Neuro Disjunctive Syllogistic Reasoning on a Sphere – How may clever monkeys have the grape?

Pahulmeet Singh

Master of Science in Computer Science

Rheinische Friedrich-Wilhelms-Universität Bonn
Deutschland

June 23, 2023

I hereby declare that this thesis was formulated by myself and that no sources or tools than those cited were used.

23 June, 2023 — Pahulmeet Singh

Date — Signature

1. Erstgutachter: Dr. Tiansi Dong
2. Zweigutachter: Prof. Dominik Bach

Abstract

Deductive reasoning to conclude based on given premises is known as Syllogistic Reasoning or Syllogism. Machine Learning approaches are often used to simulate the human reasoning process, which is also applied to syllogisms. Monkeys can also do syllogistic reasoning even though they do not speak like humans so this encourages us to look for different approaches to solve the problem. The research work in the "Learning Syllogism with Euler Neural-Networks" paper (16) pointed toward Euler Diagrams on a self-explainable neural network. Here, we propose using the 2-D Euler Diagrams on the surface of a Sphere as cones. This method solves syllogistic as well as disjunctive syllogistic reasoning which demonstrates how monkeys might solve reasoning. Each premise and conclusion is represented as individual cones on the surface of the sphere. Based on the spatial orientations of the cones, the cones can have different relations among them namely, disconnected, partial overlap, proper part, inverse part, and equal part. While working on a sphere, gives us an edge in representing the complement of a set quite easily which helps in disjunctive syllogism. Then, the neural network runs based on different loss functions and spatial relations that exist between different sets on the surface of the sphere. The network understands the five spatial relations between the sets and tries to reach a global loss of zero while making sure that all statements are met. This approach of working in 3-D on a Sphere is tested for all 256 syllogisms and lets us know if a given syllogism is satisfiable or not and provides the satisfying spatial orientation of the cones on the sphere's surface. Since we are working on the surface of a sphere with cones, it is easier to work with sets and their complements for disjunctive syllogistic reasoning.

Contents

List of Figures

Chapter 1

Introduction

1.1 Reasoning

Reasoning is one of the tests of intelligence where we are presented we different statements and draw conclusions based on the given rules in the statements. By definition from Britannica (10), 'the process of thinking about something in a logical way in order to form a conclusion or judgment is known as Reasoning or the ability of the mind to think and understand things in a logical way'. It is an essential skill for humans, allowing us to make decisions, solve problems, and communicate effectively. Reasoning can be broken down into two broad categories: deductive reasoning and inductive reasoning.

Deductive reasoning is a top-down approach that starts with a general principle and applies it to specific cases to reach a logical conclusion (3). For example, if we know that all dogs are mammals and Lyka is a dog, then we can deductively conclude that Lyka is a mammal. Deductive reasoning is often used in mathematics and formal logic, where conclusions can be drawn with certainty. Syllogistic reasoning which we are interested in, is deductive reasoning as well.

Inductive reasoning, on the other hand, is a bottom-up approach that involves drawing conclusions based on observations or patterns in data. Inductive reasoning allows us to make probabilistic predictions and generalizations based on past experiences or data. For example, if we observe that every time we throw a ball in the air, it comes back down, we can inductively conclude that all objects thrown in the air will come back down. But consider another example (9), where you go to a cafe every day for a month, and every day, the same person comes at exactly 11 am and orders a cappuccino. The specific observation is that this person has come to the cafe at the same time and ordered the same thing every day during the period observed. A general conclusion drawn from these premises could be that this person always comes to the cafe at the same time and orders the same thing. Inductive reasoning is commonly used in science and everyday decision-making.

In addition to deductive and inductive reasoning, there are other forms of reasoning that humans use to make sense of the world. Abductive reasoning (14) involves using available evidence to generate the most likely explanation for an observation or phenomenon. Analogical reasoning involves drawing comparisons between two different things to gain insight into a particular problem or situation. Logical reasoning involves using a set of premises and rules to derive a conclusion.

Regardless of the type of reasoning used, the quality of reasoning depends on the accuracy and relevance of the information used to draw conclusions. Critical thinking skills are essential for effective reasoning, as they allow individuals to evaluate the validity of information and consider alternative explanations or perspectives. Good reasoning involves recognizing and avoiding fallacies or errors in thinking, such as confirmation bias or the ad hominem fallacy.

Reasoning is a complex cognitive process that involves many different skills and strategies. Developing strong reasoning skills requires practice and reflection, as well as a willingness to challenge assumptions and consider alternative perspectives. As technology advances, machines are becoming increasingly capable of reasoning and decision-making. However, human reasoning remains a critical skill for problem-solving, innovation, and effective communication.

Moving ahead in this thesis report, we will talk about 'syllogistic reasoning' which is a type of deductive reasoning. Syllogistic reasoning is a logical process that involves drawing a conclusion based on two premises, where the conclusion follows necessarily from the premises. Syllogisms are composed of two premises and a conclusion, and the conclusion must be logically valid based on the premises. Syllogisms can be categorical, meaning they involve classes of objects or concepts, or conditional, meaning they involve propositions that are conditional on other propositions being true. Overall, syllogistic reasoning is a form of deductive reasoning because it starts with general principles (the premises) and applies them to specific cases to derive a logical conclusion.

1.2 Syllogistic reasoning

Syllogistic reasoning is a building block when it comes to logical and critical thinking. The Greek philosopher, Aristotle, also a scientist and a polymath is the one who developed syllogism and states it as "discourse in which, certain things being stated something other than what is stated follows of necessity from their being so" (7). To put it in simpler words, it is a type of deductive reasoning that is used to draw a conclusion based on two given premises. One famous example is where we conclude that "Socrates is mortal" based on given statements.
The Example:

Statement 1: Man is mortal. $\langle MajorPremise \rangle$
Statement 2: Socrates is a man. $\langle MinorPremise \rangle$
Conclusion: Socrates is mortal.

Now, here we have our 2 premises which arrive at a conclusion. The first statement is called the major premise and the second statement is referred to as the minor premise. Here, the major premise refers to a widely accepted or believed to be true general principle or proposition while the minor premise is the one that is applied to a specific situation or case. Let us consider one another example,
The Example:

Statement 1: All mammals are warm-blooded animals. $\langle MajorPremise \rangle$
Statement 2: All dogs are mammals. $\langle MinorPremise \rangle$
Conclusion: All dogs are warm-blooded animals.

Here, we introduce the concept of subject, middle, predicate, and mood. (15)

- Subject: It is the term that is being discussed or described in the premise or conclusion of the argument. In the above example, "dogs" is the subject term.
- Middle: This term builds a relationship between the major and minor premise and helps in drawing a conclusion between them. In our example, "mammals" is the middle term.
- Predicate: It is the term that is being affirmed or denied about the subject. Here in this example, "warm-blooded" is the predicate.

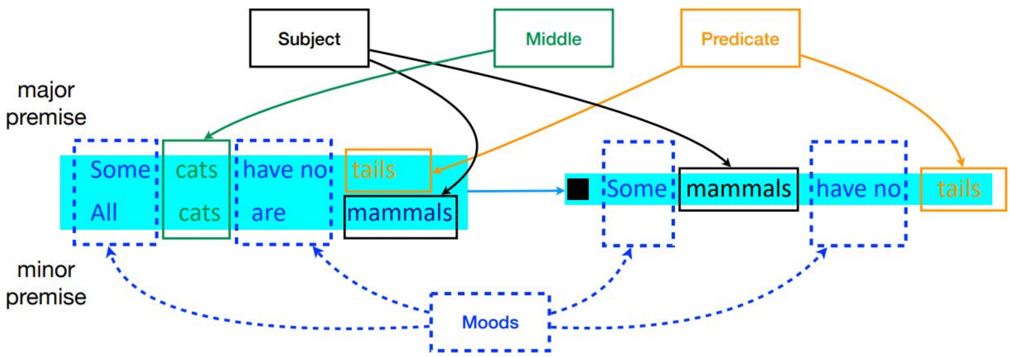

Figure 1.1: Components of Syllogistic Reasoning (15)

1.3 Moods and types of Syllogisms

The term "mood" is used to describe the general form of a syllogism, and it is represented by a series of letters that indicate the quantity and quality of the premises and the conclusion. There are 64 possible moods and out of them, only 24 are valid.

- Universal affirmative (A): These are statements that assert that a certain category is true for all members of a group. For example, "All dogs are mammals."
- Universal negative (E): These are statements that assert that a certain category is not true for any member of a group. For example, "No birds are reptiles."
- Particular affirmative (I): These are statements that assert that a certain category is true for at least one member of a group. For example, "Some cats are black."
- Particular negative (O): These are statements that assert that a certain category is not true for at least one member of a group. For example, "Some dogs are not friendly."

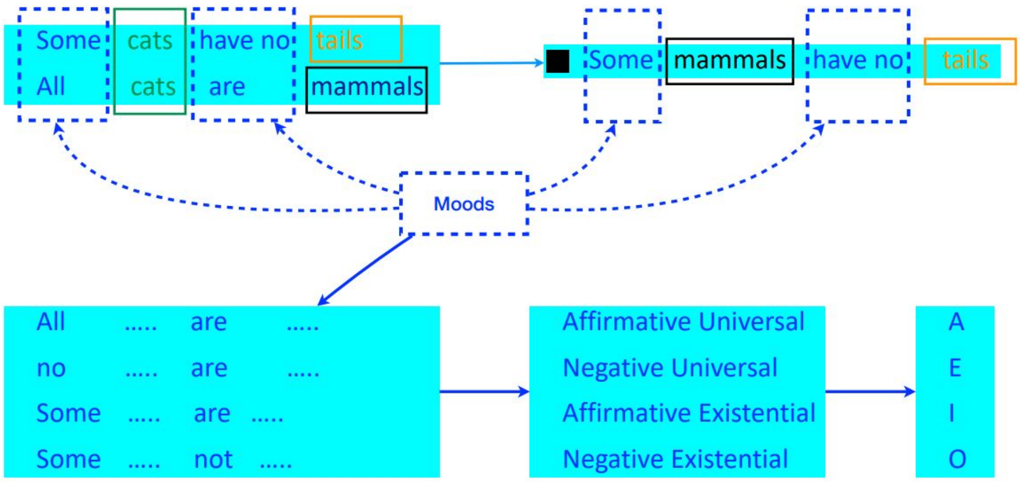

Figure 1.2: Structure of Syllogistic Reasoning (15)

In the second example with dogs, it is a AAA mood Syllogism since it has three Universal affirmative statements. Also, based on the subject, middle, predicate, and moods, there are 256 possible structures when we have 2 statements and 1 conclusion. Types of Syllogistic Reasoning

- A Syllogistic Reasoning is Valid if the premises are true, then the conclusion must be true
- A Syllogistic Reasoning is Invalid there is a case such that the premises are true, the conclusion is false
- A Syllogistic Reasoning is Satisfiable there is a case such that the premises are true, the conclusion is true
- A Syllogistic Reasoning is Unsatisfiable if it is not possible that all the premises and the conclusion are true OR if the premises are true, then the conclusion must be false

The topic has been studied for centuries and syllogistic reasoning plays an integral role in different fields where logical deductions are required. For eg, for Philosophical analysis, Lawyers and judges use it to measure the robustness or strength of an argument based on legal reasonings, used in Mathematics and other Sciences. In a similar fashion, Artificial intelligence is no different as this field has found its applications when we are working with reasoning.

