NIND Technical Whitepaper

John H. Doe, Paul Lee, M.D., Reeyan Lee

Abstract — The current wave of AI development, in the form of deep learning neural networks, does not provide a transparent system of operation nor an intrinsic method of linear logical reasoning. Proposed here is a new paradigm of information which utilizes a novel symbolic model that addresses the weaknesses of symbolic AI, and whose operation can be audited step by step. This new model can be utilized to form a reasoning engine that is able to think logically as human beings do, thus, creating an AI that is truly intelligent.

1. Introduction

Today, the waters of artificial intelligence are mainly filled by the relatively recent upsurge in deep learning, and neural networks in general, known as connectionist AI [1]. These systems demonstrate great capability for specific domains [2], outdoing world-class human experts in such areas as game playing [3] and visual recognition [4]. Two common factors in developing such "intelligences" are deep neural networks and massive amounts of data to train those networks. What forms when these technologies are implemented are black boxes, where anything resembling reasoning is represented by matrices of numbers, and the resulting "knowledge" is generally inapplicable to any other domain but that of a specific one [5]. In this paper, in contrast, we define a new paradigm that invokes the precision of symbolic AI along with methods to overcome symbolic systems' weaknesses. Definitions and functions, when implemented in our system, are perfectly transparent and we may see, at every step, why the reasoning process took the steps that it did to reach its conclusion.

Trying to correct a faulty, many-dimensional problem solved by a neural network is a challenge, as it is impossible to find exactly where the error lies [6]. However, in this novel approach that we propose, it would be perfectly reasonable to discover and surgically modify the main source of the mistake. The transparency of operation is important for domains in which verifiability is critical, and with this transparency, we are able to apply the same methods in other domains with ease. In addition, we can also observe whether the problem it solves is the actual problem we were looking to solve, rather than some coincidental solution. Biases could also be traced through the logic in our system. We would like to know the reason why, whenever the reasons may be relevant.

This paper describes a new method of machine comprehension, using the basic "unit of reasoning" which we define as a *canonical*. This method of machine comprehension is fully transparent in its operation and capable of modeling anything which can theoretically be understood by human beings. We lay out how it is so fundamental: why it can be called a "unit of reasoning" and why logical reasoning naturally follows from the traversal of the structure when this organization principle is applied to given data.



(Horizontal diagram available in Appendix)

2. Theory of Information

From the ground up and the top down we introduce a new paradigm, a functional theory of information, what it is by what it does:

Information
$$:= \Delta Potential$$

Or, information is defined to be a change in potential. We then take this definition one step further:

Measurement
$$:= \Delta Information$$

Measurement is defined as a change in information. Diving further into a functional description of information, we describe "meaning" as a "change in state." By doing so, measurement is a change in one state which then changes to another state. Described this way, we can understand that measurement has the same essential characteristics as causation. For instance, if someone is opening a door, then the person is making a specific kind of measurement in this action. We can even think of gravity as a measurement of two objects' masses, and in that measurement, the action of attraction between those two things.

3. Augmented Network/Canonical Form



Like graph models, our network, at its base, is made of nodes and links. Together, these can form a connection:



Three connections can come together as a unit in what we call a canonical:



The primitive symbols are outlined as follows:

- ? query, potential, "some": this is the base primitive (not shown here)
- {} none, nil, "not" (not shown here)
- }{ all, any, "is"
- <> bind, "has"
- >< open, "goes"

The definitions in quotes above are one way to understand how the symbols operate in our system. Essentially, they represent one level of semantics that are induced when elements are placed somewhere in this model. These primitives provide some basic semantics in which words, as nodes, can be placed in the positions in the model above. Even though we do not need all three nodes and all three links to be anything other than the "default" or unfilled values, it is fundamentally important where we place information when it is in structural form (a node/link or canonical). We call this an "augmented" network because a node can be a link, a link can be a node, and a canonical can be a link or a node. In terms of the new theory of information outlined above, we can label a canonical as follows:



The "has" side is information that sets up a position (in a matrix). When we describe information as a change in potential, we can actually think of a potential as a position. The "goes" side, a measurement, takes the operation of that potential (in what way it is understood), and changes it to a new state, thus putting the state into a new potential. The meaning of this measurement, the "is" side, is exactly the change in state that was performed, resulting in the lower right node/canonical which now holds the effect of that meaning.

