Neo4j and Cypher Query Language Introduction

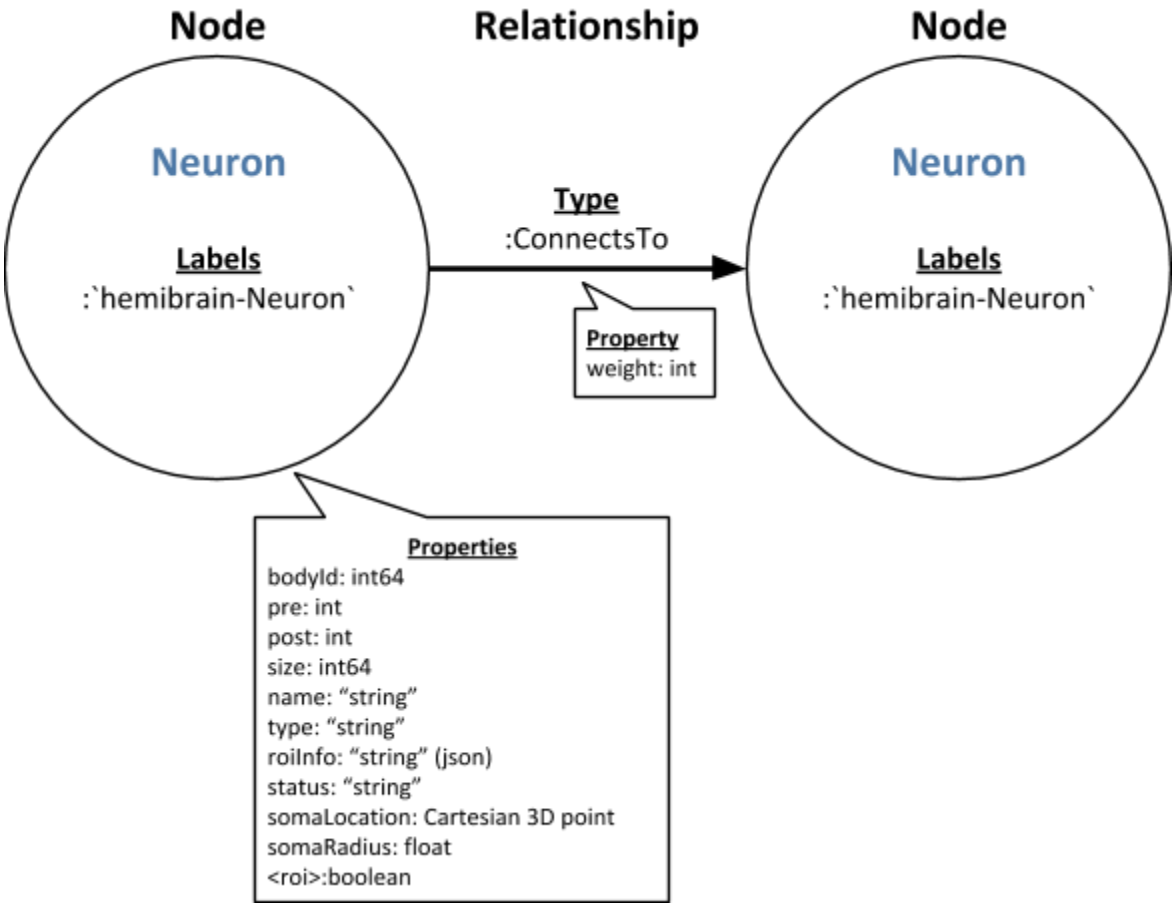
An introduction to the neuPrint data model and writing custom Cypher queries.
See Graph Data Model on neuPrint Explorer for the complete data model. Cypher documentation can be found here: https://neo4j.com/docs/cypher-manual/current/.