There are several types of syllogistic reasoning, including:

- Categorical Syllogisms: These syllogisms involve three categorical propositions that have three different terms, but the terms must be arranged in a specific order to ensure that the syllogism is valid.
- Hypothetical Syllogisms: These syllogisms involve hypothetical propositions in the premises, where one or both of the premises are in the form of a conditional statement ("if-then" statement). (8)
- Disjunctive Syllogisms: These syllogisms involve a disjunction ("either...or") in one of the premises.
- Inductive Syllogisms: These syllogisms involve propositions that are based on induction, where the conclusion is based on specific observations or examples.

- Sorites: A sorites is a chain of syllogisms that are linked together, where the conclusion of one syllogism becomes the premise of the next syllogism. (11)

Each type of syllogism has its own set of rules and conventions that must be followed to ensure the validity of the argument.

Chapter 2

Disjunctive Syllogistic Reasoning

Disjunctive Syllogisms follow the principle that if one of the given two premises is false then the other must be correct. In a mathematical sense,

Statement 1: A $\vee B$
Statement 2: $\neg A$
Conclusion: B

This states that if the nature of the premise states that, it is going to be either "A" or "B" and the second premise states that it is not "A" then it must be "B".

Disjunctive syllogistic reasoning is a type of deductive reasoning that involves a disjunctive premise and a conclusion that is derived based on the exclusion of one of the disjuncts. A disjunctive premise is a statement that presents two alternatives, either one of which must be true. For example, "Either it will rain tomorrow, or the sun will shine."

Disjunctive syllogistic reasoning proceeds in two steps. First, the disjunctive premise is presented, along with a second premise that excludes one of the alternatives. For example, "It will not rain tomorrow." Second, the conclusion is drawn by excluding the disjunct that has been excluded in the second premise. In this case, the conclusion is "Therefore, the sun will shine tomorrow."

Disjunctive syllogistic reasoning is a powerful tool for drawing conclusions based on limited information. It is often used in legal and forensic contexts, where the exclusion of one possibility can lead to the identification of a suspect or the resolution of a case. Disjunctive syllogisms can also be used in scientific reasoning, where the exclusion of one hypothesis can help to identify the most likely explanation for an observation or phenomenon.

One of the strengths of disjunctive syllogistic reasoning is that it allows for the exclusion of one possibility based on the truth of the other. This can be useful in situations where it is difficult or impossible to directly observe or measure all the relevant variables. For example, in a medical diagnosis, the exclusion of one disease can help to narrow down the range of possible diagnoses and focus the investigation on the most likely causes.

However, there are also limitations to disjunctive syllogistic reasoning. One of the main challenges is that the conclusion is only valid if one of the disjuncts is true and the other is false. If both disjuncts are false or both are true, then the conclusion is invalid. Additionally, the validity of the conclusion depends on the accuracy and relevance of the premises. If the premises are false or irrelevant, then the conclusion may be invalid.

Overall, disjunctive syllogistic reasoning is a valuable tool for drawing conclusions based on limited information. It allows for the exclusion of one possibility based on the truth of the other and can be used in a wide range of contexts, from legal and forensic investigations to scientific research and medical diagnosis. However, it is important to be aware of the limitations of disjunctive syllogistic reasoning and to carefully evaluate the accuracy and relevance of the premises before drawing a conclusion.

Disjunctive Syllogistic Reasoning is something that exists in animals as well even though they do not speak a natural language like humans. Important excerpt taken from the paper (17), "Statement of Relevance Which cognitive capacities are unique to humans and which are shared with nonhuman primates? Humans have been pondering this question for centuries. One potentially unique domain is logical reasoning—for example, solving disjunctive syllogisms: Given that A or B is true, if not A is true, then B is true. If this form of reasoning is dependent on verbal labels for logical operators, it should not be possible in nonhuman animals. We gave nonhuman primates disjunctive syllogism problems that they could solve to earn a favored food, grapes. A subset of the animals was quite successful at the task, earning grapes almost 75 percent of the time. How widespread this ability is at the population level is unknown. However, the observation that even one nonhuman primate can engage in this logical operation is proof of the existence of the cognitive capacity in nonverbal, nonhuman primates. This finding adds to the growing body of research showing what types of logic are possible in the absence of language."

To illustrate this, nine adult baboons were used in the experiment at the Seneca Park Zoo in Rochester, New York. The baboons were first trained for the experiment as shown in the figure 2.2. They were presented with 4 containers, 2 on the left and 2 on the right. The trainer would first cover the left containers so that the monkey does not know into which of the 2 containers the grape is dropped and similarly does the same thing with the containers on the right. Then, the monkey is trained to answer on a decision board as to where the grape must be.

After training, the grape is dropped into one of the left 2 covered jars and the monkey is seeing as to what is happening right in front of it. Then, they repeated the dropping of grapes behind the other 2 jars on the right. So, now all 4 jars have a 50 percent probability and now the monkey sees that if the grape is in the first jar then it does not look for the other jar. So, the monkey has some sense that the second jar is going to be empty and now starts to either pick jar 3 or 4 to find the grape. However, if the first jar is empty then the monkey would pick up the second jar knowing that there must be a grape inside it if it was not found in the first jar. So, the experiment showed that some of the clever monkeys had the capability to do disjunctive syllogistic reasoning.

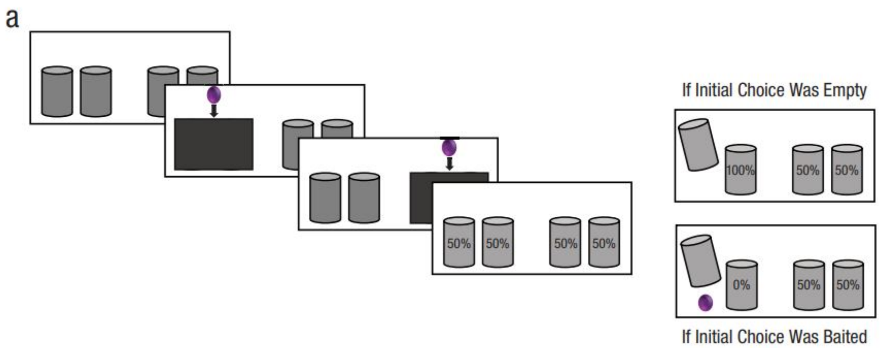

Figure 2.1: A schematic of the testing procedure and a sample trial sequence. On each trial (a), there were two sets of two baiting cylinders each. The experimenter placed an occluder in front of one of the sets and then showed a grape above it before placing the grape in one of the two occluded cylinders. This process was repeated for the other two baiting cylinders, and then the occluder was removed. The monkey could then select any of the cylinders, after which the experimenter revealed whether the monkey had found a grape. The chances that can be deduced via disjunctive syllogism are listed here on the cylinders, both for when the first choice was correct and for when it was not.(17)

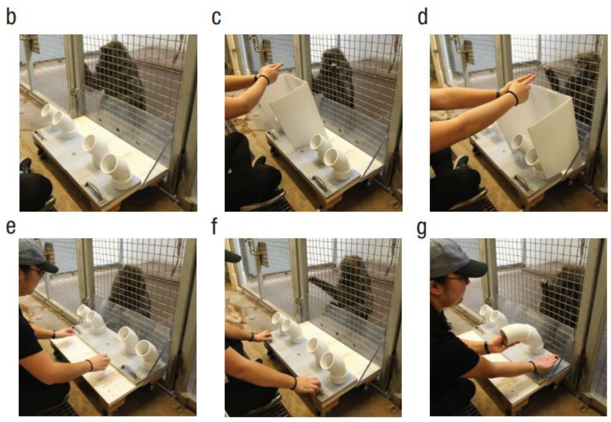

Figure 2.2: The photos illustrate moments in a sample trial. (17)

However, during the training only 4 of the baboons were able to pass the criterion and used it in the testing phase. But the experiment shows that there must be some level of reasoning going inside the animal's brain which does not require any language to be spoken. This disjunctive syllogistic reasoning can be implemented on the surface of a sphere which will be shown later in the following chapters.

Chapter 3

Euler Diagrams

Euler diagrams are a type of diagrammatic representation used to visually illustrate the relationships between sets or categories. The diagrams consist of circles or other closed shapes that overlap or do not overlap to show the relationships between the sets or categories. Euler diagrams are named after the Swiss mathematician Leonhard Euler, who developed them in the 18th century.

Euler diagrams are commonly used in logic, mathematics, computer science, and other fields where sets or categories are studied. They can be used to represent relationships between sets or categories, such as inclusion, intersection, and exclusion. For example, a simple Euler diagram might show two sets, A and B, where A is completely contained within B. The diagram would consist of two circles, one inside the other, with the smaller circle representing A and the larger circle representing B.

Euler diagrams can also be used to represent more complex relationships between sets or categories. For example, a Venn diagram is a type of Euler diagram that shows the relationships between three or more sets. Venn diagrams use overlapping circles to represent the relationships between the sets. For example, a Venn diagram might show three sets, A, B, and C, where A and B intersect, but C is completely separate from both A and B. The diagram would consist of three circles, with two overlapping and one separate.

Euler diagrams have several advantages over other forms of representation, such as lists or tables. They are visually appealing and easy to understand, even for people who are not familiar with the underlying concepts. They also provide a quick and easy way to compare the relationships between multiple sets or categories. Additionally, Euler diagrams can be easily modified to represent changes in the relationships between the sets or categories.

However, there are also limitations to Euler diagrams. They are limited in their ability to represent complex relationships between sets or categories, particularly when the relationships are not well-defined or when there are multiple possible interpretations of the relationships. Additionally, Euler diagrams can be difficult to construct accurately, particularly when there are multiple sets or categories involved.

Overall, Euler diagrams are a useful tool for visually representing the relationships between sets or categories. They provide a quick and easy way to compare the relationships between multiple sets or categories, and they are visually appealing and easy to understand. However, it is important to be aware of their limitations and to use them appropriately in the context of the underlying concepts.

Given a set A = [1,2,3,4,5,6,7] and set B = [4,5,6,7,8,9,10], we notice that there are some elements repeating in both sets and both these sets are part of a bigger set of "Natural" numbers (N). This can be represented better as a diagram for better visualization and understanding and thus form valuable logical deductions from the given sets and their relationship.

So, given the image above, we understand that "N" is a bigger set and contains 2 sets "A" and "B". To add to it, we get a visualization that there is a partial overlap between the 2 sets and there are numbers that occur in both sets.

As shown above, in the diagram the sets are represented as circles or curves and how they are spatially placed in the space defines their relationship with each other. They could be disconnected, partially overlapped, complete overlap, or an inverse complete overlap. As per our example, "A" and "B" are completely inside the set "N" and there exists a partial overlap between the sets "A" and "B".

