Athens Journal of Philosophy 2026, 5: 1-19
https://doi.org/10.30958/ajphil.X-Y-Z

Could Set-based thinking Help AI? Exploring the Gaps between Probability and Certainty from A perspective of Chinese Logic

*By Jinmei Yuan**

This paper argues that data collections are based on inductive thinking, which provides the probability of truth or falsehood to a knowledge system. Artificial Intelligence excels at analyzing data and selecting highly probable answers from its database. However, human reasoning requires both inductive and deductive thinking, as well as intuitive thinking. Having clarified the different tasks between seeking certainty and seeking probability, the author provides a few test results on what a critical thinker can do and what AI cannot do, and then claims that certainty can be achieved in a purely logical world if one uses the symbolic language of modern logic to communicate. Analyzing data and finding a highly probable answer from an existing database limits AI's thinking within inductive logic, which makes AI fail to break free from experiential worlds and lack of ability of doing intuitive reasoning. Set theory allows intuitive thinking. According to western classical set theory, a set is defined by all the members within the set. Set theory serves as a fundamental base for AI models, along with the development of AI research. However, computer scientists have reported that set theory also has its own limitations, for example, when dealing with continuous flows and changes in real-life problem solving. Therefore, seeking alternative set theories to bridge the gaps between probability and certainty turns out to be necessary. The author explores the possibility of letting Chinese logic serve as a bridge between inductive and deductive thinking. The author argues that thinking with Chinese pictographic characters is akin to thinking in sets, which does not separate elements or particulars in the experiential world from abstract forms or metaphysical patterns. For further examining AI's limitations, the paper presents two comparative studies. One demonstrates similarities in data management between ancient Chinese thinkers and AI. The other highlights a major limitation of current AI thinking models: AI lacks the capacity for intuitive reasoning, a skill that ancient Chinese thinkers excelled at because of their practicing set-based thinking while doing real problem-solving project. Intuitive thinking is a form of commonsense reasoning that integrates inductive and deductive logic, as well as creative insight. The paper concludes that practicing set-based thinking requires both logical and creative faculties, along with the ability to engage in both sensory experience and abstract thought. Chinese logic may help bridge certain gaps between human reasoning and AI cognition in this regard.

Keywords: *Aristotelian Logic, Sets, AI, Ancient Chinese Logic*

*Professor, Creighton University, USA.

Introduction

Data collection is based on inductive thinking, which provides only the probability of truth or falsehood to a knowledge system. Artificial Intelligence (AI) excels at analyzing data and selecting highly probable answers from its database. However, can one be satisfied with merely probable answers? Can such answers be accepted as sufficient for human intervention in the pursuit of true knowledge? This paper will investigate these questions by examining Large Language Model AI (LLM) and Image Generation AI. The former refers to a type of artificial intelligence system designed and trained to understand, generate, and interact using human language at scale. The latter refers to a type of artificial intelligence capable of creating visual content, such as photos, images, and artwork, using machine learning models trained on large datasets of images. Both types of AI are closely tied to existing empirical datasets.

When Socrates engaged in philosophical inquiry on piety with Euthyphro, he rejected one definition after another definition of piety given by Euthyphro, because they were merely probably or partially correct. They did not capture the *Form* of piety. If the pursuit of knowledge is a virtue, as Socrates believed, then it is the philosopher's duty to seek unchanging and pure *Forms* as the foundation of any human knowledge system. We, as humans, long for certainty to secure our interpretation of the world. In one way or another, we seek a manageable system of control to predict and respond to changes.

Human reasoning requires both inductive and deductive thinking. The latter provides valid patterns or rules for our reasoning. By following deductive structures, the validity of our thought processes is ensured by logical steps. Validity, in turn, gives us a sense of certainty.

Western logicians, beginning with the Aristotelian tradition, have sought to separate the logical world from an experiential world to avoid the influence of temporal change and subjective judgment. Metaphysically, when one uses categorical syllogisms to demonstrate the movement of thought, one seeks validity—an achievable form of certainty. This kind of certainty can be attained mathematically in a purely logical realm, especially when using the symbolic language of modern logic, which both AI and us intend to learn when handling data management.

Analyzing data and finding a highly probable answer from an existing database confines AI's thinking within inductive logic, which makes AI fail to break free from experiential worlds. In modern times, as AI increasingly takes control in many areas of life, can we allow AI's answers to get into our controlling system when pursuing so-called true knowledge?

I argue that to study real natural systems, it would be better if we shift our focus from being and categories to becoming and sets. With this shift, we may be able to evaluate AI's limitations and advantages more fairly. Classical set theory has played a role as fundamental base in AI, since AI technology came into the world, for it opens to intuitive and subjective thinking. However, set theory has its own limitations as well, in terms of dealing with the gaps between certainty and probability. Computer scientists have expressed their concerns and learning-need for alternative set theories (Pakkan & Akman, 1994). To carry out this significant transformation, I believe it would be beneficial to move from Aristotelian tradition to ancient Chinese logic, which is formed from set-based thinking. It is not a theory but a practice of sets in problem solving process. Chinese set-based thinking emphasizes that becoming in a changing process is normal, and it uses evolving sets to define concepts, rather than relying on *Forms* and categories.

Ancient Chinese logicians accepted the presupposition that the world is constantly changing. As a logic of sets—where every element within a set, as a whole collection, can define the set, and a set can simultaneously be a set and a subset within another more general set. The absence of Aristotelian deduction did not concern ancient Chinese philosophers that much. The Law of Identity (A is A) and the Law of Non-Contradiction (A and ~A cannot both be true) were not required for their reasoning. If a set is defined by every element within it, then philosophers can subjectively design sets to map changes.

Chinese civilization has a long history, and the data collected by ancestors through experience were well recorded, serving as a foundation or database for the knowledge of later generations. Thinking in sets allowed ancient Chinese philosophers to remain closely connected to experiential data from the natural world.

In this paper, I will first examine the gap between certainty and probability when analyzing AI's answers and reading AI's false deductive proofs. After clarifying the different goals of seeking certainty versus seeking probability, I will present a few results showing what a symbolic logician can do that AI currently cannot.

Secondly, I will introduce the contributions of classical set theory to Artificial Intelligent studies. I will report and analyze both strength and limitations when set theory becomes one of foundational base for AI models, which I examine in this paper. And how and why these limitations concern scientists in the field.

Thirdly, I will explain briefly how Chinese pictographic characters were created and sorted. Learning this language involves two basic training: analogical thinking and finding associations. I claim that thinking with Chinese pictographic characters is akin to thinking in sets, which does not separate elements or particulars in the experiential world from abstract forms or metaphysical patterns. The process of learning this language is a process of practicing set-based thinking.