Neo4j is a graph database
Nodes are entities in the graph
- Can hold properties
- Can be tagged with labels

Relationships are connections between nodes
- Have a type, direction, start and end node
- Can hold properties

Properties are key-value pairs
- Ex. Key = bodyId, Value = 5813063587

Neuron nodes connect to other Neuron nodes
Neuron nodes hold neuron properties (body Id, name, status, etc.)



Cypher Query Language (CQL) is Neo4j's query language

Clauses

MATCH - search the data with specified pattern

WHERE - add contents to queries

RETURN - define what to include in query result set

EXAMPLE QUERIES

Returns number of presynaptic (n.pre) and postsynaptic (n.post) sites on a neuron (n)

```
MATCH (n :`hemibrain-Neuron`{bodyId:5813063587})
RETURN n.pre, n.post
```

Returns body ID, number of presynaptic (n.pre), and postsynaptic (n.post) sites for a list of neurons

```
MATCH (n :`hemibrain-Neuron`)
WHERE n.bodyId IN [5813063587, 514850616]
RETURN n.bodyId, n.pre, n.post
```

Returns output body IDs and weight of connections for a neuron (n)

```
MATCH (n :`hemibrain-Neuron`{bodyId:5813063587})-[w:ConnectsTo]->(output)
RETURN output.bodyId, w.weight
```

Returns input body IDs and weight of connections for a neuron (n) ordered by weight (w)

```
MATCH (n :`hemibrain-Neuron`{bodyId:5813063587})<-[w:ConnectsTo]-(input)
RETURN input.bodyId, w.weight ORDER BY w.weight DESC
```

Returns output body IDs and weight of connections for a list of neurons

```
MATCH (n :`hemibrain-Neuron`)-[w:ConnectsTo]->(output)
WHERE n.bodyId IN [5813063587, 514850616]
RETURN n.bodyId, output.bodyId, w.weight
```

Returns body IDs and connection weights of inputs that are in the EB ROI for a neuron (n)

```
MATCH (n:`hemibrain-Neuron`{bodyId:5813063587})<-[w:ConnectsTo]-(input{`EB`:true})
RETURN input.bodyId, w.weight
```

Returns body IDs and connection weights of inputs that are in the EB and the FB ROIs for a neuron (n)

