A Non-Classical Syllogism with Self-control

Xi Li

Center for Logic, Language and Cognition
Department of Philosophy, Peking University
Beijing, 100871, China
xli@pku.edu.cn

November 5, 2012

Contents

1 Background and Motivation **2**
1.1 Remarks on the semantics of NAL . . . 3
1.2 Remarks on the syntax of NAL . . . 5
1.3 Other related works . . . 5
1.4 Motivation . . . 6

2 Syntax of ANS **7**
2.1 Language . . . 7
2.2 Truth Function . . . 10
2.3 Inference Rules . . . 14
2.4 Deduction . . . 19
2.5 Axioms and Reference Rules for Programs . . . 22
2.6 Axioms and Reference Rules for Updates . . . 23

3 Semantics of ANS **25**

4 ANS Reduction **28**
4.1 Axioms and Reference Rules for Updates(Continued) . . . 28
4.2 Applications . . . 30

5 Conclusion and Further Research **32**

References 32

Abstract

Pei Wang [P.Wang12] introduces a Non-Axiomatic Logic (NAL) whose semantics is experience-grounded, where meaning is implicit definition. However, if we track the whole history of the Knowledge Base and make sure that every relationships of the terms are noted down chronologically then the experience-grounded semantics can be seen as pure syntax, because formal semantics of a language should be formally defined in its meta-language and the experience-grounded semantics can't be formally figured out explicitly. The truth function $\langle f, c\rangle$ of NAL should not be confused with the binary truth value $(0, 1)$ of Classical Logic. Seen from its whole lifespan, maybe it also can be considered as a formal axiomatic system or a sequence of formal systems. In order to strengthen the reasoning power of NAL, Pei Wang allows meta-level theorems as analytic truths to be introduced to NAL as a judgement with a truth value $\langle 1, 1\rangle$, as long as they can be translated into NAL language. And they are only implicitly used in the inference process of NAL. Since no effective procedure is available to tell which meta-theorem can be introduced and which cannot, this paper tries to give a partial axiomatized variation of NAL with a lot of modifications, refinements and supplements, for example, the truth functions and most of the reference rules, are quite different from Pei Wang's NAL so as to adapt to the purpose of taking good advantage of Probability and Classical Logic and giving new names to compactly interconnected structures. Combined the ideas of NAL [P.Wang12] and KeY [Beckert,Hähnle,Schmitt07], several update operators are introduced to manage self-monitoring and self control in an explicit manner.

Keywords

NAL; DL; KeY; implicit definition; update operator

"Mathematics is the art of giving the same name to different things. . . . Moreover the physicists do just the same."

— Henri Poincaré

§1. Background and Motivation

"Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true."

— Bertrand Russell

§1.1. Remarks on the semantics of NAL

Pei Wang [P.Wang12] introduces a Non-Axiomatic Logic (NAL) whose semantics is experience-grounded, contrary to Tarski semantics, so the meaning of a term is determined by its relationships to other terms entangled with each other in the knowledge base (KB) which is incomplete under the Assumption of Insufficient Knowledge and Resources (AIKR), rather than by an interpretation that maps it to an object in an external model, and the truth value of a statement indicates the degree to which the statement is confirmed by past experience, rather than by the correspondance to a fact of the external model. In this sense, *meaning is implicit definition.* However, if we track the whole history of the KB and make sure that every relationships of the terms are noted down chronologically then the experience-grounded semantics can be seen as pure syntax, because formal semantics of a language should be formally defined in its meta-language and the experience-grounded semantics can't be formally figured out explicitly in a formal meta-language. The truth function $\langle f, c\rangle$ of NAL should not be confused with the binary truth value $(0, 1)$ of Classical Logic.

Actually, in my opinion, Tarski semantics, usually defined in a meta-theory (for example,ZFC) relative to some object theory (for example,FOL), is not as *"semantic"* as we take it for granted. If the definition of the concept *true* obeys Tarski's T-convention—"$P(c)$" is true iff $P(c)$, then the truth value of the sentence $P(c)$ must depend on the reference of the term c and the situation that $P(c)$ involved in. So it is nonsense to talk about reference without a background model and the pursuit of truth in that model, no matter whether the model is constructed by ourselves or presupposed settled in the Plato world beyond our reach. The reason that we usually can't fully capture the univocal reference of a term is that not every detail about it belongs to our KB which is incomplete and continuously (not necessarily monotonely) extendable according to our progress of scientific discovery in the physical world. To tell if a label is suitable for the moment for a certain object we observed is a problem of pattern classification, which scientists should be responsible for either at the pretreatment stage or regard it as a consequence of reasoning rather than deal with it during the process of inference. For example, when we want to know whether *whales are fish*, we have already know what are whales and what is fish, all we need to do is to compare the properties of whales and the properties of fish according to our knowledge of biology in our KB. When we want to know whether the big animal X we observed is a whale or not, we'd better try our best to discover as more properties of X as possible and then search the properties of whales in our KB, if most of the main properties of them match well, we can tell that X is a whale. Therefore, semantics in the sense of talking about the relationship between language and the physical world can't be explicitly defined and truth values can't be directly assigned. Just like in mathematics, we hope the term "truth" covers every truth in the Plato world, however, there is no effective procedure to figure out every truth for us once and for all. The concept "truth" is also infinitely extendable. Besides "truth", in fact, all of our concepts in our language, including mathematical concepts, can be seen as incomplete symbols waiting to be completed.