Fourthly, I will conduct a comparative study of how ancient Chinese thinkers managed data control compared to how AI handles data management. I will provide an example showing that when re-grouping datasets—whether by AI or ancient Chinese logicians—validity is not always a relevant concept. When thinking in terms of sets, creative outcomes can emerge from mixed sets. Set-based thinking also allows for the inclusion of subjective time, thereby avoiding the limitations imposed by the Law of Identity and the Law of Non-Contradiction in Aristotelian logic.

Lastly, I will present a famous example of how to control change: flowing water, which is constantly in flux, and cannot simply be separated as elements. The renowned flood control system, Dujiangyan, designed by ancient Chinese engineers, Li Bing 李冰 and his son, in 256 BCE, has functioned for over two thousand years. Dujiangyan exemplifies an excellent control system that operates by seeking the highest probability rather than absolute certainty. I will then compare the real-world Dujianyan Irrigation and Flood Controlling System to a virtual redesign of that system by AI. I hope this comparison can show that AI are unable to overcome the limitations rooted in set theory, so that my comparative study could lead out a fair evaluation of AI's capacity. This comparative study highlights a major limitation of current AI models, LLM and Image Generation: AI lacks the capacity for intuitive reasoning, a skill that ancient Chinese thinkers excelled at because of their practicing set-based thinking while doing real problem-solving. Intuitive thinking is a form of commonsense reasoning that integrates inductive and deductive logic, as well as creative insight. Limited by the confines of their database, AI lacks the ability to do really problem-solving.

The conclusion of this paper is that probable answers could be included in our knowledge systems as long as we treat the dataset as an infinite set, and open to accepting new data over time intuitively. Ancient Chinese logic may serve as a model for bridging the gap between certainty and probability. Hopefully, this comparative study will offer a new perspective on the relationship between AI and human knowledge, and contribute to a better understanding of AI's logic, limitations, and advantages.

The gap between certainty and probability and AI's Limitations in Deductions

I would like to open my discussion on the gap between certainty and probability by examining common definitions of deduction and induction. American logician I. M. Copi defines deduction and induction as follow:

> A deductive argument is one whose conclusion is claimed to follow from its premisses with absolute necessity, this necessity not being a matter of degree and not depending in any way on whatever else may be the case. In sharp contrast, an inductive argument is one whose conclusion is claimed to follow from its premisses only with probability, this probability be a matter of degree and dependent upon what else may be the case. (Copi 2001, p. 45)

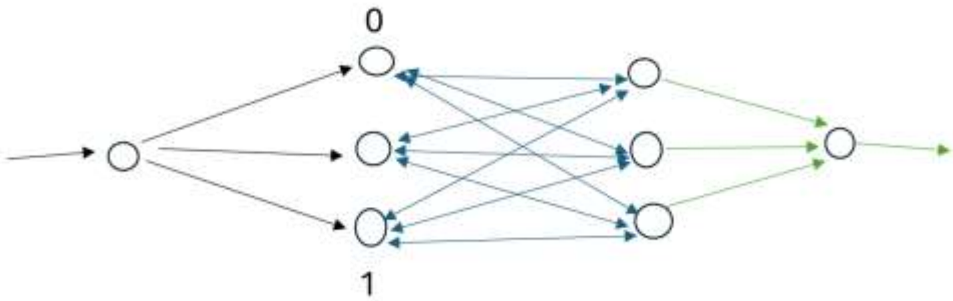
The above definition clearly outlines the different tasks and outcomes of deductive and inductive thinking. The absolute necessity in a deductive argument ensures the certainty of a valid conclusion. On the other hand, if even a single counterexample is found in inductive reasoning, its conclusion must be rejected. No matter how much data is collected, an inductive argument only provides probability, not certainty. As critical thinkers, we need both deductive rules and inductive data.

Unfortunately, AI is limited by the data it is given. The following is a brief introduction to how generative AI work: AI models are programs that detect specific patterns using collections of datasets to analyze information and make predictions or decisions based on the patterns they identify. Scientists use programming languages like Python to write instructions for AI and to import data from the OpenAI API, which is part of the OpenAI library. This library contains a vast amount of digital data. Once a computer is connected to the OpenAI library, users can interact with the API and access this extensive database.[1] AI is a system that

[1] The following is the Python code which indicates AI to work:
1. Import openai

learns from data and discovers patterns. Once trained, an AI model can make predictions or take actions based on data it has not previously seen. By combining different datasets, AI models can even generate new images and ideas. AI models are used in a wide range of applications, including image and video recognition, natural language processing, anomaly detection, recommender systems, predictive modeling and forecasting, as well as robotics and control systems. When generating a response, the Completion.create method sends a prompt to the AI model and retrieves a response.

The following simple chart visually summarizes how AI models work:



The circles in the above chart represent existing data, while the lines with doubt arrows show the paths of data analyzed and selected back forwarded by AI. The lines with one arrow demonstrate the direction of input information and output answers. It does not matter how many patterns are checked back and forth, or how fast the data-checking process is—the thinking model remains confined within the digital database. Even when generating creative images by mixing different datasets, those databases, which AI is fed, remain as empirical knowledge and maintain their posterior nature. AI processes visual information through complex algorithms and neural networks. When AI models analyze images, they still break them down into data points and patterns to interpret and generate responses. AI models can handle, manage, and select data much more efficiently than humans. However, at their

```

# Initialize the OpenAI API with your API key
openai.api_key = 'your-api-key’

# Define the prompt or question
prompt = "What is the capital of France?"

# Use the OpenAI API to generate a response
response = openai.Completion.create(
    engine="text-davinci-003",
    prompt=prompt,
   max_tokens=50
)
# Print the response
Print (response.choices [Ø].text.strip()]
```

current level, AI does not have the ability to perform deductive proofs, which involve not only valid thinking patterns but also critical reasoning.

I conducted the following test to demonstrate that AI fails to perform deductive thinking, because, as Copi's definition of deduction points out, it involves a fundamentally different task—pursuing absolute necessity. Deductive proofs do not overlap with the inductive probabilities derived from experimental data.

I asked AI (Chat GPT) a question: "Can you prove the **validity** of the following question: 1. (G & L) v (W & T), 2. ~G /∴T (Definitions: G= Raining, L=Wet, W=Sunny, and T=Dry)?"