4. Natural Language to Canonical Form

Currently we employ a natural language parser to convert the natural language input (as of this point, only English, but our model is language agnostic, and other languages will follow) into a tree structure with parts-of-speech tags and a list of universal dependencies. Our main logic uses these outputs to preliminarily put the input into canonical form. An example of the parser output (from the sentence in single quotes):

'A bird is an animal that has wings.'

```
(ROOT
 (S
   (NP (DT A) (NN bird))
   (VP (VBZ is)
    (NP
      (NP (DT an) (NN animal))
      (SBAR
        (WHNP (WDT that))
        (S
          (VP (VBZ has)
           (NP (NNS wings)))))))
   (. .)))
(det bird-2 A-1)
(nsubj animal-5 bird-2)
(cop animal-5 is-3)
(det animal-5 an-4)
```

```
(root ROOT-0 animal-5)
(nsubj has-7 animal-5)
(ref animal-5 that-6)
(acl:relcl animal-5 has-7)
(dobj has-7 wings-8)
```

The canonical form created with the data above is outlined below. The key in this process is that all communication to and from Mind takes place in the form of natural language. Even entering the ontologies comes down to inputting natural language sentences in order to define how the logic should form. Some canonical formation depends on some semantics of the input from the parser, but many can form naturally from the type of dependencies in which they are organized. For example, (nsubj animal bird) is combined with (cop animal is) to form the right side of this canonical (the left side to be added later):



5. Upper Ontology

An upper or uppermost ontology is a set of predicates by which the fundamental ideas of the world and its functionings are distilled into the system [7]. A global ontology is presumed to have unchanging, or rarely changing, pieces of information as a basis for the substance of its reasoning [8]. For example, physical objects have mass, and their masses invoke an increase in gravity by virtue of how near they are to each other [9]. We develop or obtain this upper ontology in order to assimilate new information in terms that the system already understands. There is a balance between the most fundamental of elements in which everything else can be put in terms with, and the need to be expressive enough in implementation that it is not a great inconvenience to navigate through the canonical structures formed from many very elementary canonical forms.

Theoretically, the primitives above can define all things. However, trying to break down unwieldy structures into primitives can be complex. Therefore, in order to avoid such difficulty, we can start with basic units such as: person, place, time, thing, attribute, and function. Though we know a place could also be called a thing, we draw this line in the sand so that we do not add unnecessary abstraction in the interpretation of whatever comes our way.

We can leverage existing upper ontologies to see what might be the most efficient means by which to encode all the basic ideas of what exists, and how processes work. Providing a rationale for why or how, when adding a new ontology to the ontological register, is another step to further build basic ontologies. Generally, there should be an explanation for all of its various pieces. In canonical terms, the rationale of any item of information is represented in the bottom side of the canonical.

6. Compartmentalization

There are four levels where information can be stored: global, local, user, and session.

• **Global**: holds the upper ontology, which are predicates that are globally relevant, unchanging knowledge which is at the root of the understanding of everything else

- Local (or domain information): can exist in a subhierarchy which holds the knowledge applied to a specific domain of interest
- User: holds the profile and history of the user's specifics, and is used to help contextualize information as it pertains to that user
- Session: is transient information that is in effect between the time the user logs in to when that user logs out

7. Contextualization & Ontology Versioning

The context of any data we work with is intrinsic in the fundamental definition of information we have outlined. Contextualization is key to being able to optimally traverse the network of canonicals, as well as the revision of information when new evidence out-dates the current knowledge within the system. In general, we do not delete such old data, but rather we contextualize them as deprecated. This is the essence of ontological versioning. In fact, if we so desire, we can run processes against its old way of thinking, for the purpose of comparing the results between the old and the new. This is one prime advantage we hold in contrast to older symbolic systems which are based on a static pool of rules to draw from. Mind continually learns.

Technically and implicitly, every part of a canonical is context for other parts. The explicit context node is specifically designed to aid in sorting like information, for ease of retrieval as well as use in logical processes.