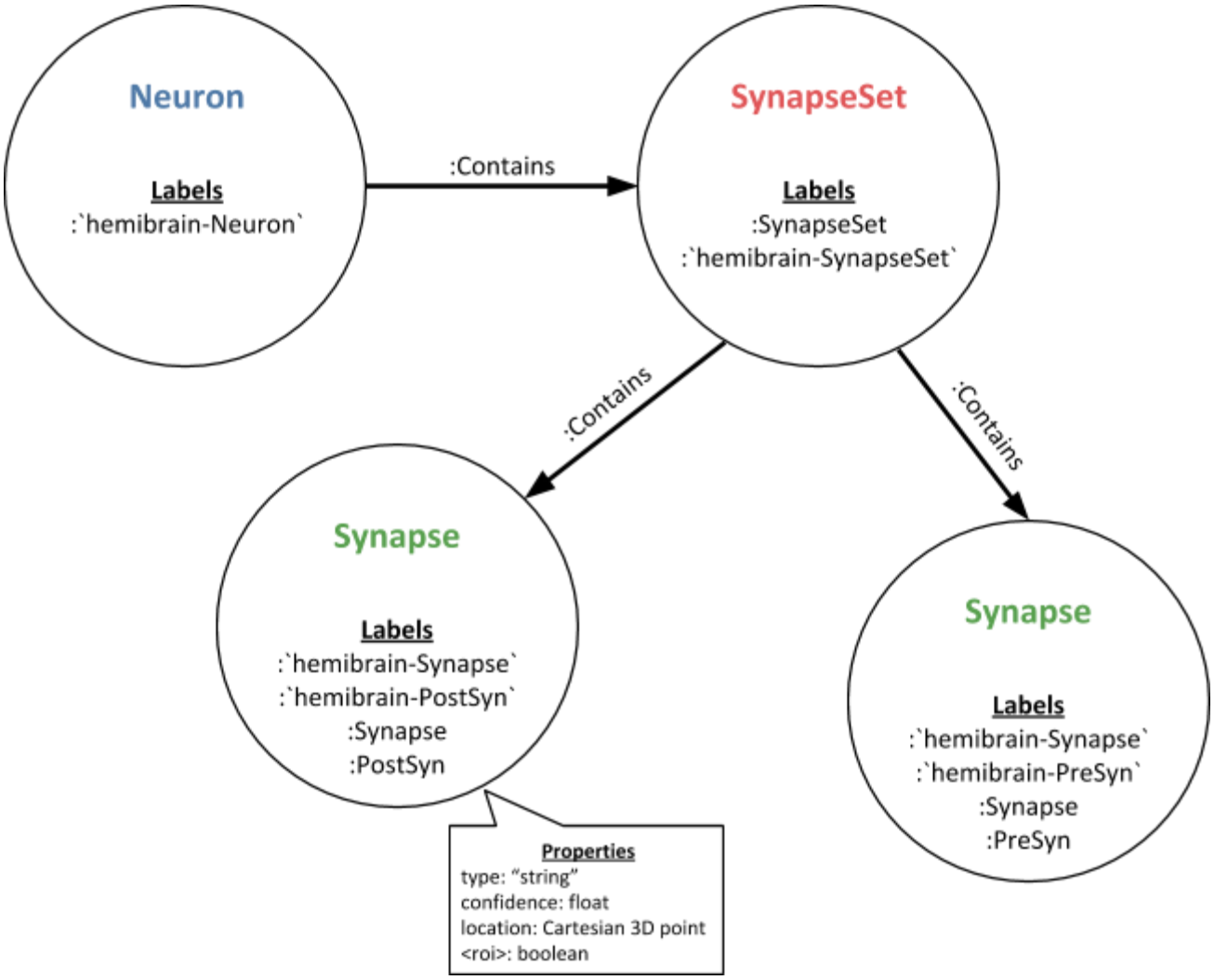
```
MATCH (n:`hemibrain-Neuron`{bodyId:5813063587})<-[w:ConnectsTo]-(input) WHERE input.`EB` AND
input.`FB`
RETURN input.bodyId, w.weight
```

Returns all named neurons (excludes automatically generated names, which end with "*")

```
MATCH (n:`hemibrain-Neuron`)
WHERE NOT n.name ENDS WITH "*" RETURN n.bodyId, n.name
```

Neuron nodes contain synapse set nodes which contain all the synapse nodes for that neuron.

Synapse nodes hold synapse properties (location, roi, etc.)



