Nodes, Relationship & Properties.

Neo4j Create Nodes :-

Node is a data or record in a graph database. In Neo4j, the CREATE statement is used to create a node. You can create the following things by using CREATE statement.

* Create a single node.
* Create multiple nodes.
* Create a node with Label.
* Create a node with multiple Labels.
* Create a node with properties.
* Returning the creating node.

(i) Create a single Node :-

To create a single node in Neo4j, specify the name of the node along with CREATE statement.

Syntax.

```
CREATE (node_name);
```

↳ optional.

Eg `CREATE (single);`

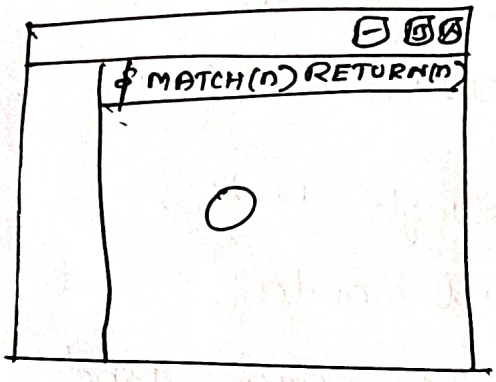
↳ You can see that a node is created.

Verification :- created 1 node, completed after 1ms.

Execute the following code to verify the creation of the node type :.

```
MATCH (n) RETURN (n)
```

After Running .



(ii) CREATE Multiple Nodes : -

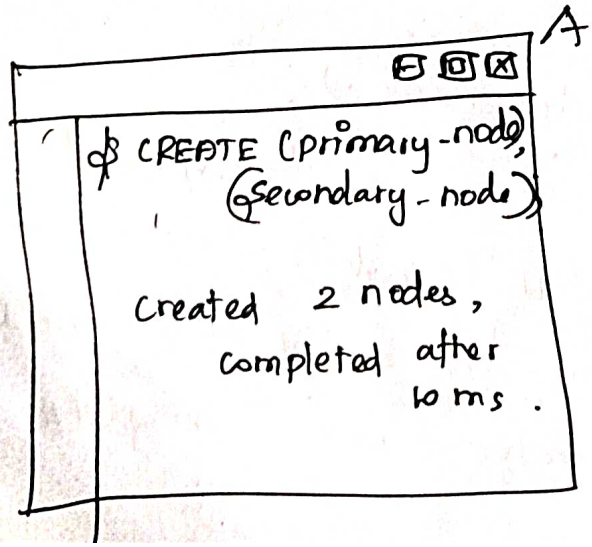
To create multiple nodes in Neo4j, Use CREATE statement with the name of nodes separated by a comma .

Syntax .

```
CREATE (node1), (node2) (node 1) ??? ...
```

Eg) Create 2 nodes : (primary - node), (secondary - node);

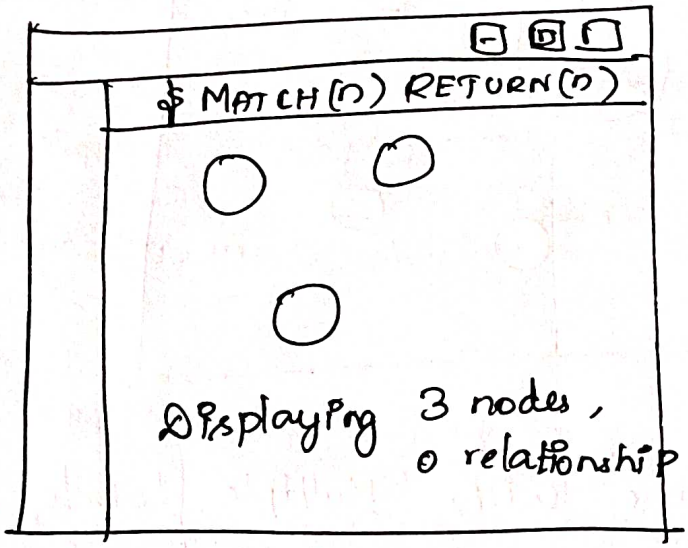
After running .