The diagrams are used to visually illustrate logical relationships and make them easier to understand. They can be used to represent different types of logical statements, including categorical statements used in syllogistic reasoning. Euler diagrams are widely used in logic, mathematics, and computer science to represent sets, classes, and categories, and to reason about their relationships. They are also used in other fields, such as linguistics, biology, and social sciences, to model and analyze various phenomena.

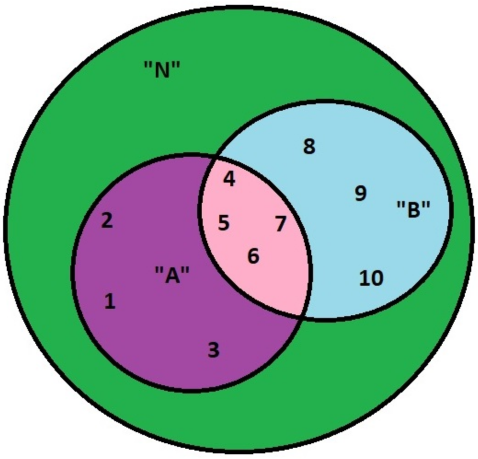

Figure 3.1: Figure depicts the Euler Diagram using the sets 'N', 'A' and 'B'

3.1 Difference between Venn Diagrams and Euler Diagrams

One of the main differences between Euler diagrams and Venn diagrams is their complexity. Euler diagrams can be used to represent simple relationships between sets, such as inclusion, intersection, and exclusion. In contrast, Venn diagrams can represent more complex relationships, such as partial overlap, subsets, and complements. Venn diagrams can also be used to represent the relationships between three or more sets, whereas Euler diagrams are typically limited to two sets. (6)

Another difference between the two is the way they represent empty sets. In Venn diagrams, empty sets are represented by regions with no overlap between circles. Euler diagrams, on the other hand, do not show empty sets explicitly, as they only depict the relationships between sets that actually exist.

A third difference between Euler diagrams and Venn diagrams is their construction. Euler diagrams are typically constructed by drawing circles or other closed shapes that represent the sets or categories and then overlapping them as necessary to show the relationships between them. Venn diagrams, on the other hand, are typically constructed by drawing circles that overlap in specific ways to show the relationships between the sets or categories.

Finally, Euler diagrams and Venn diagrams differ in their use in practice. Euler diagrams are commonly used in fields such as logic, mathematics, and computer science to represent simple relationships between sets or categories. Venn diagrams, on the other hand, are more commonly used in fields such as statistics, probability, and data visualization, where more complex relationships between sets or categories are often encountered.

In conclusion, Euler diagrams and Venn diagrams are two types of visual representations used to illustrate the relationships between sets or categories. Although they share some similarities, they differ in their complexity, the way they represent empty sets, their construction, and their use in practice. Both diagrams are useful tools for representing relationships between sets or categories, and the choice between them depends on the specific context and the complexity of the relationships being represented.

Both diagrams are based on the set theory. The main difference between Venn and Euler diagrams is that a Venn diagram shows all possible logical relationships between sets, while an Euler diagram only shows existing relationships. In other words, in a Venn diagram, you have to depict each intersection between each set, even if the intersection is empty, while in an Euler diagram, you only depict intersections that are not empty. Suppose we have three sets:

A = [1, 3, 5, 7]

B = [2, 4, 6, 8]

C = [5]

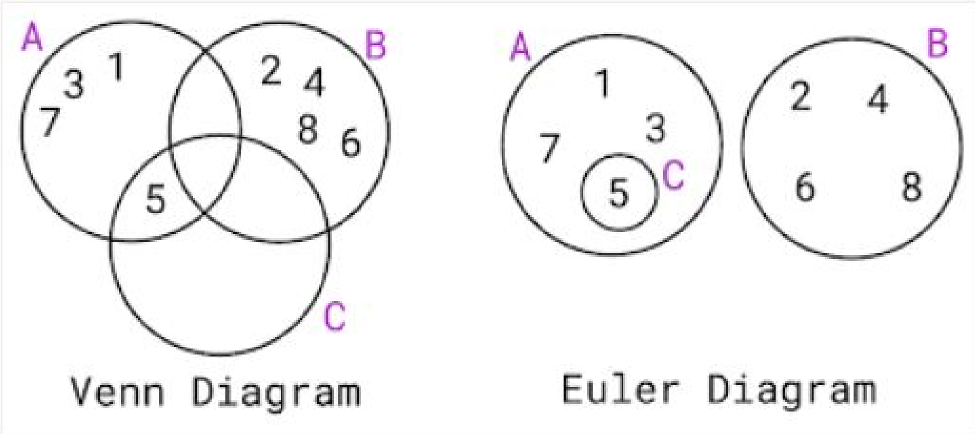

Figure 3.2: Difference between a Venn Diagram and an Euler Diagram (4)

3.2 Representing statements as Euler Diagrams

In order to introduce Euler diagrams for solving syllogistic reasoning, we need to represent the statements as Euler diagrams as shown in figure 3.3. But there are two problems, one is that we do not have a one-to-one mapping, and the second is that the statement 'Some X are Y' and 'Some X not Y' have a common Euler diagram possible where set 'Y' is inside set 'X'.

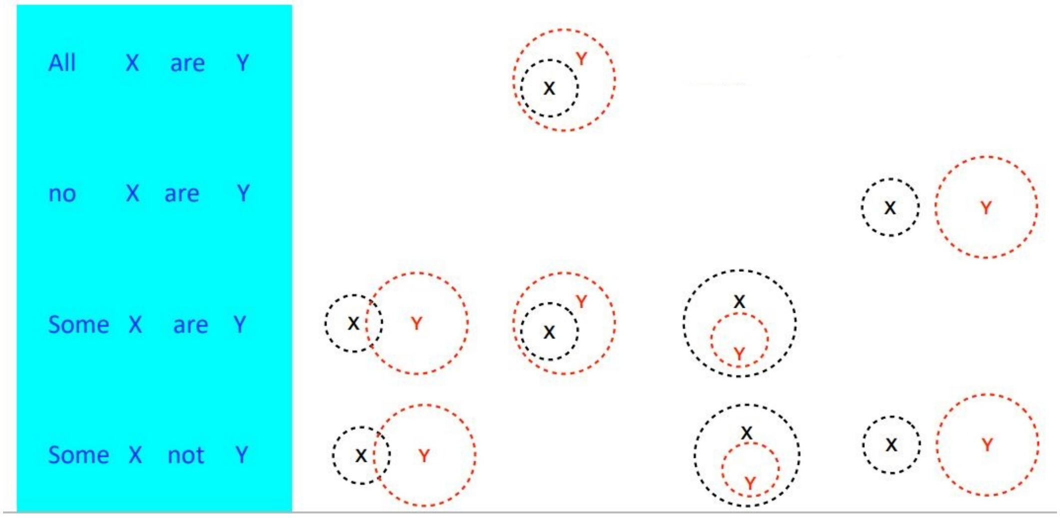

Figure 3.3: Mapping between Syllogistic Statement and Spatial Statement (15)

So, a better way to represent the statements is shown in figure 3.4 where we have one-to-one mapping and also unique Euler diagrams.

Now that we have our Euler diagrams representing statements, it is easier to convert the natural language into a pictorial form. We only need these diagrams to be passed to our Euler Neural Network(ENN) and the mathematical formulae will be applied to these diagrams to transition from one spatial state to another based on the target relation. Each of the Euler diagrams will have a certain center point in space and radius, both these center points and radius will be changed as per the needs of the algorithm to satisfy the target relation, if possible. Here, you will see the Euler Diagrams in a 2-Dimensional plane but we need to work in 3-Dimensional space, and for that, we will need to work with the coordinate system on the surface of the sphere. In 3D, we worked on a sphere and cones on its surface which are shown in figure 3.5 below:

Here, we have a Sphere with a radius of 'R'. There exists a cone on the surface of the sphere with its center point 'C'. The center point 'C' is represented using R, θ, ϕ, and α. More on this coordinate system in the following chapter.

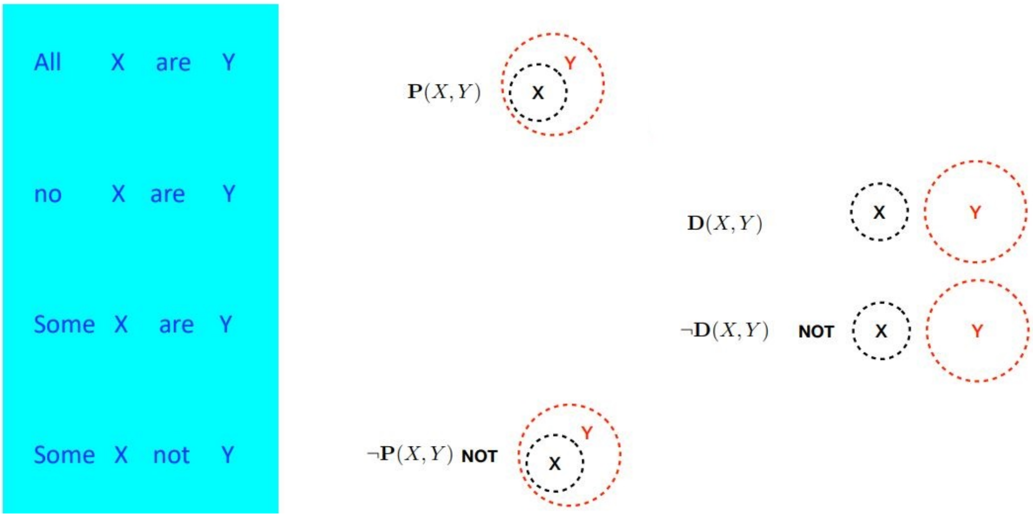

Figure 3.4: One-to-one mapping between Syllogistic Statement and Target Spatial Statement (15)

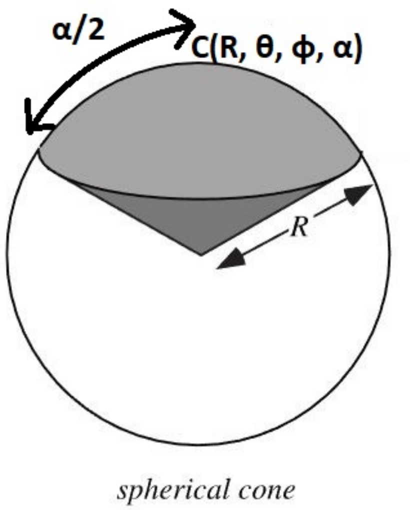

Figure 3.5: Representing the cones used in the algorithm (12)

Chapter 4

Co-ordinate System on the Sphere

A sphere is a 3-dimensional geometric shape that is analogous to a circle in a 2-D plane, hence, it is a round shape with no edges or corners and the surface of a sphere is smooth and continuous. The surface of the sphere has a set of points that are equidistant from the center of the sphere. This is a fundamental concept of mathematics and requires a corresponding 3-D coordinate System which is used to work with the position of points on the surface of the sphere.

In mathematics, when we are working with the position of points in space or on the surface of a geometric shape, a coordinate system is the go-to mathematical tool. While working in 3D space and especially on the surface of a sphere, we can work with cartesian coordinates or polar coordinates.