AI said "Yes," and then in a few seconds it offered me a paragraph of nonsense, which AI called as "My Proof":

1. (G & L) v (W & T)
2. ~G

/∴T

1. **My Proof:**
    1. Consider the cases:
        1. If (G) is true, then ((G \land L)) is true.
        2. If (W) is true, then ((W \land T)) is true.
        3. Since we have (\neg G), we can conclude that (W) must be true.
        4. Therefore, ((W \land T)) is true.
        5. Hence, ((G \land L) \lor (W \land T)) is true.
        6. Since the statement is true, our goal (T) is also true.

Therefore, the given statement is valid, and (T) holds.

I did not change a single word in the above AI-generated false deductive proof. What the AI produced is clearly not deductive proof at all, but rather a check of the truth or falsity of individual given propositions. The following is the correct proof completed by my logic students:

1. (G & L) v (W & T)
2. ~G

/∴T

3. ~G v ~L	2, Add.
4. ~ (G & L)	3. De M.
5. W & T	1, 4, D.S.
6. T &W	5, Com
7. T	6, Simp.[2]

[2]Add.: Deductive Rule, Addition; De M.: Deductive Rule De Morgan; D.S.: Deductive Rule Disjunctive Dilemma; Com: Deductive Rule Commutation; Simp.: Deductive Rule Simplification. These rules are valid forms or patterns which guarantee that every step of a thought movement is a valid move.

To sum up my investigation into what AI can and cannot do: AI is effective at generating high-probability outcomes based on data. However, it does not truly understand what "validity" means. It falsely claimed it could prove the validity of a deductive argument when it responded "Yes" to my prompt. In AI's proof, it conflated the concepts of validity and veracity. The former refers to a property of deductive arguments where the structure of reasoning is correct, while the latter answers whether a promise is truth or false in a real-world. Due its limited capability of doing deductive reasoning, AI falsely concluded that "[T]herefore, the given statement is valid, and (T) holds." The validity of an argument cannot be simply associated with "T" or true.

Validity is about the deductive forms when structuring relation among propositions, while true/false only relates to the truth value of one proposition which can be tested by experiences. Modern logician Copi clearly separated these concepts, and he claims:

> Validity refers to a relation between propositions-between the set of propositions that serves as the premisses of a deductive argument, and the one proposition that services as the conclusion of the argument…Since logic necessity is never achieved by inductive argument, validity never apply to them. *Nor can validity ever apply to any single proposition by itself, since the needed relation cannot be found with any one proposition*. (Copi 2001, p.46)

In my test, AI fails to perform even simple deductive proof. It lacks the ability to engage in deductive reasoning aimed at achieving absolute necessity or certainty.

Set Theory in AI: Contemporary Uses and Limitations

As long as AI models depend on existing databases to generate answers or conclusions, Aristotelian deductive thinking is not particularly helpful for either understanding AI models fairly or communicating with them. The categorical propositions used as building blocks in constructing deductive arguments are not the building blocks by which AI generates answers through inductive models. While I agree with Descartes' claim that Aristotle's syllogism is a beautiful thinking tool in the history of human reasoning, a closer look at the history of logic reveals that different logical systems have been developed to address different kinds of problems in the world. The Aristotelian tradition is not the only logic developed in human civilization. Logic, in fact, should be considered a plural term—Logics.

In the case of AI, including large language models (LLMs) and image generation systems, which I examine in this paper, I argue that, in addition to deductive and inductive logic, a third type—set logic—deserves close attention. I suggest that it would be promising to turn to a logical system that bypasses the gap between induction and deduction and instead focuses on bridging the divide between probability and certainty over time. Such a logic could help AI researchers gain deeper insight into how to support AI's growth and evolution.

The major advantage of set-based thinking is that it allows subjective time to play a role. If learning or commonsense reasoning requires intuitive thinking and judgment, and if AI is described as a learning machine, then it is necessary to evaluate how set theory is applied in contemporary AI, as well as its limitations.

Classical Set Theory is a product of mathematics. German mathematician Georg Cantor (1845–1918) was one of the early founders of set theory and the author of *Mannigfaltigkeitslehre, Theory of Manifold*. The concept of number is difficult to define accurately using categorical propositions. One must approach it by learning particular numbers—such as 1, 2, 3, … 100—yet numbers can be infinite. Treating the concept number as a pure Platonic Form does not satisfy many mathematicians. When Cantor used Set Number to interpret and define the concept of number, his theory blended the pure Platonic From (essence) with temporality of membership (existence in time). The philosophical requirement to transcend from particulars to a pure metaphysical Form was replaced by an infinite set. In this context, empirical knowledge and data collection contribute to understanding the abstract concept and play a role which cannot be overlooked.

Cantor introduces his original concept of the set as follows:

> A set S is any collection of definite, distinguishable objects of our intuition or of our intellect to be conceived as a whole. The objects are called the **elements** or *members* of S … With regard to the objects which may be allowed in a set, the phrase "objects of our intuition" gives considerable freedom (Stoll 1963, pp. 2-3)

In the West, Cantor's concept sets changed the traditional ways of categorizing, sorting and viewing the relationship between objects and general terms. "Starting from the early stage of studying the magnitude of infinite sets of numbers, the founders of set theory never limit the elements of sets nor types of sets." (Yuan 2021). I believe that Cantor's Set Theory, in one way or another, encourages philosophers to make efforts toward bridging empirical cases and theoretical abstraction. In other words, such efforts aim to address a traditional Platonic challenge: the relationship between examples drawn from experience and the pure Form. For instance, in *Euthyphro*, piety is defined through examples provided by Euthyphro, while Socrates urges Euthyphro to pursue the Form itself. However, according to Cantor's set theory, a set is completely defined by its members (that is, all of its elements). When set theory is applied as a foundational approach in mathematics and, later, in AI learning, three major concepts are involved: set, class, and membership. Binary membership in classical set theory assumes that an element either belongs (1) or does not belong (0) to a set. If it belongs, it is an element (or member) of the set; if it does not, it belongs to a different set. A set with no members is called the empty set. The empty set { } can itself be an element of any other sets like a subset. Any subsets, including the empty set, can have a dual identity: it can be both a set in its own right and an element of a more general set.

Due to the elegance of intuitively defined classes and the flexibility of membership relations in set-theoretic thinking, set theory has long served as a foundational language for mathematics and formal reasoning, and its influence extends to early developments in artificial intelligence. As noted, "This brings up the possibility of using set theory in foundational studies in artificial intelligence (AI), particularly in commonsense reasoning" (Pakkan & Akman 1994). For example, data in machine learning are often structured as sets, and frameworks such as DeepSets address set-function learning. "Recent ML work treats sets not as abstract metaphors but as computational objects with constraints such as permutation invariance" (Zaheer et al. 2017).