Verification :-

Match (n) RETURN (n)

After running.



(iii) ~~CRE~~ Create a Node with a Label :-

In neo4j, a Label is used to ~~create~~ classify the nodes using Labels.

CREATE statement is used to create a Label for a node in neo4j

Syntax

CREATE ($node : label)

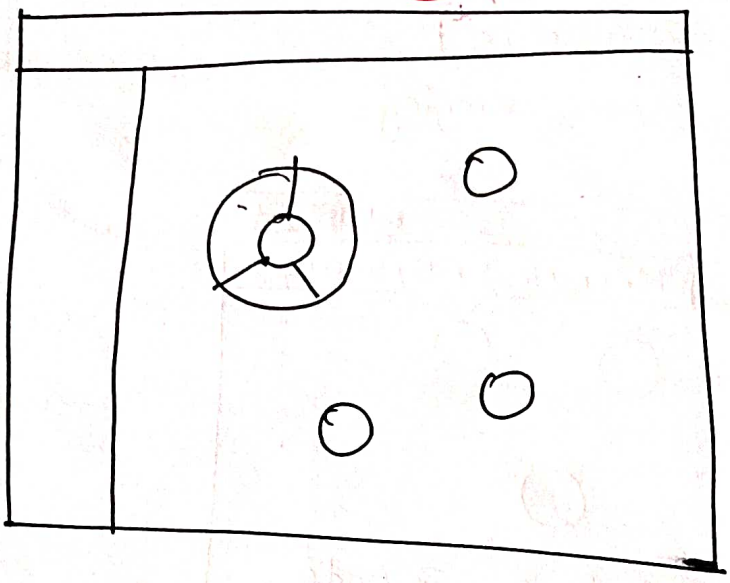
Eg)

CREATE (Kalam : Scientists)

↳ Added 1 label, created 1 node, completed after 174 ms.

Verification.

MATCH (n) RETURN (n)



(iii) Create a Node with Multiple Labels :-

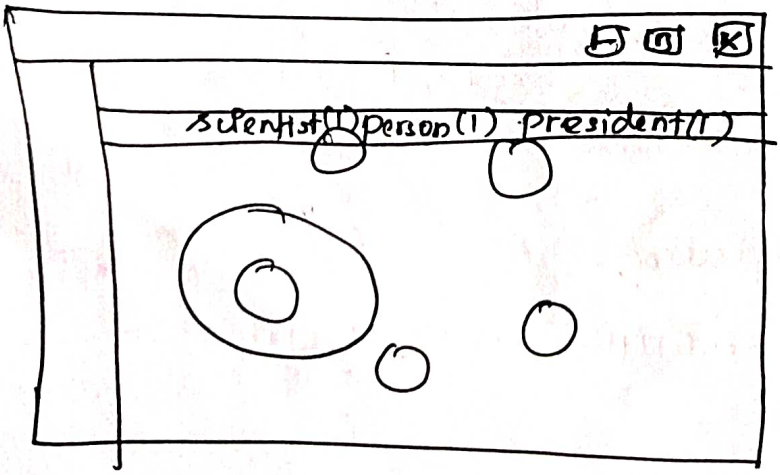
To create multiple labels with a single node, you have to specify the Labels for the node by separating them with a colon.

Syntax.

CREATE (node : label1 : label2 : .... labeln)

Eg) create a node "Kalam" with Label "person", "president" and "scientist"

CREATE (Kalam : person : president : scientist)



(V) create a Node with Properties :-

In Neo4j, properties are the key value pairs which are used by nodes to store data.

CREATE statement is used to create nodes with properties, you can just have to specify these properties separated by commas within curely braces " { } "

Syntax :

CREATE (node : Label { key1 : value , key2 : value .... })

Eg) lets create a node "Ajeet", having the following properties.