In operation, we use a method of constraint propagation to reason with specific elements of a problem, the predicates in which logic is known by the system in a generalized form. When we apply constraints in the form of applying canonicals to other canonicals, we are contextualizing by applying general logic to specific cases. This operation is detailed below when we apply predicates to specific configurations in **The 3 Logical Reasoning Embodiments**.

8. Semantic Fluidity/Ontological Topology

In mathematical topology, a coffee mug and a donut are considered equivalent, as certain properties are preserved between the two, such as the fact that they each have one hole. In ontological topology (o-topology), two expressions are equivalent semantically [10]. Even when the canonical forms do not line up between their structures, it does not mean that they must have different meanings.

There are several ways that synonymies are modeled in our system. The most direct method understands that even among synonyms, there are nuances between words that keep the usage of both words current, even as they might be relatively interchangeable. In this way, one can set up a definitional canonical where the effect side says that one word means another, and then the subtlety which divides their meanings is discerned by the qualifying feature.

Another method of o-topology comes from the fact that if two canonicals share the same inductive side, then they are functionally equivalent. In this way, two terms that function identically are then said to be synonymous to each other.

Where there might not be the simpler means of otopology available to us, we again reiterate the importance of the upper ontology: if we can understand connectivity in practical terms, any new information in terms of the information we already understand, and if two expressions are understood in the same way, as far as what connections it has to the knowledge it already comprehends, then we can say that the new information and information already understood are synonymous as well. The novel part of our approach is that the ability to reconcile new information can also be learned. When this ability can handle any arbitrary text, we may say that it has "learned how to learn." This is one component of **Critical Mass**.

O-topology solves a big problem with brittleness that has been associated with symbolic AIs. For example, when talking to a chatbot backed by an AI, a query will fail if you do not express the query using the exact words it understands or in the exact structures that it can parse [11]. With otopology, you do not need the "magic word" to get what you want.

Since we have detailed how the network of connected canonicals is augmented and topological in nature, we can call the structure as a whole, an "augmented topological network."

9. The 3 Logical Reasoning Embodiments

The three distinct ways that a human being reasons are deduction, induction, and abduction [12]. These are all embodied in our data structure, and so we call it a "unit of reasoning." We say it embodies the three forms of logical reasoning by labeling the sides of the canonical as the means by which that type of reasoning occurs. Another problem of symbolic AI is that previous generations of "reasoning engines" focused almost entirely on only the deductive aspect of reasoning [13]. A feature of our model, over other traditional symbolic AI, is that if more constraints (logical structures) are applied, the basis for our contextualization lets us triangulate solutions to problems faster and more accurately.

Because of our three part embodiment, we claim that anything which can be understood, anything that makes some sort of logical sense, can be modeled in this one structure. One way of understanding this structure is by looking at neural networks. Neural networks are made to resemble the operation of neurons and the neural circuitry in the human brain [14], which is to say, from the bottom up. Our structure seeks to model the reasoning of the brain [15], as we know it, from the top down. If you follow the process of reasoning as exemplified by this system, it is clear how the reasoning occurs [16].

A simple example is shown below that uses every node and link but the inductive link shows us how one can define the statement, "A bird is an animal that has wings.":



A bird can then be identified abductively by the question, "What kind of animal has wings?"

This method of abduction is similar to diagnosing a disease, in which a doctor asks the patient to state the symptoms in order to identify which illness best fits the conditions. Just as the symptoms are used to identify the disease, the features of a particular animal aid in its identification. The canonical of the query (the system rephrases the query so as to better match the structure we search against, as ontological topology allows us to do: What is the animal that has wings?):



Another way to look at the canonical:



A bird is effectively an animal, and it has the feature: wings.

Another example, which highlights the "function" connection, can be conceived by a simple case of cause and effect. This more elaborate example shows an operation of the inductive link via what we call an "S-Canonical" (the "S" shape is highlighted):



We see an augmentation in the middle of this figure, "C," which is the canonical shown here:



We defined measurement, and therefore causation, as being capable of changing. In the example shown, the "C" node/canonical changes the initial state (location) of a person "at" place1 so that the person is now "at" place2. The rationale for this development can be defined in C: Condition (The C-canonical can also be C: Cause, and the resultant would be the effect). These types of nested canonicals are one way the canonical forms scale.