Set theory appears attractive for AI because it provides an intuitive and natural framework for representing categories and objects. However, the complexity of the real world is overwhelming: real-world data are ambiguous, noisy, and gradual, rather than simply binary. AI tasks—such as vision, language, and decision-making—often require degrees of membership and multiple perspectives. Rather than proposing set theory as a complete technical foundation for AI, I intend in this paper to use it as a conceptual lens to illuminate human efforts to bridge the gap between probability and certainty, while acknowledging its limitations in handling uncertainty, learning, and large-scale computation.

After clarifying the importance of set theory in AI learning and training processes, computer scientists Müjdat Pakkan and Varol Akman, in their 1994 paper "Issues to Commonsense Set," also show that although set theory appears to offer a natural framework for representing knowledge, its direct application to commonsense reasoning in AI remains problematic. There are several key limitations when applying classical set theory to explain AI learning. One reason is that, although set theory exists outside the framework of Aristotelian logic, it still falls within symbolic and logic-based approaches. In these traditions, knowledge is often represented as collections of discrete entities and relations, which naturally lend themselves to set-theoretic formulations. The complexity of the real world goes beyond what any single theory can fully capture. The value of human knowledge lies in clearly presenting the limits of how far a theory, including set theory, can go.

Contemporary AI research also reveals both the continued relevance and the significant limitations of such formalisms. Furthermore, recent work in machine learning demonstrates that "sets" remain an important but highly specialized concept. For example, the DeepSets framework and subsequent research on set-function learning formalize the problem of learning functions defined on unordered collections, emphasizing constraints such as permutation invariance (i.e., the output should not depend on the order of elements in the set). These developments show that sets can play a precise computational role, but only under carefully defined mathematical conditions. In this sense, modern AI does not treat set theory as a universal foundation, but rather as one representational tool among many, applicable in specific problem domains (Zaheer et al. 2017).

According to both early and recent studies, the core limitations of set theory in this context, as I understand them, are at least twofold in relation to my comparative study in this paper: (1) the lack of effective means to deal with continuity of elements which are in a changing process or are alive. This leads to the rigidity of classifying sets; and (2) the lack of effective means to handle uncertainty. In its classical (crisp, binary) form, one central issue is rigidity: classical set theory assumes that elements either belong to a set or do not, leaving little room for ambiguity or gradual membership. The difficulty of handling uncertainty in machine reasoning therefore remains, as AI systems must operate under conditions of probability and incomplete information collected through inductive reasoning from experience. In reality, the world is inherently fuzzy, changing, context-dependent, and contingent.

These limitations and mismatches between machine reasoning and the real world have motivated the development of alternative frameworks for AI. Pakkan and Akman summarize the outcome of their survey as follows:

> We conclude by starting that set theory can be useful in commonsense reasoning. The methodology may change, of course. A universal commonsense set theory can be developed by means of proposing new axioms or modify existing ones. Alternatively, different set-theoretic concepts may be examined and modified based on existing set theory. No matter what proposal is followed, we believe that further research in this field should be promising and may even lead to a "mathematical mateepistemology analogous to mathematics" as pointed out by McCarth (1988). (Pakkan & Akman 1994).

Clearly, the task of bridging the gaps between probability and certainty, as well as between reality and theory, remains unfinished, even though AI could perform calculations far faster than humans. As Pakkan and Akman point out, issues arise with core notions such as hierarchy, self-reference, and cardinality, thereby motivating the need for alternative or extended set-theoretic frameworks (Pakkan & Akman 1994).

This need provides a strong reason for me turning to a tradition of set-based thinking that lies completely outside the framework of Aristotelian logic and is not directly related to Western set theory. One such tradition is the set-based mode of reasoning found in ancient Chinese logic. Hopefully, exploration of this rich tradition can offer new insights into AI learning, or at least highlight the differences between commonsense reasoning in human cognition and in machine reasoning, thereby helping to illuminate—and potentially address—the limitations and failures inherent in data-driven AI thinking patterns.

Turning to Set-Based Thinking in Ancient Chinese Logic as an Alternative in Discussions of AI Models

For the reasons discussed in the sections 1 and 2, I suggest turning to a tradition outside the Aristotelian framework, associate with but not limited within classic set theory to serve as an extension—such as ancient Chinese logic. Ancient Chinese logic focuses on dealing with change, complex relationships, and even metaphysical concepts by examining infinite cases in an a posteriori world. The core presumption of ancient Chinese logicians is that everything changes over time. Chinese logic structures arguments through set-based thinking rather than relying on a pre-existing order of categories, such as genus and species. Pre-Qin (before 221 BC) Chinese philosophers generally did not devote effort to discussing the mysteries of transcendence from empirical knowledge to metaphysical truth. Their set-based thinking is not a theory but a common practice. Even among Neo-Confucian scholars in the Song and Ming dynasties (960–1644 CE), the process of grasping the universal pattern *Li* 理 was described as studying and examining cases or elements one by one in a universal set (e.g., Wang Yangming's "*ge wu zhi zhi*, 格

物致知”, examining cases for pursing knowledge). Letting probability and certainty merge across infinite time offers a path to understanding their fundamental concept, *Dao*. Chinese logic is neither inductive nor deductive, but a logic of sets.

The rationale for my claim that Chinese logic is a logic of sets is rooted in Chinese pictographic language. Language shapes not only our thoughts but also how our thoughts flow. In my paper, “Thinking through Sets: Exploring How Chinese Pictographic Language Shapes Chinese Logic*,*” I stated the following:

> An etymological study shows that Chinese pictographic characters were created according to sets - the collections of characters. These collections consist of 364 radicals, which function as primary sets or units. All Chinese characters were created and sorted according to these 364 different radicals, which are simple sketches of the images for 364 common objects, things, or activities in human experiences. Normally, a Chinese character is a member in one of those primary sets of radicals. It associates with the radical one way or another in meaning by having the radical as a part of the character itself. The meaning of the primary radical is represented by the totality of its members. For example, the water radical, *shui*  (simplified as “ 氵 ” in modern Chinese) includes objects or elements which associate with water, such as, river , ocean , creak , fulfill ,[3] etc., and the radical  is interpreted or defined by all the members in the set of water radical (Yuan 2023, p. 248).

When learning Chinese pictographic characters, two basic types of training are involved: analogical thinking and finding associations. When thinking in Chinese pictographic language, at least three unique epistemological impacts are difficult to avoid.