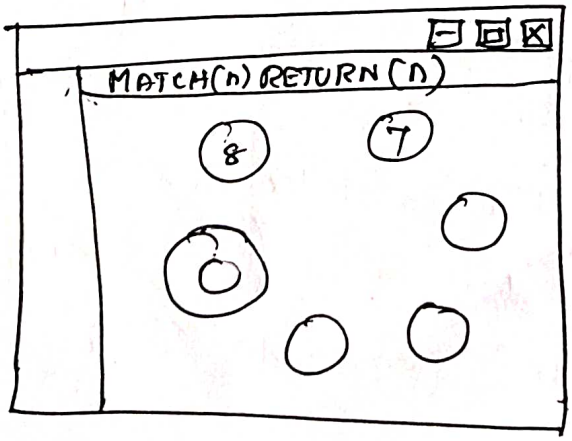
CREATE (Ajeet : Developer { name : "Ajeet kumar", YOB : 1989, POB : "Mau" })

After Running .

Added 1 Label, created 1 node, set 3 properties, completed after 605 ms.

Verification :-

MATCH(n) RETURN (n)



(vi) Returning the created Node :-

→ MATCH (n) RETURN (n) command is used to view the created nodes. This query returns all the existing nodes in the database.

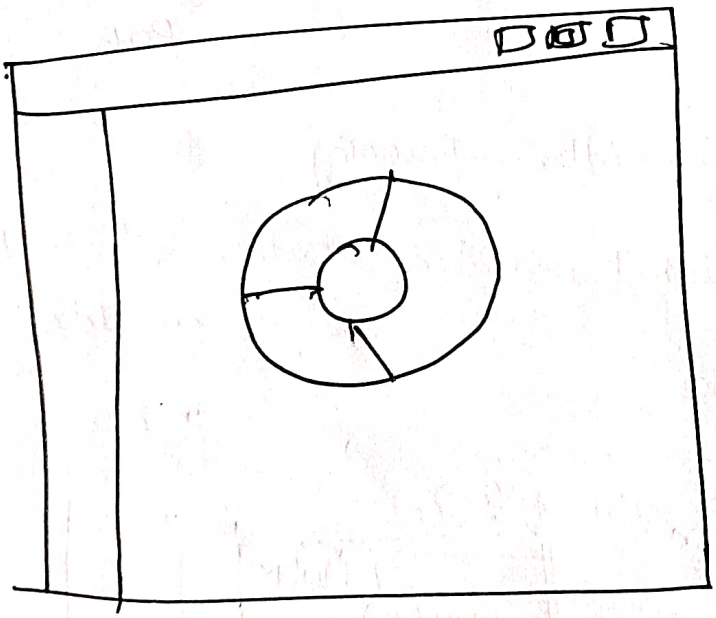
⇒ But if you want to return the newly created node use the RETURN command with CREATE command.

Syntax :-

CREATE (Node : Label { properties .... }) RETURN node.

Eg)

CREATE ($onoo : trainer { trainer: "sonoo Jaiswal", YOB: 1987, POB: "Faizabad" }) RETURN $onoo.



(i) Delete a node :-

In neo4j ~~Delete~~ DELETE statement is always used with MATCH statement to delete whatever data is matched.

The DELETE command is used in the same place.

we use the RETURN clause in our previous eg.

Eg) MATCH (Kohli : person { name : virat kohli" }) DELETE Kohli.

(ii) Delete Multiple Nodes :-

You can delete multiple nodes by using MATCH and DELETE commands in a single statement.

Eg MATCH (a : student { name : "chris Grey" }),
b : Employee { name : "Mark Twin" })

DELETE a, b

(iii) Delete all nodes :- To delete all nodes from the database.