For example, let us say that Tom is the person, home is place1, and office is place2:



We instantiate the s-canonical with these singular (one sided) canonicals and thus (piecewise), we see that the satisfaction of a pattern is itself a pattern, thereby constraining the values of the structure. These constraints (such as "Tom is a person") are propagated through the reasoning structure we invoke:



Tom starts "at" home, and then if "C," Tom "goes" to the office, then Tom is now "at" the office. The deductive side is not set to any specific value, and thus just operates as a pass-through. However, the two values on either side of the "}{" must still reconcile. We need to be able to connect one value to the other and keep the integrity of the canonical loci. Note that the functionality of a canonical form, in its most basic process, is merely to satisfy the values in order, and when these values are satisfied, to follow in the path as if going from state to state in a finite state machine or automaton. The main difference between such structures is the need to satisfy the node's value to be "in" or "on" that state.

10. Known Unknowns & Disambiguation

When we have formed the basic ontology in terms of canonical forms, one interesting property emerges: when we have a partial match of a certain canonical, we can then project what information is necessary to fulfill the rest of that pattern. Intelligence, artificial or not, in one sense can be understood as the ability to complete an incomplete pattern.

One facility the system has, is the ability to match specific partial patterns. Matching partial patterns contribute to understanding problem solving methods. The process by which an incomplete pattern is completed is itself a pattern, and fulfills the idea as first introduced above: the satisfaction of a pattern is itself a pattern. All actions which transpire in the scope of the system are understood in some way. This would be the result of what was meant when our system has "learned how to learn." There must be a way for anything that makes sense, to make sense to us (More on that idea is below, in **Critical Mass**). If these things make sense to us, then they can make sense to those who observe or peek in, for we are logically connected everywhere to everything, wherever we can reach and connect.

When there is more than one valid value that turns up in a search for a piece of information (an ambiguity), we can treat that superpotential as an incompleteness. In order to disambiguate in cases where we have knowledge of what data is needed for an unknown to be known, we apply existing context(s) or post the problem to the community to resolve its difficulty. This process can be applied irregardless of whether the solution is completely unknown or if there are multiple possible solutions [17]. The kind of information that will solve the unknown is borne out by the slots where the specific data would satisfy the completion of the pattern in which there is a missing piece.

For example, the question, "Where is Tom?" can be refit topologically as "Tom at ?":



From the canonicals defined above in **The 3 Logical Rea**soning Embodiments, we saw that the ? on the lower left can be satisfied by both the places "home" and "office." We also saw that the way the two possibilities are connected is through the C-canonical, whether Tom went to the office or not, and when we obtain the resolution (let's say, from the natural question given to us by Mind, "Did Tom go to the office?" and we answer, "no"), by the directionality of the canonicals in question (if we say "no," then the C-canonical cannot have happened yet), the system can conclude the answer, "Tom is at home."

11. Natural Language Generation

As more and more inputs are arranged into canonical form, we will keep track of the transformations that occur from the unstructured to the structured forms. In so doing, we can reverse the process, so as to create grammatically correct sentential representations with which to answer the user or to request more information. This is yet another example of its ability to learn by observation of the dialog that transpires between the system and its users. Whatever goes in can then come out. Idiomatic structures, once understood, can be formed as a response back to the user, where it might be relevant to do so.

12. Transparency of Operation

Since the logical process through a canonical follows the three operations of that canonical (as we have shown), we only need to keep track of what nodes, links, and canonicals are traversed upon the input of some problem or query. It is easy to see that a metatool, which can look up the processes followed, is both desirable and readily feasible. Anyone with permission would have the ability to first see what logic is present, how that logic is used, and be able to make modifications if there were any fault in the logic. With Mind's ability to summarize, the tool could let an authorized agent zoom in or out when dealing with a complex process, to root out possible errors.