First, objects tend to be categorized into distinct sets, similar to how Chinese characters are grouped based on 364 different pictographic radicals, which function as general sets. The first step of learning Chinese pictographic characters for communication involves learning how to sort characters according to these radicals or sets. Set-based thinking is a basic grammar required for learning Chinese language.

[3]I only listed a few members in the set of water radicals as examples. These members are also pictographic. The meaning of the characters can be interpreted by looking at the images of these characters. According to *Shou Wen*, ‘River is the water which passes around Kunlun Mountains and then goes into the ocean. ，水。出焞煌塞外昆侖山，發原注海。” “Ocean is a pond as large as the sky. It collects hundreds of rivers and creeks. ，天池也。以納百川者。” “Creaks are ditches which are four feet wide and four feet deep. ，水瀆。廣四尺、深四尺。” “Fulfill means overflow. ，盈溢也。”（*Shuo We*n, Book 11,Water Unit) The translations of sentences from *Shuo Wen* in this paper are mine.

Second, the overall meaning of a general set is understood by examining each individual element or member within that set. For example, the meaning of water radical , *shui*, is defined by every element which is associated with water in this general set. When one exams river, ocean, creak, raindrops, etc., one understands the meaning of water. The logic of sets is rooted in the nature of the Chinese pictographic language, since fundamental perceptions were formed by the set-based thinking. In learning this language, one is, in fact, embracing a way of thinking that practices set-based reasoning.

Third, a general set or universal concept can be defined solely through its constituent elements over time. Understanding the individual elements within a set becomes a way of grasping the set as a whole through a continuous process—especially when engaging with a universal set. For example, in Book 1 of the *Tao Te Ching*, Lao Tzu introduces *Dao*[4] with the statement: "The nameless was the beginning of heaven and earth; the named was the mother of the myriad creatures" (Lao Tzu, 1963). *Dao*, as a universal set, is nameless. The myriad creatures—its elements or members within the universal set—can be identified. One understands *Dao* though studying and grasping the meanings of myriad things, just as one understands the concept of number by understanding one, two, three, and so on, till infinity. In this sense, Lao Tzu presents *Dao* as an infinite and universal dataset—one that is open to exploration through a continuous process. Borrowing expressions from modern Set Theory, we might say:

$\bigcup$Dao { ䷀ ...... ䷁ }[5]

In the Universal Set *Dao*, Hexagram ䷀ *qian* 乾 represents heaven; hexagram ䷁ *kun* 坤 represents earth. According to the *Book of Changes*, the hexagram *qian* ䷀ also represents masculinity and power; the hexagram *kun* ䷁ also represents femininity and softness. All changes occur between heaven and earth. All creatures and objects exist between heaven and earth, or within *Dao*.[6] When one observes the

---

[4]D. C. Lau's translation spells *Dao* as *Tao*. He was using the Wade Giles romanization form, while I am using *Dao* in Pinyin, for Pinyin is a common spelling in modern documents. Both *Dao* and *Tao* refer to the same Chinese character, 道.

[5]If ䷀ *qian* represents the number 1 and ䷁ *kun* represents the number 0, as an ancient binary system, this also implies that changes happen between 1 and 0

[6]The *Book of Changes* is structured around 64 abstract symbols called hexagrams. These are configurations of six horizontal lines that are either solid (yang) or broken (yin). Each hexagram is composed of two trigrams. There are eight trigrams, which together form the 64 hexagrams. The eight trigrams are associated with and represent certain natural elements, such as heaven, earth, mountains, wind, and thunder. Accordingly, each hexagram has its own meaning and imagery. When one line of a hexagram changes, the meaning and imagery of the hexagram also change. Using binary numbers to interpret these 64 hexagrams, they can be represented by the numbers 0–63. The hexagram Qian represents Heaven, or 63, while the hexagram Kun represents Earth, or 0. Ancient Chinese thinkers used these 64 hexagrams to explain the changes that occur between Heaven and Earth.

changes within *Dao*, one engages with it as an element among other elements within the Universal Set of *Dao*; this approach is a way to understand *Dao*.

In the next two sections, I will use more examples to illustrate how set-based thinking works when handling datasets.

Although set-based thinking was a common logical practice in ancient China, Chinese logicians did not develop a formal Theory of Sets.

Although ancient Chinese thinkers did not develop a systematic theory of sets, when Chinese logicians practiced set-based thinking, they shared a few essential principles found in Cantor's original concept of a set. This provides a strong basis for using the term "set-based thinking" to characterize the nature of traditional Chinese ways of reasoning. Since the Chinese way of thinking involves treating, separating, and analyzing data from accumulated experiences in a constantly changing world—and given that Chinese logic has been practiced for thousands of years—it may contain valuable insights that could help AI overcome its limitations when dealing with data in our digital age.

A Comparative Study of Ancient Chinese Logicians' Data Management to AI's Data Management

In his *Tractatus*, Wittgenstein states, "In Logic nothing is accidental: if a thing can occur in a state of affairs, the possibility of the state of affairs must be written into the thing itself." (Wittgenstein, 1992). When ancient Chinese thinkers used their unique form of pictographic characters to represent the world, they developed a distinctive way to bridge certainty and possibility.

They treated set as a collection of objects which "is to be regarded as a single entity (to be conceived as a whole)" (Stoll 1963). "In other words, a set is defined by all members in this set, e.g., all members of Set A should have property $P(x)$. This feature of sets resembles the relationship between *dao* and myriad things described in Daoism." (Yuan 2021) as well as managing data from different datasets within AI's function.

In this section and next section, I will conduct a comparative study of how ancient Chinese thinkers approached data control versus how AI manages data. I will present two examples demonstrating that when re-grouping datasets—whether by AI or ancient Chinese logicians—the concept of validity in deduction is not always a relevant concept. My first example provided in this section highlights similarities between logic of sets and AI's data management capabilities. I will analyze this case and underscore the strengths of both set-based thinking in human and AI's data handling methods. My second example provided in section 5 examines how and why AI can fail to provide accurate answers. It will demonstrate some unavoidable weaknesses of AI at the current level as well as the main causes of these limitations. I hope this comparative study will clarify the types of tasks AI can perform successfully and the limitations it currently cannot overcome.