A Cartesian coordinate system, also known as a rectangular coordinate system, is a coordinate system that specifies each point uniquely in a plane by a pair of numerical coordinates, which are the signed distances from the point to two fixed perpendicular directed lines, measured in the same unit of length. The coordinates can be positive, negative or zero, depending on the position of the point with respect to the origin, which is the intersection point of the two perpendicular lines. In the two-dimensional system, the two lines are usually given the names x-axis and y-axis, while in the three-dimensional system, an additional z-axis is added perpendicular to the x-y plane. The Cartesian coordinate system is named after René Descartes, a French mathematician, and philosopher who first introduced it in the seventeenth century (2). It is widely used in mathematics, physics, engineering, computer graphics, and other fields for representing and analyzing geometric figures and functions.

A polar coordinate system is a two-dimensional coordinate system in which each point is determined by a distance from a fixed point, called the pole, and an angle from a fixed direction called the polar axis. In other words, instead of specifying the position of a point using the coordinates (x, y) as in the Cartesian coordinate system, a point in the polar coordinate system is represented by the pair (r, θ), where r is the radial distance from the origin to the point, and θ is the angle measured counterclockwise from a fixed reference direction to the line segment connecting the origin to the point. The polar coordinate system is particularly useful for representing and analyzing circular and rotational motion, as well as shapes and phenomena that exhibit radial symmetry. It is also often used in mathematical and scientific fields such as physics, engineering, and geometry. Now while working on the surface of the sphere, the cartesian system(using x,y,z) to locate points in space is not ideal however, using the polar coordinate system (r, θ, ϕ) is the way forward as they make it easier to navigate on the surface of the sphere. It is defined by three values, which are the radius, the polar angle, and the azimuth angle (13). The radius is the distance from the center of the sphere to the point on the sphere. The polar angle is the angle between the positive z-axis and the line joining the point to the origin, while the azimuth angle is the angle between the positive x-axis and the projection of the line joining the point to the origin on the x-y plane. These values can be represented using the notation (r, θ, ϕ), where r is the radius, θ is the polar angle, and ϕ is the azimuth angle.

Here, as an example, consider the diagrams below to better understand how the polar coordinates look in 3-D space and how to navigate the location of points on the surface of the sphere.

As seen in the figure 4.1, point 'C' is represented as (ρ, θ, ϕ) in the 3-Dimensional space and it could easily be converted to the cartesian coordinate system when we need the x,y, and z coordinates. Here we are working with a sphere so the radius of the sphere (ρ) will remain fixed. But by changing the angles theta(θ) and phi(ϕ), we can move the point on the surface of the sphere with ease.

The spherical polar coordinates are a system of curvilinear coordinates that are natural for describing positions on the surface of a sphere. Θ is called the azimuthal angle in the x-y plane from the x-axis with $0 <= \theta < 2 * \pi$. Φ is the polar angle from the positive z-axis with $0 <= \phi < pi$. 'r' is the distance from the origin to any point on the surface of the sphere (13).

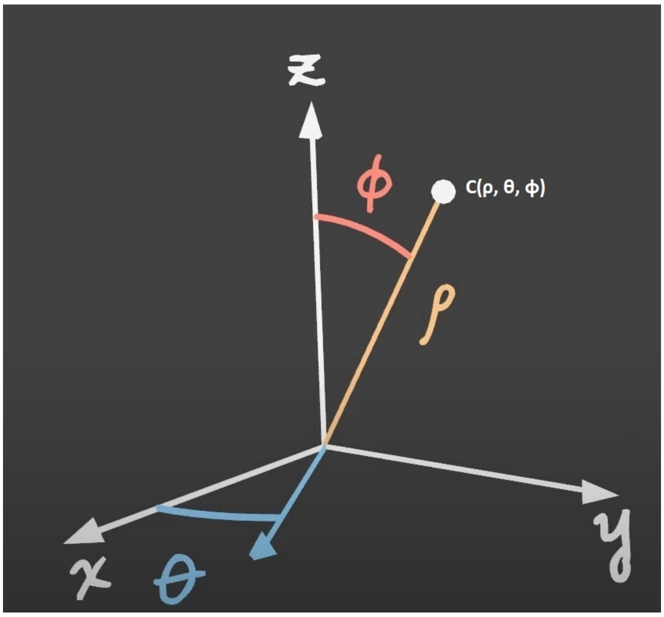

Figure 4.1: 3-Dimensional Polar Coordinate system (1)

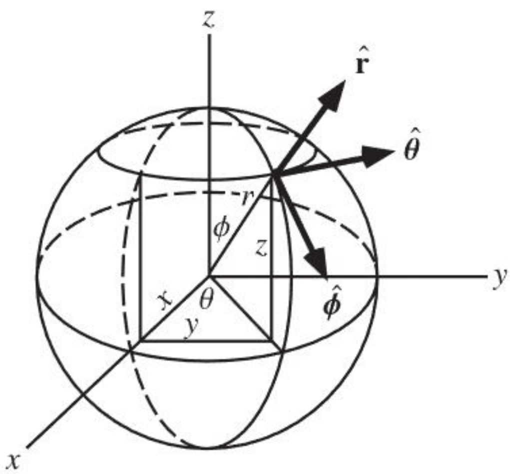

Figure 4.2: 3D representation of the Sphere and the coordinate system (13)

Chapter 5

Mathematics used for the Algorithm

Let us start with normal 2-D Circles on a plane and then later transition to 3-D cones on a sphere. A Circle 'v' is structured by a central point and a radius, (15)

- the orientation of the central point O_v is $\vec{\alpha}_v$
- the distance from the central point to the origin is l_v
- the radius is r_v

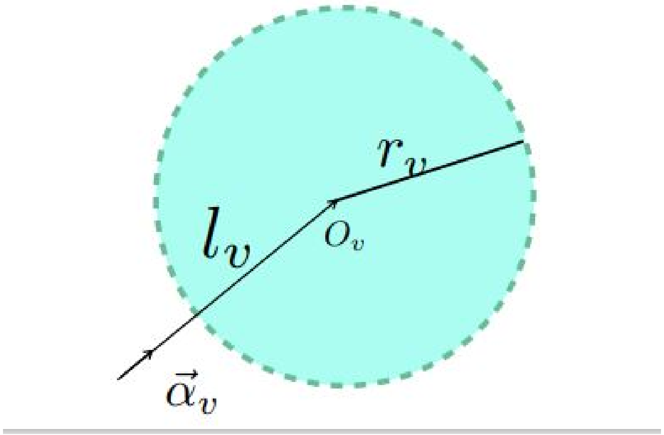

Figure 5.1: Geometric Structure of a Sphere (**?**)

Now, all the points inside this circle are the points of interest for our algorithm. Now 'V' is defined as the set point satisfying

$\|\vec{p} - l_v\vec{\alpha}\| < r_v$

Moving ahead towards 2 circles and the geometric relationship between them.

Disconnect: the distance between their central point vectors is greater than or equal to the sum of their radii. (15)

$$\|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| \geq r_v + r_w \tag{5.1}$$

Proper Part Of: Circle 'w' is completely inside Circle 'v' (15)

$$\mathbf{P}(\vec{v}, \vec{w}) \triangleq r_v + \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| - r_w \leq 0 \tag{5.2}$$

Partial Overlap: Sphere w partially overlaps with sphere v (15)

$$\mathbf{PO}(\vec{v}, \vec{w}) \triangleq \|l_w\vec{\alpha}_w - l_v\vec{\alpha}_v\| < r_v + r_w \wedge \|l_w\vec{\alpha}_w - l_v\vec{\alpha}_v\| > |r_v - r_w| \tag{5.3}$$

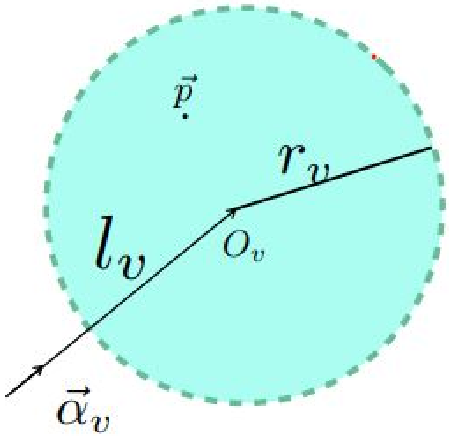

Figure 5.2: Geometric Structure of a Sphere (15)

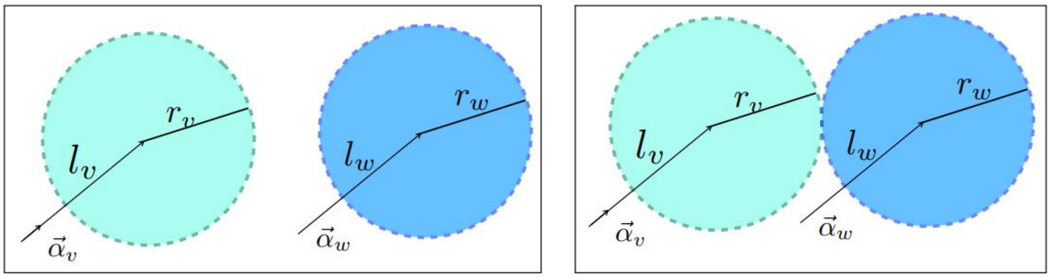

Figure 5.3: Disconnect Geometric Relation (15)

5.1 Loss Functions

Clearly, we do not have one unique loss function and we require different loss functions if we want to transition from one relation between the 2 circles to another.

Now, let us have a look at various loss functions needed to transition from one spatial relation to another. (15)

$$\mathcal{L}_{\overline{\mathbf{P}}}^{\mathrm{PO}}(\vec{v},\vec{w}) \triangleq \begin{cases} r_v - r_w - \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| & \neg\mathbf{PO}(\vec{v},\vec{w}) \\ 0 & \mathbf{PO}(\vec{v},\vec{w}) \end{cases} \tag{5.4}$$

$$\mathcal{L}_{\mathbf{P}}^{\mathrm{PO}}(\vec{v},\vec{w}) \triangleq \begin{cases} r_w - r_v - \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| & \neg\mathbf{PO}(\vec{v},\vec{w}) \\ 0 & \mathbf{PO}(\vec{v},\vec{w}) \end{cases} \tag{5.5}$$

$$\mathcal{L}_{\mathbf{D}}^{\mathrm{PO}} \triangleq \begin{cases} \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| - r_v - r_w & \neg\mathbf{PO}(\vec{v},\vec{w}) \\ 0 & \mathbf{PO}(\vec{v},\vec{w}) \end{cases} \tag{5.6}$$

$$\mathcal{L}_{\mathbf{PO}}^{\overline{\mathrm{P}}}(\vec{v},\vec{w}) \triangleq \begin{cases} r_w + \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| - r_v & \neg\overline{\mathbf{P}}(\vec{v},\vec{w}) \\ 0 & \overline{\mathbf{P}}(\vec{v},\vec{w}) \end{cases} \tag{5.7}$$

$$\mathcal{L}_{\mathbf{PO}}^{\mathrm{P}}(\vec{v},\vec{w}) \triangleq \begin{cases} r_v + \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| - r_w & \neg\mathbf{P}(\vec{v},\vec{w}) \\ 0 & \mathbf{P}(\vec{v},\vec{w}) \end{cases} \tag{5.8}$$

$$\mathcal{L}_{\mathbf{PO}}^{\mathrm{D}}(\vec{v},\vec{w}) \triangleq \begin{cases} r_v + r_w - \|l_v\vec{\alpha}_v - l_w\vec{\alpha}_w\| & \neg\mathbf{D}(\vec{v},\vec{w}) \\ 0 & \mathbf{D}(\vec{v},\vec{w}) \end{cases} \tag{5.9}$$