13. Critical Mass

In atomic theory, when the correct radioactive matter reaches a certain mass, an autocatalytic process is initiated [18]. Similarly, we describe ontological critical mass (and its trappings, namely, the reasoning facilities) as the containment/comprehension of enough information and functionality for Mind to learn on its own. If something is beyond the capability of the current state of the system to understand, it must, on its own, figure out just what that thing is in relation to what *is* known. For this to happen, as it relies on ontology, the knowledge basis must be in at least one sense "complete." There must be a way to be able to open up all possibilities. This is not an intractable problem.

What things can exist and what can those things do? These two questions cover most of the bases of what can be understood. The only qualifier to understanding is what we can break down the understanding of something into. Nothing is magic. Either we have all the principles by which new information can be fathomed, or we need new principles (The latter case is described in more detail below, in **Metatheoretics**).

On the practical side, if we can create an interface to the outside world (in cyberspace) for it to search the web and scrape the contents of the websites it finds, then in theory it should be able to do its own research on topics it knows very little about to solve some problem or query.

In preparation for the "critical mass" phenomenon, we need to include within Mind's functionality a means by which it can discern good information from bad. This can be seen as another aspect of it having "learned how to learn." The Mind needs to be able to understand and automatically weed out false information, and in situations in which a piece of information is beyond Minds ability to reasonably check its verity, it needs to turn to the community for assistance. Through assistance, Mind learns how to perform certain verity checks on its own, or which techniques it can apply at a future time when faced with similar challenges. The core concepts within its upper ontology (if not also other, perhaps mid-level ontologies) need the ability to critique information being input, and there would surely be sandboxing when the information source is perhaps questionable in its reliability.

14. Metatheoretics

The ultimate goal of this approach is to be able to create an artificial scientist that is able to create its own hypotheses and theories, and to perform experiments that reach new conclusions about the world. These functionalities can be founded upon simple statements, comprehensions of basic principles, and/or scientific theories. At the base of its functionality, we will set up rules for generalizing, which allow specific events to trigger the formation of broader predicates. We will focus on observing why events happen as they do. For example, when a penicl falls to the floor, the phenomenon can be explain by gravity, as all objects with mass are affected by gravity and all physical objects have mass. Therefore, at the most basic level of theorization, we can hypothesize that anything we drop that falls to the ground is a physical object with mass.

As Mind learns more about the world, it is able to understand the relationships between new information and its prior knowledge, and understand what can be measured, how things are connected, and how things make sense. As a result, Mind can then formulate its own hypotheses.

The art of crafting theories depends primarily upon finding a relationship between disparate phenomena. If Mind does not understand why something happens, it will try to obtain the information elsewhere to see if it can make sense of it. In addition to this capacity, it knows what to do if something is not resolved or if a resolution cannot be achieved with the scholarship that already exists. In the first iterations of the development of Metatheoretics, we will feed it situations so it can attempt to discover existing theories. We input past insights that humans used to create those theories into Mind, but we limit its reach of knowledge. Therefore, though it can see the observations that led to those theories, the conclusions of said evidence are not given. This "training" would be equivalent to an AI's university education. After it graduates, and understands how the great theories of the past came to be, it will be able to do some real world work. By allowing Mind to draw its own connections with limited knowledge, it will also have theories about how such theories are made. This is the essence of Metatheoretics.

When Mind can come up with its own scholarship and reaches that height of artificial scientific method, we can truly say that we have created an intelligence. It will be able to do research on a level unprecedented by human reasoning. Experts in domains such as drug discovery or materials engineering have the entirety of all the existing scholarship in their fields including all the papers from the beginnings of their field's establishment. Yet, Mind will have all of that same information "in its head," properly annotated, and checked for logical missteps based on the understanding it has from the entirety of its knowledge. Mind will then be in a position to discover something truly new by formulating novel hypotheses and devising experiments to bear them out. Ultimately, because we all can see how insights are reached by Mind, we are in a position to advance our own intelligence. This is the essence of Mind.

15. Conclusion

We have defined the fundamental *unit of reasoning*, described why it is so named, and described how it may be operated with and upon to be able to reason via the three types of logical reasoning. We have shown how it is, as a whole, an "augmented topological network," its capacity for augmentation and the topological nature of the system. By this structure we have shown how crisp, qualitative, linear, and logical reasoning may be performed, and how the problem of symbolic brittleness is solved with ontological topology. Ultimately, Metatheoretics promises that which has never before been within reach— the development of an adaptive and autonomous entity that will explore the vastness of the world and then, on its own, determine how it works.