EXAMPLE QUERIES

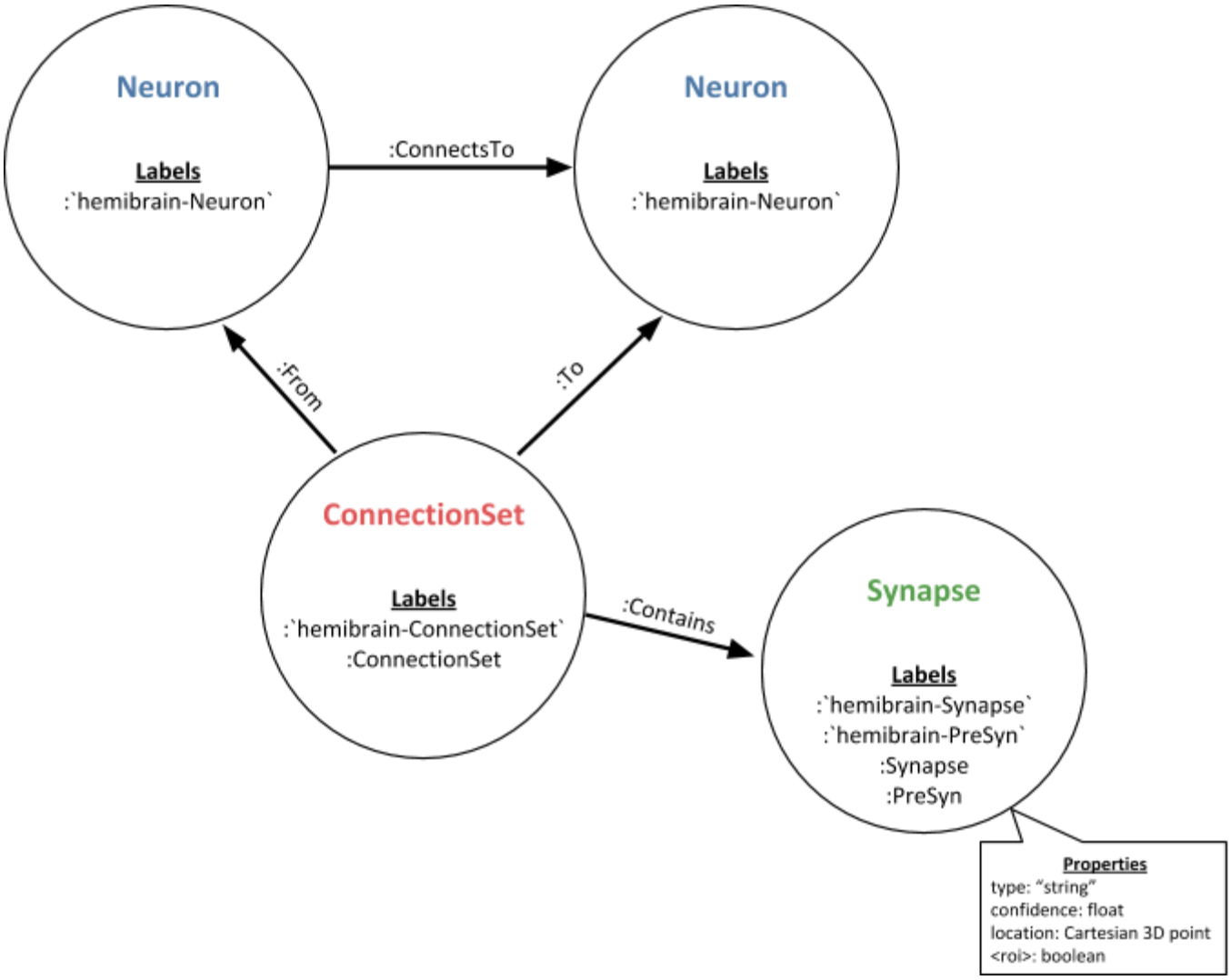
Returns x y z location of all synapses in the (L)SPS ROI

```
MATCH (s:`hemibrain-Synapse`)
WHERE s.`(L)SPS`
RETURN  s.location.x AS X, s.location.y AS Y, s.location.z AS Z
```

Returns type and x y z location of synapses in the FB ROI on a neuron (n)

```
MATCH (n:`hemibrain-Neuron`)-[:Contains]->(:SynapseSet)-[:Contains]->(s:`hemibrain-Synapse`)
WHERE n.bodyId = 541127846 AND s.FB
RETURN  n.bodyId AS ID, s.type AS TYPE,  s.location.x AS X, s.location.y AS Y, s.location.z AS Z
```

Two connecting neuron nodes contain a connection set node.
The connection set node contains the synapse nodes for that particular connection.



EXAMPLE QUERIES

Returns the number of postsynaptic densities in the AVLP ROI involved in the connection from (m) to (n)

```
MATCH (n:`hemibrain-Neuron`{bodyId:950880008})-[c:ConnectsTo]->(m{bodyId:5813027276}),
(m)<-[:From]-(cs:ConnectionSet)-[:To]->(n), (cs)-[:Contains]->(p:PostSyn{`AVLP`:true})
RETURN count(p)
```