MATCH (n) DELETE (n)

(iv) create Relationship :-

CREATE statement is used to create relationship between nodes. These relationship define direction, type and form patterns of the data.

It defines mainly three things.

* Creating Relationship.
* Creating Relationship between existing nodes.
* Creating Relationship with label and properties.

(iv) Creating Relationship :-

While creating a relationship, relationship should be specified within square braces "[ ]", depending on the direction of the relationship it is

placed between hyphen "-" and arrow "->" as shown below,

Syntax:

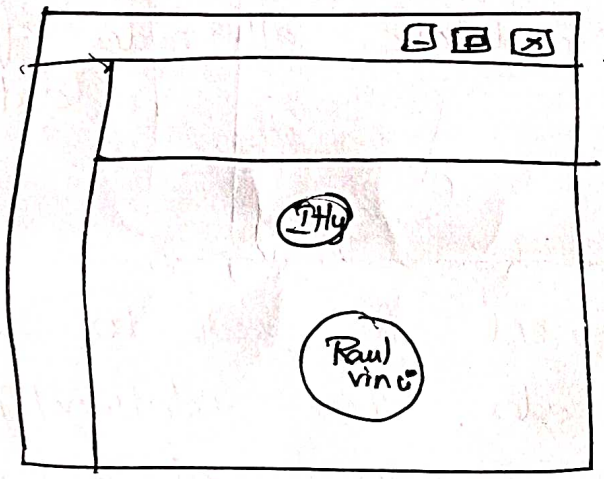
CREATE (node1) - [: RelationshipType] -> (node2)

Eg)

CREATE (Raul: player { name : "Raul Vinci", YOB : 1973, POB: "Milan" })

CREATE (It : country { name : "Itly" })

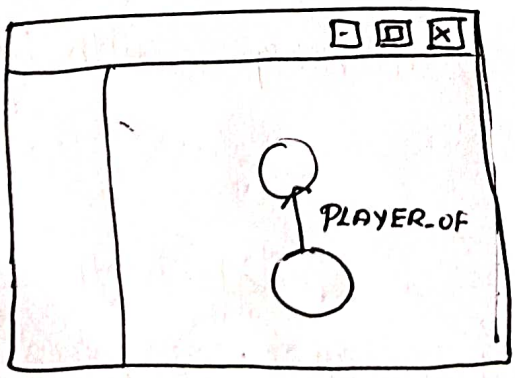
RETURN Paul, It.



Now create a relationship "PLAYER_OF" between these two nodes.

CREATE (Paul) - [r: PLAYER_OF] -> (It)

RETURN Raul, It



Create a relationship between existing nodes.

MATCH statement is used to create relationship between the existing nodes.

Syntax:-

```
MATCH : (a : labelofnode1),(b : labelofnode2)
WHERE a.name = "nameofnode1" AND b.name = "nameofnode
CREATE (a)-[: Relation] → (b)
RETURN a, b
```

Eg)

```
MATCH (a.Player), (b: country) WHERE a.name =
        "Raul vinci" AND b.name = "Ptly"
CREATE (a) - [r : FOOTBALLER_OF] → b
RETURN a, b
```

Create a Relationship with label and properties.

⇒ CREATE statement is used to create a relationship with label and properties.

syntax:-

```
CREATE (node1) - [Label : Rel_Type {key1 : value1 , key2 : value2
                                    ..... n}] → (node2)
```

Eg)

```
CREATE (Kohli: person : player)
```

Then create some properties with same node.

```
CREATE (kohli : player {name : "virate kohli" , yoB : 1988 ,
                        POB : "Delhi" })
RETURN Kohli
```

CREATE (Ind : Country {name : "India"})

Now create a relationship with label and properties.

MATCH (a: Player), (b: country) WHERE a.name =

"Virat Kohli" AND b.name = "India"

CREATE (a) - [r: BATSMAN - OF {Matches : 5 , Avg: 90.75}]

→(b)

RETURN a, b