It's been stated that when AI solutions are created to solve problems, we are no longer dealing with an AI, but merely algorithms [19]. And these algorithms generate only an automated sense of reasoning, which leaves us with the fears that stem from unpredictability in our deep learning systems today. But, what if we now pose the solution to close that gap? What lies before us goes beyond algorithms, and hinges upon understandability, reasoning, and accountability. With the ability to reason, for itself, we then will be able to isolate and receive explanations for its thoughts, answers, solutions, and actions.

References

- Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. Neural Networks. 61: 85117. arXiv:1404.7828 . doi:10.1016/j.neunet.2014.09.003
- [2] Schuster, M., Johnson, M., Thorat, N. (2016, November 22). Zero-Shot Translation with Googles Multilingual Neural Machine Translation System. Retrieved from https://research.googleblog.com/2016/11/zero-shot-translationwith-googles.html

- [3] AlphaGo. (n.d.). Retrieved from https://deepmind.com/research/alphago/
- [4] Kurzweil AI Network. (2017). AI system performs better than 75 percent of American adults on standard visual intelligence test. Retrieved from http://www.kurzweilai.net/ai-system-performs-better-than-75percent-of-american-adults-on-standard-visual-intelligence-test
- [5] Essays, UK. (November 2013). Advantages and Limitations of Neural Networks. Retrieved from https://www.ukessays.com?vref=1
- [6] Masters, T. (1999). Practical neural network recipes in C. San Diego: Acad. Press.
- [7] Hoehndorf, R. (2010). What is an upper level ontology?. Retrieved from http://ontogenesis.knowledgeblog.org/740
- [8] Informatica, Vakgroep Toegepaste Informatica, En Meersman, Robert A. Meersman, R. (2003). Semantic Ontology Tools in IS Design. 1609. 10.1007/BFb0095088.
- [9] Leighton, J. P. (2004). Defining and describing reason, in The Nature of Reasoning (eds Leighton, J. P. and Sternberg, R. J.) Cambridge University Press
- [10] Carter, J. S., Saito, M. (1998). Knotted surfaces and their diagrams. Oxford: American mathematical Society.
- [11] Block, Ned (1995). The mind as the software of the brain. In Daniel N. Osherson, Lila Gleitman, Stephen M. Kosslyn, S. Smith Saadya Sternberg (eds.), An Invitation to Cognitive Science, Second Edition, Volume 3. Cambridge MA: MIT Press. pp. 377-425.
- [12] Thagard, P., Shelley, C. (n.d.). Abductive reasoning: Logic, visual thinking, and coherence*. Retrieved from http://cogsci.uwaterloo.ca/Articles/Pages/Abductive.html
- [13] Dreyfus, H. L. (1994). What computers still cant do: A critique of artificial reason. Cambridge, MA: MIT Press.
- [14] Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model For Information Storage And Organization In The Brain". Psychological Review. 65 (6): 386408. CiteSeerX 10.1.1.588.3775 . doi:10.1037/h0042519. PMID 13602029
- [15] Johnson-Laird, P. N. (1995). Mental models: Towards a cognitive science of language, inference, and consciousness. Cambridge: Harvard Univ. Press.
- [16] Hintikka, Other Disci-J. J. (n.d.). Logic And In Encyclopaedia Retrieved plines. Britannica. from https://www.britannica.com/topic/philosophy-of-logic/Logic-andother-disciplines
- [17] Lesk; M. 1986. Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In Proc. of SIGDOC-86: 5th International Conference on Systems Documentation, Toronto, Canada.
- [18] The Editors of Encyclopaedia Britannica. (2018). Critical mass. In Encyclopaedia Britannica. Retrieved from https://www.britannica.com/science/critical-mass
- [19] Kahn, J. (2002). It's Alive!. Retrieved from https://www.wired.com/2002/03/everywhere/

Paul Lee, M.D. Cofounder & CEO - Mind AI Phone: +1 (949) 533 9346 E-mail: paul@mind.ai

For more information and the full whitepaper, please visit www.mind.ai.

16. Appendix



Fig. 2. System architecture