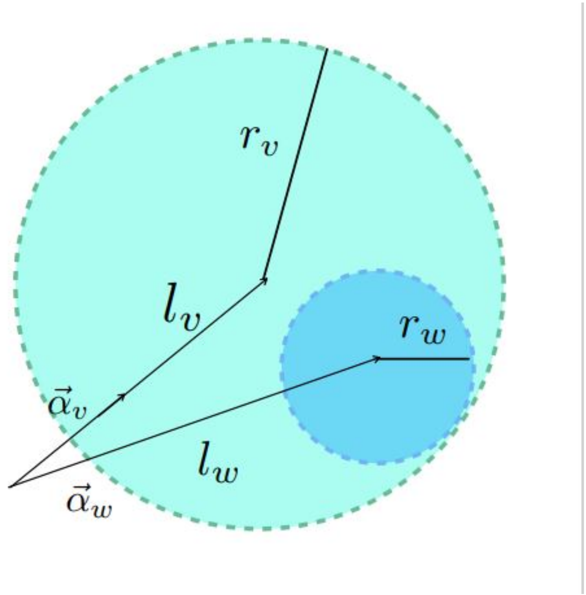

Figure 5.4: Proper Part Geometric Relation (15)

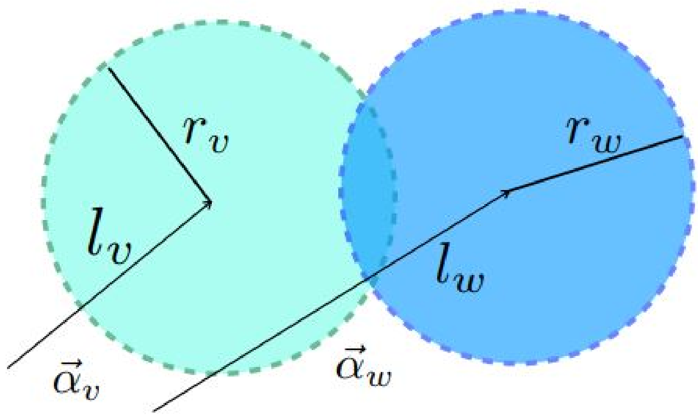

Figure 5.5: Partial Overlap Geometric Relation (15)

5.2 Transition from 2D to 3D space

Now, in order to transition these formulae from a 2-D plane to a 3-D Sphere, we need some changes. First and foremost, we start by using Polar Coordinate System. In the polar coordinate system, we define a point on a sphere as, p = (ρ, θ, ϕ) As defined earlier, rho is the radius of the sphere, theta is the angle with the X-Y plane and can span from [0,2*π] and phi is the angle with the Z-axis and can span from [0,π]

With this point 'p' we can define the center of a cone on a sphere but we need something that gives us some information about the radius of the cone. To define the radius, we introduce an angle 'alpha'(α) from the origin of the sphere.

Here, $\alpha/2$ is the radius of the cone and is calculated as,

R1 = $((\alpha/2)/360)$*2*$\pi * \rho$

Also, the point 'p' need to be converted from polar to cartesian coordinates, (13)

$p(\rho, \theta, \phi) - > C(x, y, z)$

$$x = \rho * \sin(\phi) * \cos(\theta) \tag{5.10}$$

$$y = \rho * \sin(\phi) * \sin(\theta) \tag{5.11}$$

$$z = \rho * \cos(\phi) \tag{5.12}$$

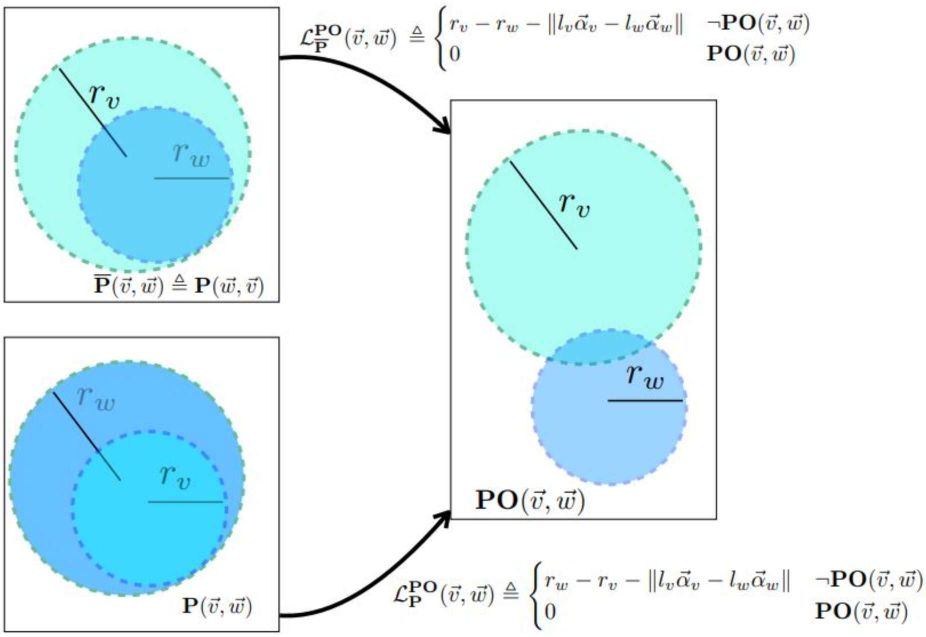

Figure 5.6: Loss function to map from Proper Part(or Inverse Proper Part) to Partial Overlap relation (15)

$$r = \frac{\rho^*\pi^*\alpha}{360} \tag{5.13}$$

Now, the center(C) can be anywhere on the sphere and the radius of the cone is dependent on the angle alpha. Let us assume one extreme case where we need a cone based on a syllogistic statement that covers everything inside it. In that extreme case, alpha can be 2*pi and can cover the whole sphere. A second case would be where the cone should span just a single point so it could be slightly more than zero. So, based on these 2 cases, our angle 'alpha' spans across (0,2*pi]

Next up we have the distance along the surface of the sphere which is a curved surface. Every time the radius of the cone is discussed or the distance between cones is discussed, it is always the curvilinear distance along the surface of the sphere. Let us assume we have 2 points c1(x1,y1,z1) and c2(x2,y2,z2) on a sphere of radius rho. So, distance 'd' is given by (5) ,

$$d = \rho * \arccos\left(\frac{x_1x_2 + y_1y_2 + z_1z_2}{\rho^2}\right)$$

Now, that we have the coordinates, distance, and radius formulae, we can use the loss functions defined for the planar circles to cones on the surface of the sphere. This also helps in checking the status between 2 cones.

$$Disconnect = \begin{cases} true, d - (r_1 + r_2) \geq 0 \\ false, otherwise \end{cases}$$

$$PartialOverlap = \begin{cases} true, d > |r_1 - r_2| \, \& d < |r_1 + r_2| \\ false, otherwise \end{cases}$$

$$ProperPartX = \begin{cases} true, d + r_1 \leq r_2 \\ false, otherwise \end{cases}$$

5.3 Rectified Spatial Units (ReSU)

Rectified Spatial Units (ReSU) are used by the neural network which helps in the transitions required to move from one neighborhood's spatial status to another. (16)

To make use of ReSU, we need to know the radii of cones($r1$ and $r2$) along the surface of the sphere and the circular/arc distance("d") along the surface of the sphere between the 2 cones. These ReSU are used by the Euler Neural Network to switch between the spatial statuses of cones. For example, if we need the spatial status between 2 cones to change from disconnected to overlap then decreasing the value of $ReSU_D^O$ will push the relation between them to the relation of being overlapped. The tolerance used in the formulae is a small positive value.

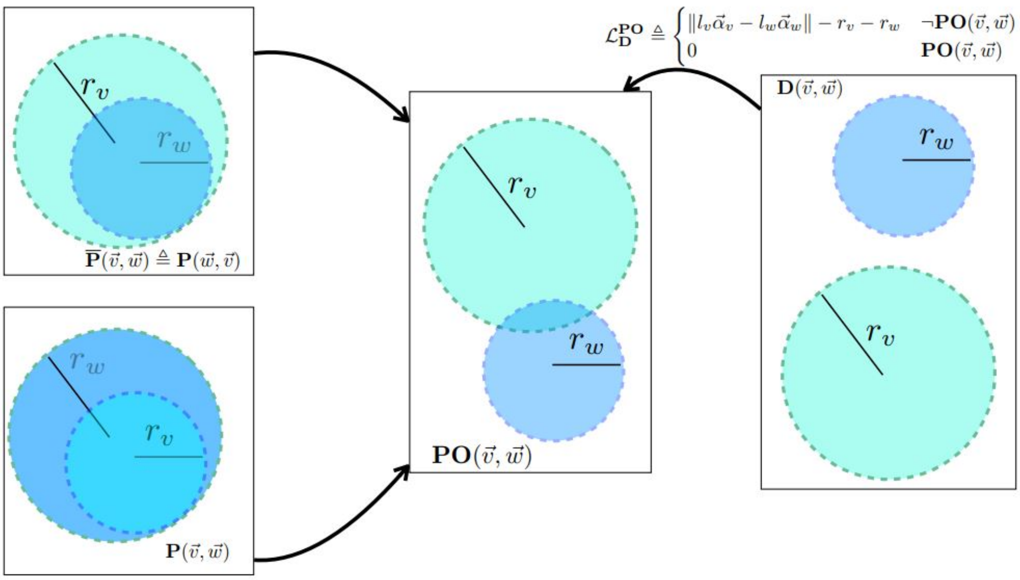

Figure 5.7: Loss function to map from Disconnect to Partial Overlap relation (15)

Here, we present the 6 ReSU used for the transitions.

$ReSU_D^O = \max\big(0, d - (r_1 + r_2 - tolerance)\big)$
$ReSU_O^D = \min\big(0, d - (r_1 + r_2 + tolerance)\big)$
$ReSU_O^P = \max\left(0, d + r_2 + tolerance - r_1\right)$
$ReSU_P^O = \min\left(0, d - r_1 - r_2 + tolerance\right)$
$ReSU_O^{IP} = \max\left(0, d + r_1 - r_2 + tolerance\right)$
$ReSU_{IP}^O = \min\left(0, d - r_1 - r_2 + tolerance\right)$

Now as we can see that we do not have a unique loss function for our problems so we have all different loss functions defined. These will be used depending on the transition map with the help of 'ReSU' functions.