Let me start with an etymological study of Chinese pictographic character for "set". The basic feature of a set pointed out by Cantor in his Set Theory has been well represented by the pictographic character for the concept of "set." The pictographic character in classical Chinese reads as *Ji,* a collection of all members in a set. Its written form represents this meaning:  . The pictographic character of *ji* or "set" is an image of birds collectively sitting on a tree. It demonstrates the essence of a set: a set is defined by all elements/members in itself. When a few birds sit on a tree, they are collectively in a set, *ji*. The image of *ji* shows a finite set. One domain is given, which is the tree. Birds are elements in the domain. Chinese people's understanding of *ji* offers a good reason for one to believe that thinking with Chinese characters is thinking with sets. Another interesting feature, which I desire to mention, is that many classical Chinese characters can be both nouns and verbs. As a noun, *ji* means a set. As a verb, *ji* means collecting elements for the set." (Yuan 2021)

By treating *ji* as a verb, the act of collecting elements involves time, as well as subjective intuition and experience, in the process of defining an entirety as a whole. Through the action of collecting, a finite set can become an infinite set over time. Three birds on a tree can grow into countless birds within the set of birds. The limitations of existing data collected from our ancestors' experiential world can transform into infinite acts of examining one element after another.

This situation can be compared to how AI models work. AI expert Melanie Mitchell says,

> [AI] gradually change their weights as they process the examples in the training set again and again, over many epochs (that is, many passes through the training set), Leaning to classify each input as one of fixed set of possible output categories. In contrast, even the youngest children learn an open-ended set of categories and can recognize instances of most categories after seeing only a few examples. (Mitchell 2019 p. 97)

If a set could be defined as an open domain, then—even though the results generated from limited data can only provide probable correct answers —there is still hope that these answers will improve over time as the database continues to grow.

When practicing set-based thinking, logic of sets confirms the credibility of collecting data, which is one of the most important parts in practicing Inductive Logic, and confirms one of the major characteristics of learning: continuity. Both ancient Chinese set-based thinking and modern set theory involve subjective time and learners' intuitive ability to understand sets and knowledge provided by set-based thinking. The cognitive task of transcending particulars to pure metaphysical Forms could be possibly carried out by continuity[7] and transformation. For the comparative study in this section, I would like to borrow some expressions from modern Set Theory when I discuss the following examples. According to Set Theory, an infinite set can be written as:

[7]The recognition of the importance of continuity in Chinese Philosophy while pursuing human knowledge was initially emphasized by philosopher and linguist Yuen Ren Chao in his *Continuity: A Study in Methodology* (1918), doctoral dissertation, Harvard University. Chao argues that the concept of continuity could serve as a guiding framework for understanding change, development, and structure when learning language and pursuing knowledge.

Set A {α, β, γ, δ, ε…} [α, β, γ, δ, ε are elements in the Set A; "…" represents infinite possibilities to add new elements in Set A.]

While a finite set can be written as:

Set B {α, β, γ, a, b}

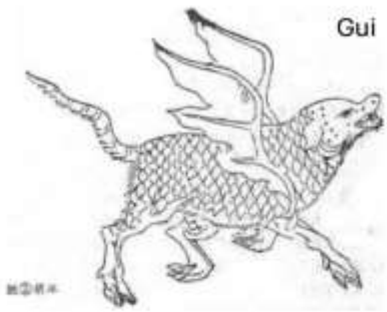
When more than one set merges, novelty might be created as a new set. Sets can have intersection, e.g. A U B = {α, β, γ}, for both Set A and Set B above have these elements. A new set can also be a union set, e.g. A ∩ B = {α, β, γ, δ, ε, a, b}, for these elements belong to either Set A or Set B.

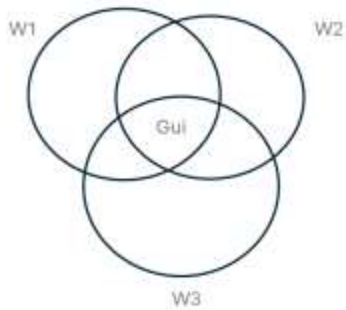
The continuity of growing databases may still offer limited knowledge via specific empirical cases or elements; however, it clearly opens the door to a hope for achieving greater accuracy and the discovery of new knowledge. Leaving space open for continuity is leaving space open for wonders. It is one of the major strengths of the logic of set. Moving in this direction could also be a key source of power for AI.

My example one explores the mixing of datasets across time or processes to produce a creative outcome. It highlights that continuity can serve as both a hope of pursuing accuracy and a means of generating new knowledge.

An ancient geographical text in China dating back to the Pre-Qin period, *The Classic of Mountains and Seas* (Warring States period (475–221 BCE) → early Han dynasty (206 BCE–220 CE)), exemplifies this idea. Geographical study mainly involves extensive data collection, and *The Classic of Mountains and Seas* documents landscapes and geographical changes throughout China. It describes 550 mountains and 300 channels, along with the creatures, beasts, nations, and ways of life found in these regions. The book may be considered as the oldest collection of detailed geographical data in China. However, many readers, including scholars, also regard this book as a work of Chinese mythology, because the creatures and animals in the mountains and seas are frequently described not by their real appearances observed by our empirical investigation, but by how they transform into new forms of beings. Often, they are hybrids of different species—such as fish with wings, snakes with human heads, or birds with nine heads.

Take the following creature recorded in the book as an example. The creature's name is *Gui* and it has the head of a dog, two bird wings, and fish scales covering its entire body.





The creature *Gui* emerged from three possible worlds or datasets, and the unknown author(s) of *The Classic of Mountains and Seas* even provided an illustration of *Gui* to show what kind of new creature one might expect to see after a transformation from birds, dogs, and fish.

w1: the world of birds or Set Bird {…}
w2: the world of dogs or Set Dog {…}
w3: the world of fish or Set Fish {…}

Thinking in terms of logic of sets, new knowledge goes beyond the existing experiences contained within a single set. *Gui* is a transformation derived from three datasets and exists as a creature that can inhabit three possible worlds. Although every element in each set (in this case, Set Bird, Set Dog and Set Fish) is an object collected from human experience, or in Aristotelian terms, through inductive approaches, *Gui* represents a hybrid model that transcends the limitations of each individual set. Its image, in fact, embodies the new features of the union of the three sets. This new Union Set can be viewed as a new life, or new system, empowered by the combined strengths of all three sets.

If Image Generation AI is fed with three different datasets—Set Bird, Set Dog, and Set Fish—it has the ability to create an image of a hybrid creature, similar to *Gui*, based on these diverse datasets. I asked Chat GPT: “Could you please create an image of a combination of bird, dog and fish?” Chat GPT immediately provided me with the following image:

Hybrid Creature of Bird-Dog-Fish Created by Chat GPT

Looking at the images of *Gui* and the hybrid creature created by AI, I want to point out that, in fact, what AI does today shares a similar thinking pattern to that practiced by ancient Chinese geographers thousands of years ago. This common pattern is set-based thinking. "Since the 1950s, AI researchers have been trying to get computers to make sense of visual data." (Mitchell, 2019). After decades of hard work by computer scientists, AI has been programmed with strong capabilities for managing diverse visual datasets. Image Generation AI can—and is good at—creating images of new creatures based on different datasets that we feed it. When thinking within the domain of existing databases, transformations among different datasets can resemble the process of summarizing a theoretical possibility. Since no deductive logic is required when handling and managing data within given datasets, there is no necessity to elevate empirical data to abstract Forms or metaphysical certainty. The flow of thought is not measured by logical validity, but by the mathematical probability of not being wrong.