5.4 Complement of Set on the Surface of Sphere

Given that we have a set 'p' represented as a cone on the surface of the sphere represented by angles (θ, ϕ, α), we can easily calculate the complement cone of the set 'p' as shown below:

$\theta_c = \pi + \theta$
$\theta_c = \begin{pmatrix} \theta_c - 2 * \pi, & \theta_c > 2 * \pi \\ \theta_c, otherwise \end{pmatrix}$
$\phi_c = \pi - \phi$
$\alpha_c = 360 - \alpha$

Let us assume that Figure 5.10 shows $p \vee q$ as the complete sphere. The blue region is the set 'p' and its complement $\neg p$ is the green region. Then, the green region must be the set 'q'.

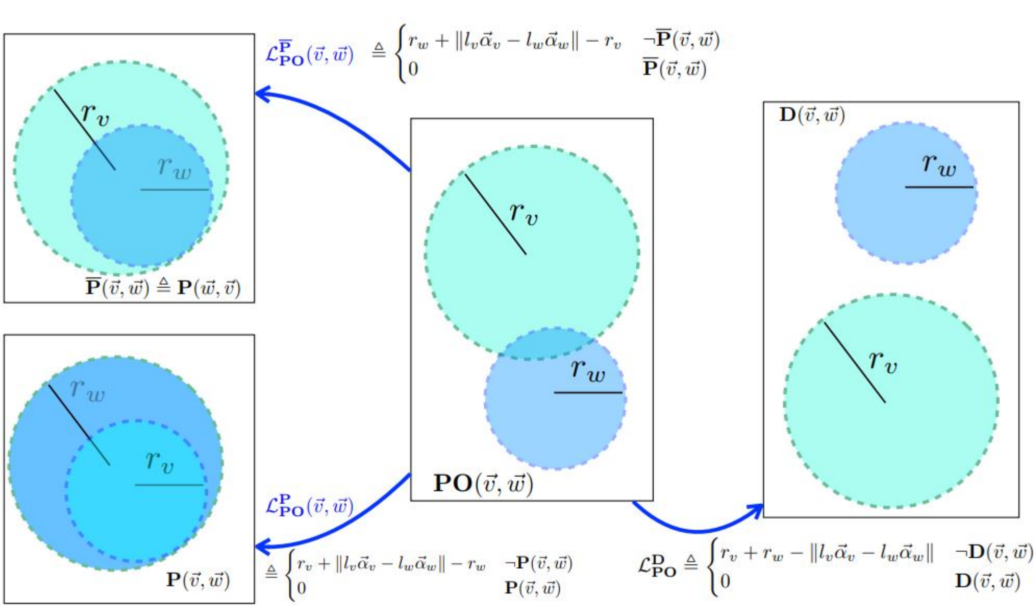

Figure 5.8: Loss functions to map from Partial Overlap to other spatial relations (15)

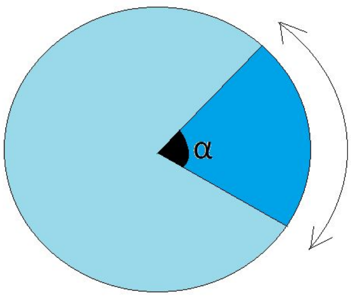

Figure 5.9: Representing the cone angle alpha from origin(center of sphere)

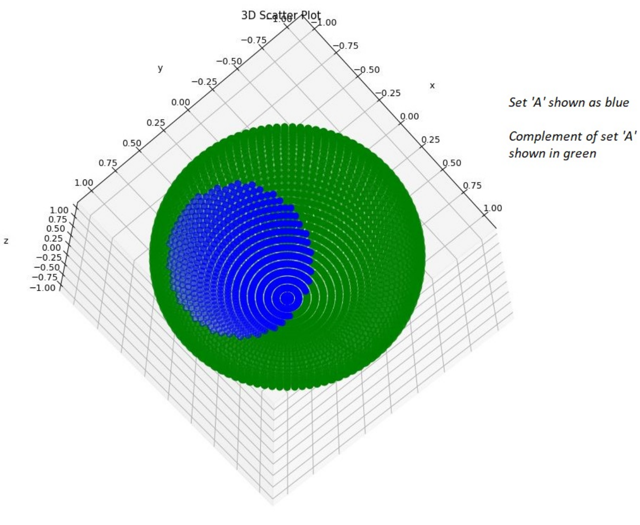

Figure 5.10: Set 'A' and its Complement

Chapter 6

System Architecture

Now that we have our mathematical formulae defined, we need to design the system architecture based on the transition map. The mapping will be followed by the loss functions that are required to move from one spatial relation to another during mapping. Next, the flowchart will depict how the euler neural network tries to solve the problem.

6.1 Transition Map

The transition map is the mapping the euler neural network follows to move from one spatial state to another. As an example, consider the 2 cones A and B and both cones are disconnected from each other on the surface of the sphere. Now as per the need of the target relation, the cones need to be 'Proper Part of' i.e. cone B is inside cone A. For this, the algorithm will run a loss function to go from 'Disconnected' to 'Partial Overlap' and then to 'Proper Part of'.

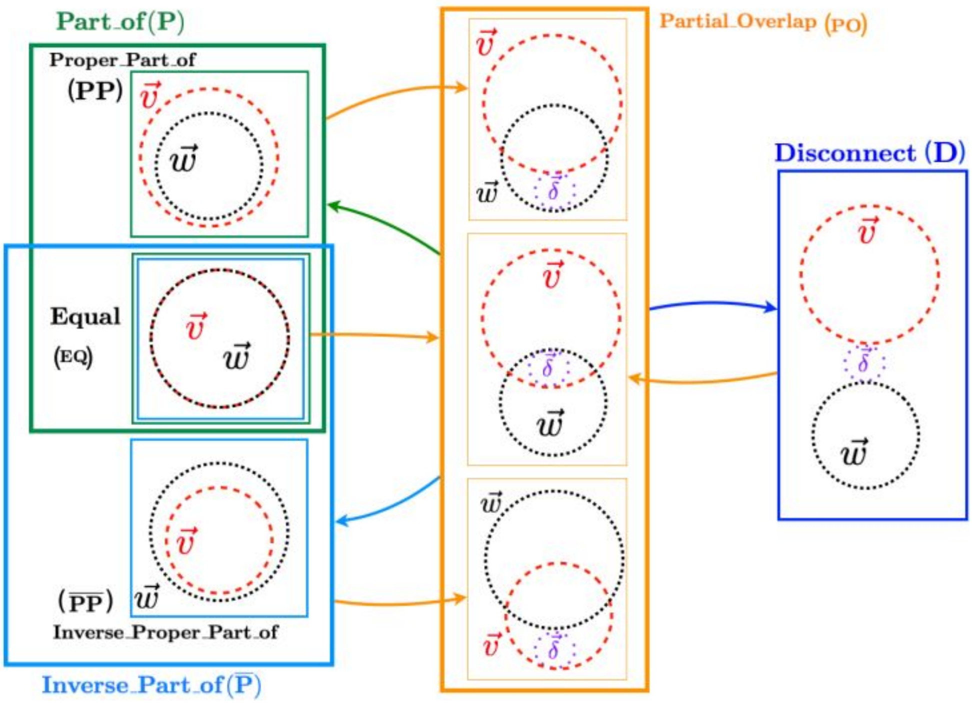

Figure 6.1: Figure Representing the complete transition map. The δ(radius of a tolerance circle) represents the tolerance values required to achieve a particular state i.e. as an example if the 2 circles are disconnected then there must be a distance of 2*δ between them(15)

6.2 Self-Explainable Euler Neural Network(ENN)

Once, we have three syllogistic statements, each term is represented in the form of a cone as shown in Figure 6.2. Also, each is translated into spatial relations between spheres. Now, this way we have Euler Diagram Embeddings and this is self-explainable(15).

The network runs with the end goal of reaching a global loss of zero or until the maximum number of iterations is reached. If the global loss is zero then that means the target spatial relation between the cones is possible and hence, the syllogism is satisfiable.

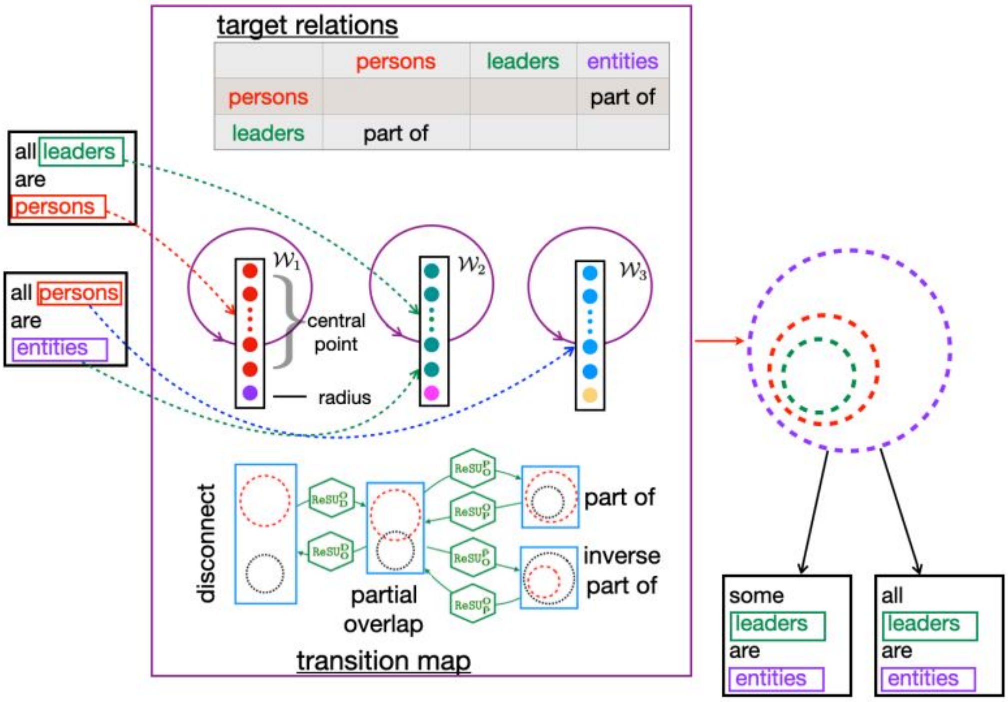

Figure 6.2: Self-explainable neural network for syllogistic reasoning (15)

6.3 Loss function transition table

The loss functions once defined mathematically, are the building blocks of our algorithm. There is no single loss function that runs the Euler neural network to completion and it depends on the transition map. In many cases, there is a need to follow one loss function by another to reach the target spatial relation. All the loss functions that are required to transition and in what order are mentioned in the figure6.3.

6.4 Flowchart

Here is the flowchart 6.4 of how the Euler Neural Network picks up 2 cones at one time and tries to reach a global loss of zero, if possible. The flowchart shows that the algorithm runs a total of three rounds. During each round, two out of the three cones are picked in order and one is fixed. Then, the loss function runs based on the target spatial relation and since we are working with 2 cones at a moment, they will always give the local loss of zero and complete the spatial transition.

6.5 Algorithm

Given the three cones on the sphere, they have a target relation among each other that must be reached to reach a global loss of zero. Once, we get that loss to zero then we can say that the two statements and their conclusion

Spatial Relations	Disconnected	Partial Overlap	Proper Part	Inverse Proper Part
Disconnected	-	$\mathcal{L}_{\mathbf{D}}^{\mathrm{PO}}$	$\mathcal{L}_{\mathbf{D}}^{\mathrm{PO}}$ $\mathcal{L}_{\mathbf{PO}}^{\mathrm{P}}$	$\mathcal{L}_{\mathbf{D}}^{\mathrm{PO}}$ $\mathcal{L}_{\mathbf{PO}}^{\overline{\mathrm{P}}}$
Partial Overlap	$\mathcal{L}_{\mathbf{PO}}^{\mathrm{D}}$	-	$\mathcal{L}_{\mathbf{PO}}^{\mathrm{P}}$	$\mathcal{L}_{\mathbf{PO}}^{\overline{\mathrm{P}}}$
Proper Part	$\mathcal{L}_{\mathbf{P}}^{\mathrm{PO}}$ $\mathcal{L}_{\mathbf{PO}}^{\mathrm{D}}$	$\mathcal{L}_{\mathbf{P}}^{\mathrm{PO}}$	-	$\mathcal{L}_{\mathbf{P}}^{\mathrm{PO}}$ $\mathcal{L}_{\mathbf{PO}}^{\overline{\mathrm{P}}}$
Inverse Proper Part	$\mathcal{L}_{\overline{\mathbf{P}}}^{\mathrm{PO}}$ $\mathcal{L}_{\mathbf{PO}}^{\mathrm{D}}$	$\mathcal{L}_{\overline{\mathbf{P}}}^{\mathrm{PO}}$	$\mathcal{L}_{\overline{\mathbf{P}}}^{\mathrm{PO}}$	-

Figure 6.3: Loss Functions to transition from one spatial relation to another

are satisfiable. However, at the start of the algorithm, all three cones will have some starting relation with each other generated at random and the goal is to reach their target relations. These relations can be:

Let us assume that there exist 3 sets, S, P, and M.

- Disconnect: no 'S' are 'M'
- Partial Overlap: some 'S' are 'M'
- Part Of : all 'S' are 'M'
- Inverse Part Of : all 'M' are 'S'
- Equal Part: all 'S' are 'M' and all 'M' are 'S'

These relations are clearly represented in the Figures below.

Now the transition map is used by the euler neural network to achieve a global loss wherever possible for all 256 Syllogisms. As per the map, let us assume that the given target relation between the two cones is one of 'Proper Part Of' and currently the two cones are 'Disconnect', so in order to move from 'Disconnect' to 'Proper Part Of', we first run a loss function that changes the status of two cones to one of 'Partial Overlap' and then another loss function that changes it to 'Proper Part Of'. Similarly, other relations can be transitioned based on the current status of the two cones.

Clearly, we can always reach a loss of zero between two cones but here we have three cones on a sphere, and many times, it will happen that once we reach a target relation between 2 cones and later decide to work with the third cone, our initial relations will get disturbed so we need to find a way to avoid this problem.

As a solution to this problem, we start with fixing one cone and not changing its position in space. We have 3 cones with initial random spatial relations among each other and their corresponding target relations. So, as a start, we work with cones 1 and 2, fix cone 2, and start by moving and changing the size(if need be) of cone 1. Once the relation is satisfied, we move on to cones 2 and 3, fix cone 2 again, and start by moving and changing the size(if need be) of cone 3. When the loss between cones 2 and 3 is zero, we move toward cones 1 and 3. Again, cone 3 is kept fixed and cone 1 is changed. This entire process is repeated a second time and the third time but during the third try, when we are working with cones 1 and 2, we start with fixing Cone 1 and changing Cone 2.

The given figures 6.6 and 6.7 depicts how the algorithm changes the spatial relations of the cones to satisfy the target relation. As shown in figure 6.6, the algorithm generates 3 random cones on the surface of the sphere. Meanwhile, the output figure 6.7 shows that the set 'S' is partially overlapping with sets 'P' and 'M'. Also, at the same time sets 'M' and 'P' are disconnected. This gives us a global loss of zero and we can say that the given statements and conclusion are satisfiable.

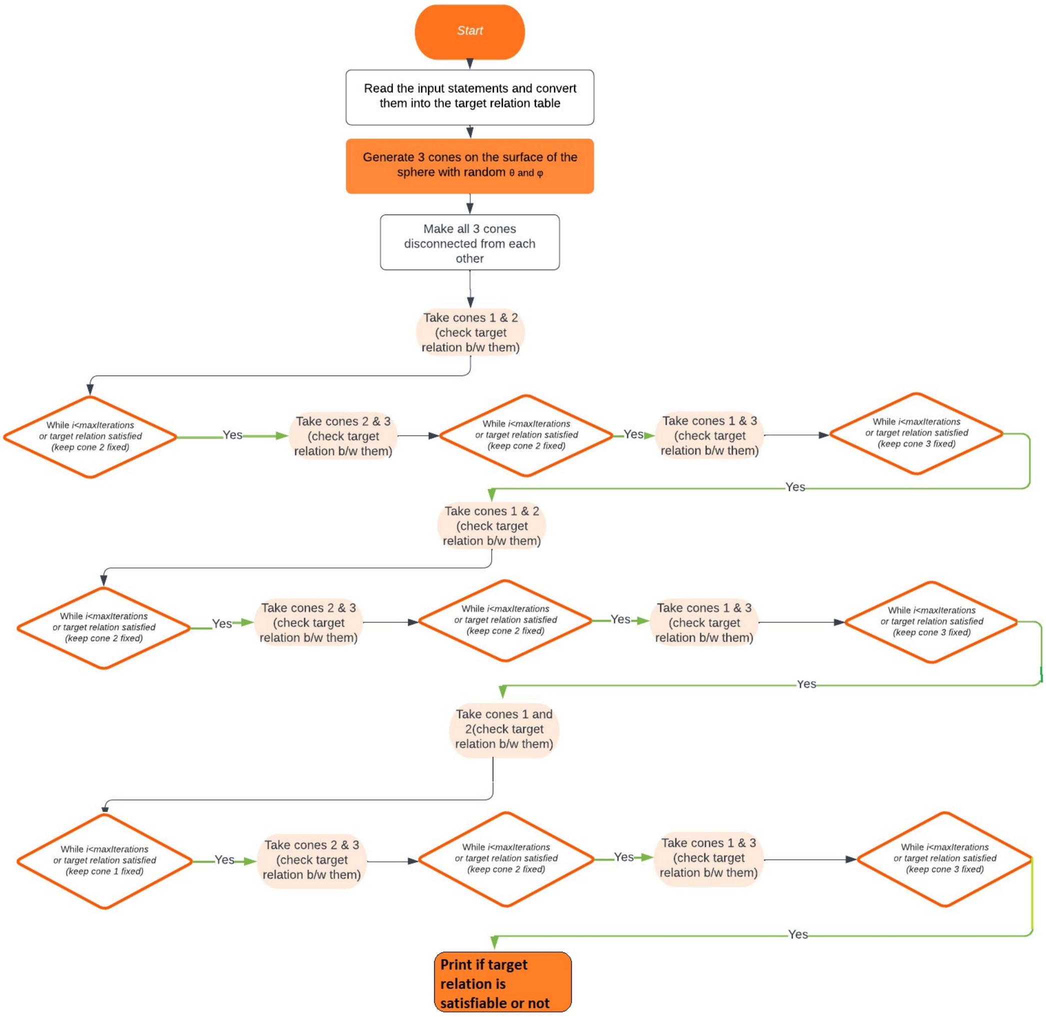

Figure 6.4: Flowchart depicting the control flow of the algorithm

6.6 Speeding up the algorithm for better and faster results

Optimization used in the program to speed up the process and reduce iterations: While working with Part Of relation, the algorithm changes the radius of the smaller cone by a factor of 5.5, i.e. let us assume that cone 1(smaller cone, say alpha angle = 69.99 degrees) just becomes Part Of cone 2(bigger cone, say alpha angle = 70 degrees), the algorithm as an optimization, changes the cone 1 radius to cone 2 radius/ 5.5, cone 1 radius = cone 2 radius/5.5 cone 1 radius = 70/5.5 = 12.72

This helps in speeding up the algorithm as otherwise, the algorithm spends a lot of time reducing the radius of the smaller cone or increasing the radius of the larger cone. Also, specifically the factor '5.5' was chosen as this difference between the radii of two cones helps in solving some complex cases.

The algorithm is written in Python. As for the experiments, the Euler Neural Network uses a learning rate of 0.01 and runs for a maximum of 72000 iterations in total in the worst-case scenario. To add to it, the algorithm uses three tolerance values when transitioning from one spatial relation to another. The tolerance values are basically the minimum distance that must exist between 2 cones to call them disconnected or partial overlap or proper part as shown in figure 6.1. Furthermore, the alpha angle of the cones generated at random is always fixed at the start of the algorithm(70 or 80 degrees) and the radius of the Sphere is óf unit 1.

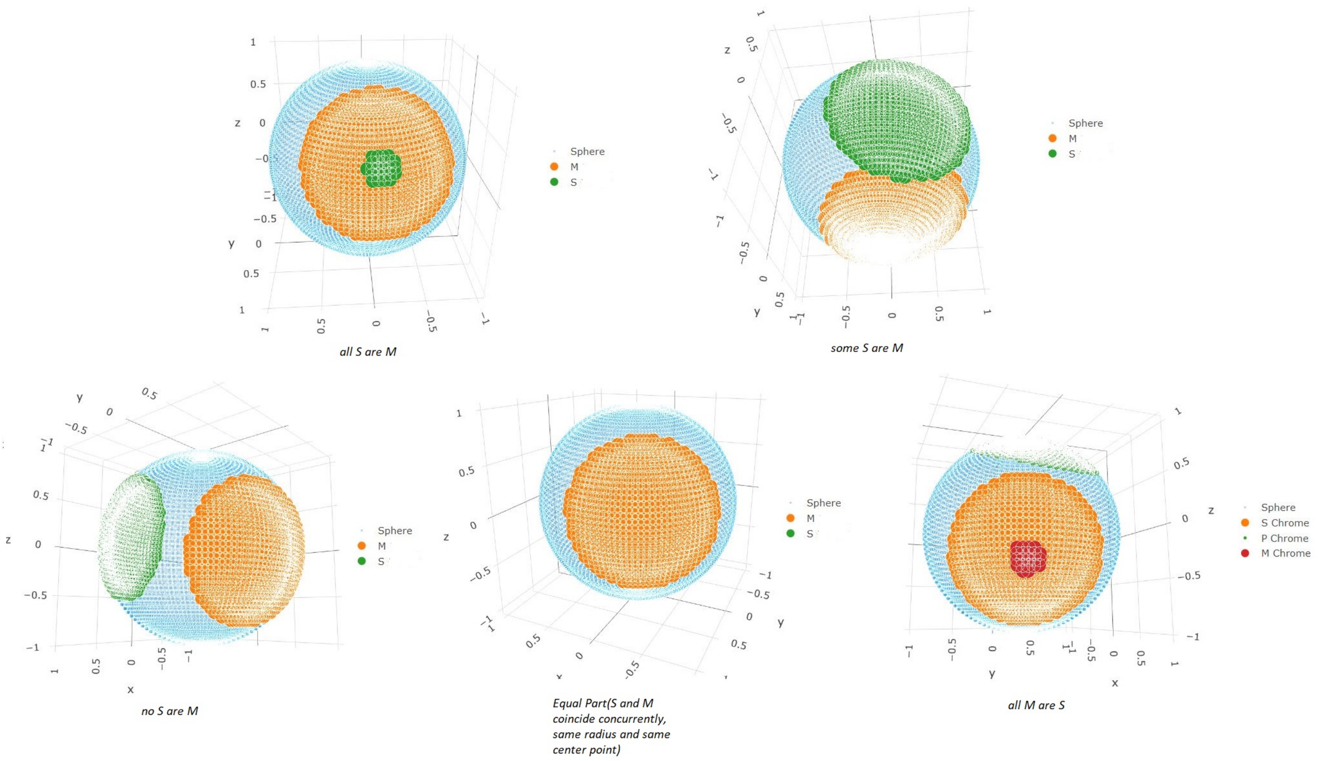

Figure 6.5: Different Spatial relations depicted by the algorithm using Javascript