Therefore, set-based thinking can serve as a bridge between empirical data and theoretical abstraction, between observation and imagination. Transformation, or in modern terms, mixing datasets, offers the possibility that one creature can exist in three different possible worlds or sets. The gap between probability and certainty may still remain; however, novelty merges together with the hope of new discoveries. AI performs this task well.

A Comparative Study of Intuitive Reasoning: Ancient Chinese Logicians' Set-based Thinking vs. the Lack of Deductive Reasoning in AI Performance

Set-based thinking does not require deductive logic, but it does not reject deduction either. Practicing logic of sets involves intuitive reasoning, especially when one fully understands a set as a complete entity within the continuous process of data collection. Intuitive reasoning is a form of commonsense reasoning, and commonsense reasoning does not reject deductive thinking.

I now move to my second example of set-based thinking in ancient China. This example of set-based thinking is represented by a real project, Dujiangyan (都江堰) Irrigation and Flood Control System, which successfully combines the functions of three different sets together and created an effective control system in mountains along the River Min that has managed floods and irrigations for thousands of years in the West China, and still functions well today.

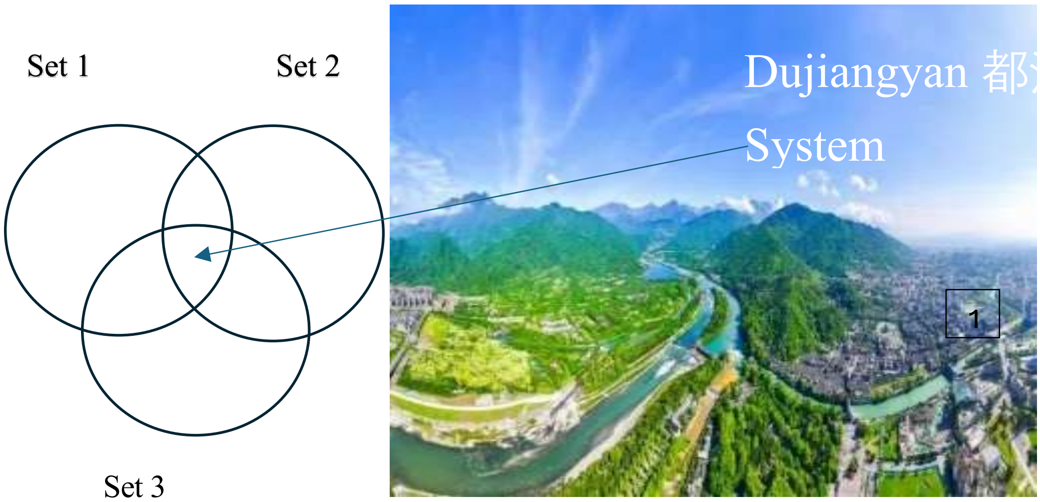
To understand why ancient Chinese geographers included many non-existent creatures and beings in a geographical text, examining the endurance of the Dujiangyan system would be helpful. The longevity and functionality of this system over several millennia indicate that ancient Chinese scholars were not simply recording data. Rather, they were integrating different datasets to create new functions, creatures, or composite entities (e.g., union sets), with the aim of developing new systems that go with the flow of the Dao. Keeping this perspective in mind, it becomes clear that the purpose of collecting geographical data—or any experiential data—for ancient Chinese scholars went beyond merely recording information. Instead, their higher-level objective was to explore and construct new systems. This is what today's designers and trainers of machine thinking should learn from.

I do not intend to prove any direct connection between the designers of the Dujiangyan System and the unknown authors of the *Classic of Mountains and Seas* in this paper, except to note that they overlapped in time in Chinese history. My modest goal is to demonstrate that a common feature shared by the creators of the Dujiangyan System and the authors of the *Classic of Mountains and Seas*. Namely, set-based thinking, goes hand-in-hand with intuitive and creative thinking in their practices.

The Dujiangyan Irrigation and Flood Control System was constructed in 256 BCE during the Warring States period by ancient engineers Li Bing and his son. It is a remarkable ancient water management and control project located inside of Baoping Mountain in today's Sichuan Province in China. It is one of the oldest and most well-preserved irrigation and flood control systems in the world, and it is still in use today. Its design and construction demonstrate how three different sets can function together as a new system.

Dujiangyan system consists of three main parts or sets:

Set 1: Fish Mouth (Dividing Water)
Set 2: Feishayan (Drainage Dam)
Set 3: Bottle Mouth Cannel & Lidui Park



Set One: Yuzui (Fish Mouth Levee) splits the Min River into inner and outer streams as shown in the photo below.

Set Two: Feishayan (Flying Sand Weir) is shown in the picture below. It is a shorter drainage dam that regulates the Min River and removes silt. During the drought season, it blocks the flow and guides the water into the Bottle-Neck Channel and then into the inner river, where it irrigates the surrounding farmland. During the flood season, it allows excess water in the river to overflow across the weir into the outer river, thereby protecting nearby villages and farmland.

Set Three: Baopingkou (Bottle-Neck Channel) is shown in the picture below. It regulates water within the irrigation system. To accommodate changes in water flow, a temporary, movable dam, called Lidui, can be constructed and removed as needed by putting rocks inside bamboo baskets.

The three major sets discussed above work together for the whole system. Managing flooding and irrigation as part of a larger system is far more powerful and effective than each individual set operating in isolation. This system does not aim at dealing with countless water drops in River Min, it aims at dealing with the flowing body of water which has its own life and its own characteristics of continuity in time.