Algorithm 1 Syllogism on a Sphere

Input: A table of target relations between the cones (Disconnected, Partial Overlap, Proper Part, Inverse Proper Part, Equal Part).
A function *"CreateCones"* randomly chooses values of $\theta(0,2*\pi)$, $\phi(0,\pi)$, and a fixed value of angle $\alpha = 70$ to generate 3 random cones on the surface of the sphere
The target spatial relations among the 3 cones are passed to the algorithm
A simple disconnect relation is run for all cones to make them all disconnected.
while rounds < maxRounds **and** gLoss > 0 **do**
    Run the transition function between Cones 1 and 2 with cone 2 fixed to reach the target relation
    Run the transition function between Cones 2 and 3 with cone 2 fixed to reach the target relation
    Run the transition function between Cones 1 and 3 with cone 3 fixed to reach the target relation
end while
Now, target spatial relations are applied to this disconnected orientation of cones
while rounds < maxRounds **and** gLoss > 0 **do**
    Run the transition function between Cones 1 and 2 with cone 2 fixed to reach the target relation
    Run the transition function between Cones 2 and 3 with cone 2 fixed to reach the target relation
    Run the transition function between Cones 1 and 3 with cone 3 fixed to reach the target relation
end while
The above while loop is made to run for a second time
while rounds < maxRounds **and** gLoss > 0 **do**
    Run the transition function between Cones 1 and 2 with cone 1 fixed to reach the target relation
    Run the transition function between Cones 2 and 3 with cone 2 fixed to reach the target relation
    Run the transition function between Cones 1 and 3 with cone 3 fixed to reach the target relation
end while
The algorithm displays the result as Satisfiable if global loss zero is reached or not.
The end spatial orientation that exists among the 3 cones is presented on a sphere.

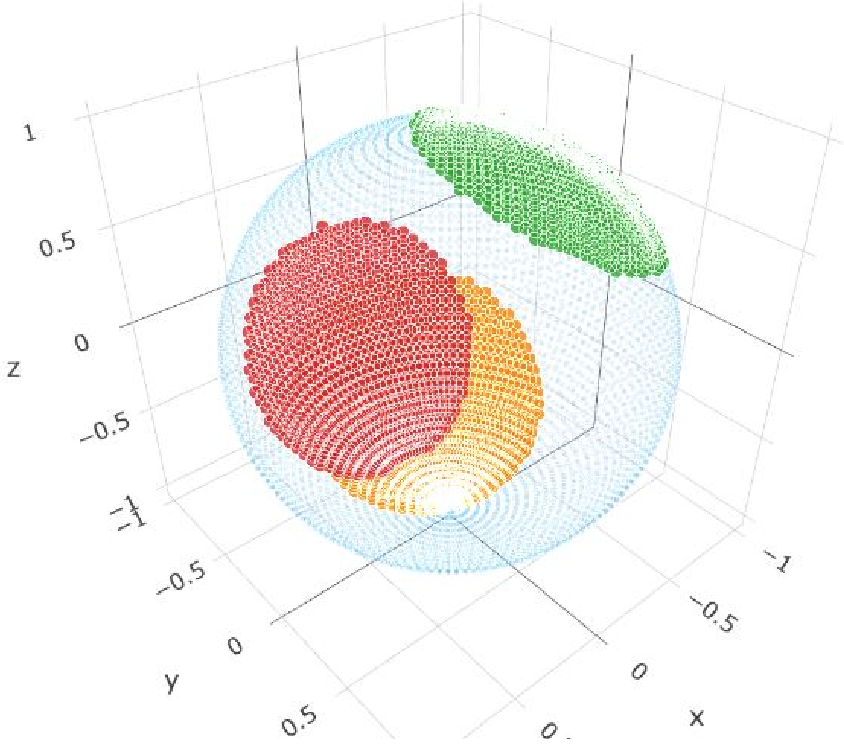

Figure 6.6: Random cones generated on the surface of the sphere by the algorithm at the start(depicted using Javascript)

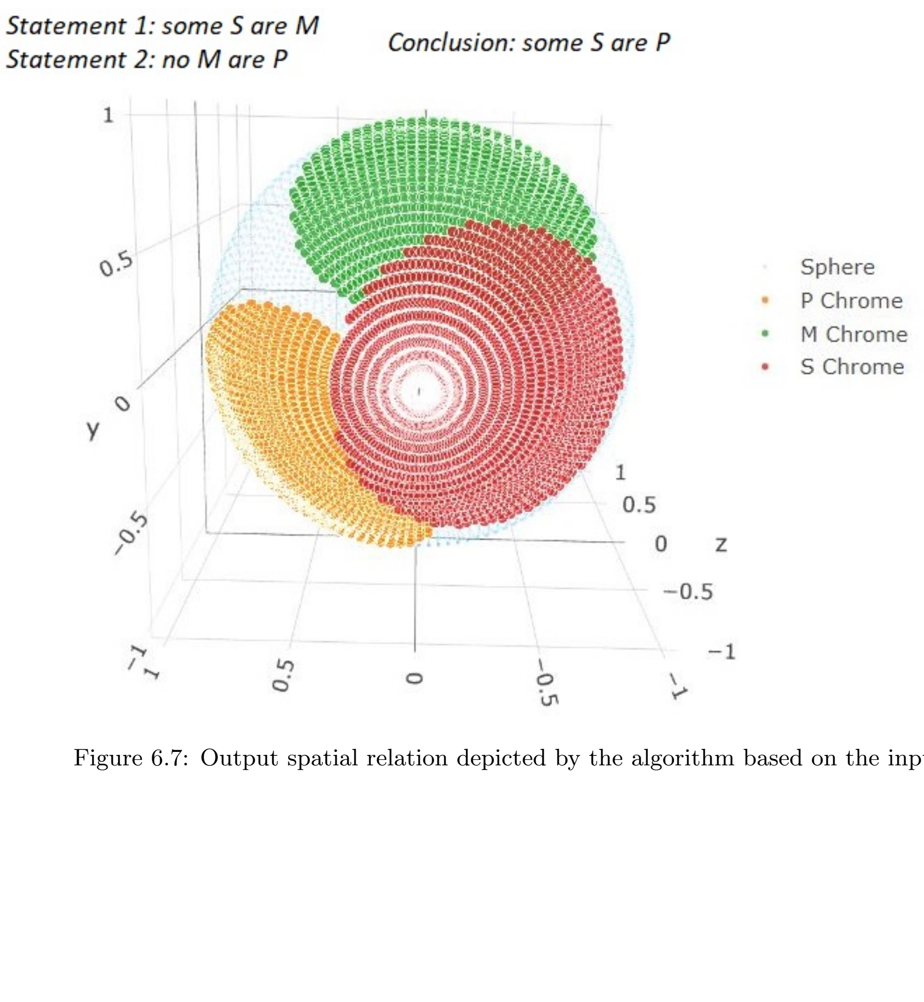

Figure 6.7: Output spatial relation depicted by the algorithm based on the inputs

Chapter 7

Results and Conclusion

The Euler Neural Network on the surface of the sphere runs on all 256 possible syllogisms and gives a 100 percent accuracy of detecting if they are satisfiable or not. It gives 24 non-zero global losses where syllogisms are not possible. To add to it, the benefit of working on a sphere gives us an edge for disjunctive syllogistic reasoning as we can work with complements on a sphere quite easily and then deduce the spatial relations between complements and the given statements and conclusion. As for the complement, if we can represent a set with a cone on the surface of the sphere, then the rest of the sphere is the complement of that set. Furthermore, the network only requires loss functions and the corresponding transition functions and does not require geometric rotations which were needed when working on a 2-D plane. Thus, moving to a higher dimension has helped in simplifying the algorithm and working with disjunctive syllogisms.

In addition, the Euler Network is analogous to how monkeys might do syllogistic reasoning as we have not used natural language as an input for our algorithm, and every computation is based on Euler diagrams and the spatial relations that exist between them and how easy it is to move from one spatial relation to another based on our target relations thus helping us in knowing if the given statements with their conclusion are satisfiable or not.

Bibliography

[1] 3d polar coordinates: https://www.youtube.com/watch?v=ex$_g 2w4e5lqt = 87s$.

[2] Caretsian coordinate system by rene descartes: https://www.encyclopedia.com/education/news-wires-white-papers-and-books/descartes-and-his-coordinate-system.

[3] *Deductive Reasoning: https://www.indeed.com/career-advice/career-development/inductive-vs-deductive-reasoning.*

[4] Difference between euler and venn diagrams:http://denis1251.blogspot.com/2016/06/the-difference-between-venn-diagram-and.html.

[5] *Distance along surface of sphere: https://math.stackexchange.com/questions/1304169/distance-between-two-points-on-a-sphere.*

[6] Euler vs venn diagrams: https://creately.com/blog/diagrams/venn-diagrams-vs-euler-diagrams/.

[7] History of syllogism: https://www.britannica.com/topic/history-of-logic/syllogisms.

[8] Hypothetical syllogism: http://beisecker.faculty.unlv.edu//courses/phi-102/hypotheticalsyllogisms.htm.

[9] *Inductive vs Deductive Reasoning: https://www.dictionary.com/e/inductive-vs-deductive/.*

[10] *Reasoning Definition: https://www.britannica.com/dictionary/reasoning.*

[11] Sorites: https://www.thoughtco.com/sorites-argument-1691977.

[12] Spherical cone: https://mathworld.wolfram.com/sphericalcone.html.

[13] Spherical coordinates on sphere: https://mathworld.wolfram.com/sphericalcoordinates.html.

[14] Types of syllogisms:https://www.indeed.com/career-advice/career-development/types-of-reasoning.

[15] Tiansi Dong. Explainable ai and application, b-it, university of bonn, lecture series.

[16] Tiansi Dong, Chengjiang Li, Christian Bauckhage, Juanzi Li, Stefan Wrobel, and Armin B. Cremers. Learning syllogism with euler neural-networks, 2020.

[17] Cantlon JF Ferrigno S, Huang Y. Reasoning through the disjunctive syllogism in monkeys, 2021.