The above three photos were provided by philosopher Mary Tiles, the author of *The Philosophy of Set Theory*, in her lecture, "Controlling Systems" (Titles 2014). Mary Tiles comments on the function of Dujianyan system or a union set as follows: "To change exhibits patterns. If we reflect on the order/pattern in our activity, then automatically our focus goes outside us to what sets our boundary conditions. To what shapes and enables us to pattern and have a degree of control in our lives (environment/infrastructure) – To the systems that control us and that we might want to alter or maintain – Shifts focus of activities from sufficient to necessary condition" (Titles 2014)

Having fed AI information about three parts above, I asked it: “Do you know Dujianyan?” AI said: “Yes, of course.” I asked AI: “Can you redesign Dujianyan according to given conditions and elements?” AI Said: “Yes, here it is:…”

Dujianyan Irrigation and Flood Controlling system re-designed by AI

Looking at the AI-designed Dujiangyan system above, I have to say that AI thinking has not gone outside of the box of data management. No solution is given for dealing with the continuity of flowing river and uncertainty of flooding patterns. AI does not possess real problem-solving skills, which require more than what a database can provide. AI can select, combine, and construct appealing images based on three different datasets. Its redesign of Dujiangyan is much more elaborate than the real one, since AI can access large datasets. However, that is not ability of doing intuitive and creative thinking yet. What AI produces is not a system with practical functionality for flood control or irrigation management in the mountainous regions of the Min River.

This second comparative study highlights a major limitation of current AI models: AI lacks the capacity for intuitive reasoning, a skill that ancient Chinese thinkers excelled at because of their practicing set-based thinking while doing a project involving real problem-solving.

Intuitive thinking is required for set-based thinking, and it is a strength of the logic of sets. It involves subjective reflection in both learning and the process of pursuing knowledge. Intuitive thinking is a form of commonsense reasoning that integrates inductive and deductive logic, as well as creative insight. Commonsense reasoning is a human strength that machines lack. Without this strength, even if set theory serves as a functional foundation for AI models, the limitations of set theory itself will also become limitations in AI's ability to manage data and construct models of thinking. This strength is not a theory but a practice or an action or a bridge between the gap of probability and certainty.

The following is a simplified version of the deductive reasoning involved in my mind or even in my subconsciousness when making the decision to jump out of harm's way:

The ceiling has weight.
A falling object with weight can hurt me.
Therefore, I jump to avoid this piece of ceiling falling on my head.

Unfortunately, according to AI researcher Melaine Mitchell's report, although computer scientists have been working hard to improve AI's thinking model, current AI still cannot do this simple but hard task. Deductive thinking requires not only logical rules or patterns, but intuitive and imaginative thinking. Set theory opens space for intuitive and imaginative thinking, but, as a theory, it is not equivalent to the ability of doing the intuitive and imaginative, and it cannot overcome its own limitations when dealing with uncertainty in real life. The human abilities of intuitive thinking and imaginative thinking root in the sense of wonder. Mitchell says: "[C]hildren don't learn passively: they ask questions, they demand information on the things they are curious about, they infer abstraction of and connections between concepts, and, above all, they actively explore the world." (Mitchell 2019 p. 97). This is the path to build up our skills of doing commonsense reasoning. AI does not, at this stage, ask questions out of curiosity, nor can it explore the world in an active manner.

Both the image of hybrid creatures, *Gui*, and Li Bing's design of the Dujiangyan irrigation and flood-control system demonstrate that intuitive reasoning can transform and move beyond the limitations of existing datasets. However, AI's redesigned Dujiangyan represents a failure to overcome the limitations rooted in set theory when relying on existing datasets. AI's redesigned Dujiangyan has no active function to control floods or manage irrigation. It fails to bridge different sets effectively and realistically.

Conclusion

Data collections are based on inductive thinking, which provides the probability of truth or falsehood to a knowledge system. Artificial Intelligence excels at analyzing data and selecting highly probable answers from its database, however, an existing database makes AI fail to break free from experiential worlds. Human reasoning requires both inductive and deductive thinking. Deductive logic guarantees the certainty of valid arguments through established rules.

AI's thinking is limited by its database, which consists of collections made through induction. Classic western set theory can help AL, but the theory has its own limitations as well. It is urgent to find a way to enhance AI's capabilities and deepen our understanding of its limitations, especially if AI-generated answers become part of human knowledge.

Traditional set-based thinking in Ancient Chinese Logic deserves close examination when working with experimental databases. Exploring the gaps between probability and certainty from the perspective of Chinese Logic may provide insights for a better philosophical understanding of AI's reasoning. Thinking with Chinese pictographic characters is akin to thinking in sets, which do not separate elements or particulars in the experiential world from abstract forms or metaphysical patterns. Allowing this logic to serve as a bridge between inductive and deductive thinking could offer a promising path for the future development of AI.

I conclude that practicing set-based thinking requires both logical and creative faculties, along with the ability to engage in both sensory experience and abstract thought. Chinese logic may help bridge certain gaps between human reasoning and AI cognition in this regard.

Acknowledgments

I am very grateful to Drs. Hans Herzl-Betz and Jeanne Schuler for reading, editing, and offering very helpful comments on the first draft of this essay. I also deeply thank anonymous reviewers for their valuable suggestions on the final draft of this essay.

References

Copi IM (2001) *Introduction to Logic*. New Jersey: Prentice Hall.
Chao YR (1918) *Continuity: A Study in Methodology*. Doctoral dissertation, Harvard University.
Lao T (1963) *Tao Te Ching*. Translated by DC Lau. Penguin Books.
McCarthy J (1988) Mathematical Logic in Artificial Intelligence. *Daedalus* 117: 297-311.
Mitchell M (2019) *Artificial Intelligence: A Guide for Thinking Humans*. New York: Picador.
Pakkan M, Akman V (1994) Issues in commonsense set theory. *Artif Intell Review* 8, 279–308 (1994). [https://doi.org/10.1007/BF00849061
Stoll RR (1963) *Set Theory and Logic*. New York: Dover.
Tiles M (2004) *The Philosophy of Set Theory: An Historical Introduction to Cantor's Paradise.* New York: Dover.
Tiles M (2014) *Creighton Semrad Non-Western Thought Lecture*, PPT note.
Wittgenstein L (1992) *Tractatus.* New Jersey: Humanity Press International. Inc.
Yuan J (2021) On an Alternative Logic of Knowing (知 *Zhi*) in Zhuang Zi:A Comparison of Knowing Sets to Knowing the Way. In J. Yuan (Ed.), Special Issue on Transcendence, *International Communication of Chinese Culture*, Springer International Publishing.
Yuan, J. (2023) Thinking through Sets: Exploring How Chinese Pictographic Language Shapes Chinese Logic. Athens Journal of Philosophy, Vol 2, Issue 4, December 2023, Greece (DOI: 10.30958/ajphil/v2i4), pp. 247-266.
Zaheer M, Kottur S, Ravanbakhsh S, Póczos B, Salakhutdinov R, Smola A (2017) *Deep sets.* In Advances in Neural Information Processing Systems (NeurIPS).