Recall the acceptable framework of mathematics nowadays. What is the definition of a model of FOL? It is an ordered tuple $\langle D, \eta \rangle$ defined in set theory (ZFC), where D refers to a nonempty set, and η maps the terms to the elements of the nonempty set, while predicate symbols to the elements of the Cartesian product of the nonempty set. Even we would like to talk about some "concrete" domain composed of "concrete" objects such as "bottle","car","Caesar", actually all of them are nothing real but some other *labels*, or *terms* in other words. What is called a model is just another algebra system defined somewhere else, most of the time in ZFC. But should we believe in ZFC? What is ZFC? It is a formal system consists of some axioms about "set" in the language $\mathscr{L}_\in$ based on FOL. Then what is FOL? It is also defined by several axioms and some reference rules which implicitly define the logical symbols "$\neg$", "$\wedge$", "$\vee$", "$\rightarrow$", "$\leftrightarrow$", "$\forall$", "$\exists$". Therefore, ZFC (based on FOL) should not only be regarded as a formal system built to define the concept "set", but to develop an implicit definition of "$\in$" together with "$\neg$", "$\wedge$", "$\vee$", "$\rightarrow$", "$\leftrightarrow$", "$\forall$", "$\exists$" simultaneously. According to Löwenheim-Skolem theorem, if a first order formal theory, for example ZFC, has infinite models, then it has at least a model whose domain is only countable, while Cantor theorem guarantees the existence of uncountable set, which means that ZFC is incapable of fully capturing its intended interpretation. The "intuitive notion of set" is beyond ZFC's reach. We even can't figure out a standard model for ZFC. If the concept "set" itself is ambiguous, how can we count upon Tarski semantics defined in ZFC to fix the univocal references for our terms? Just as Russell said "Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true". Maybe all of our concept, including mathematical concept, are implicitly defined. Consistency is nothing but relative consistency. But, when we talk we have to stand somewhere to talk, or we would fall into infinite regress. Maybe ZFC is the proper place safe and abundant enough to sustain us according to our experience. So when we talk about models we usually presuppose that we work together in the same framework of ZFC. We have no choice but to build a Plato world for ourselves by our own hands based on our practice in the imperfect physical world. In other words, only in the KB does one's belief have a rational justification. The agent owes his entire belief to the KB, and has his being within it alone. Whatever worth and spiritual reality he possesses are his solely by virtue of the KB.

From this point of view, the ultimate semantics must be grounded on either the physical world or Plato world, a formal semantics must be based on some meta-language (meta-theory) we've built ourselves. Pei Wang's experience-grounded semantics can be seen as pure syntax if we look into the whole lifespan of it, and in some sense, some Tarski-style semantics can also be built for NAL from some meta-level theory as long as its formal meta-language can be given explicitly.

§1.2. Remarks on the syntax of NAL

Pei Wang's NAL is a variant of Term Logic, a radical extension of Aristotle's Syllogism, including inference rules such as deduction, induction, abduction, exemplification, revision, conversion, analogy, comparison, resemblance, intersection, union, difference, negation and so on. The reason that Pei Wang call it a non-axiomatic logic is that it is a finite real-time open system, statements can be added or forgotten at any time. However, from the view of its whole life cycle, it also can be considered as a (sequence of) formal axiomatic system(s). In order to strengthen the reasoning power of NAL, Pei Wang allows meta-level theorems as analytic truths to be introduced to NAL as a judgement with a truth value $\langle 1, 1\rangle$, as long as they can be translated into NAL language. But they are only implicitly used in the inference process of NAL, and NAL only contains empirical truths with confidence lower than 1, refusing to store any absolute truth as *axiom* on the object-level. In my opinion, there is no difference with stealing to play such a trick, because no effective procedure is available to tell which meta-theorem can be introduced and which cannot. This kind of trick can only be played by its creator, just like an Oracle to a Turing machine. Therefore, the cleverness in this sense has noting to do with NAL's intelligence. To be capable to reason with "analytic truth", NAL should acquire its ownership. To be qualified to be an intelligent agent, the system should be able to obtain its own "absolute truths" through uncertain inference by itself, though, philosophically speaking, it is a mission impossible. Even we human beings can never surely know that we are proceeding in the right direction to approximate the "absolute" truth. However, we can make assumptions according to our experience to pretend that our hypothesis would have truth value 1, under which we make implicit definitions and create our own abstract theory which is very useful in some particular field. Although we live in a coarse physical world, the presupposition of a Plato world (and possible worlds) is, pragmatically speaking, a useful convenience.

In other words, for the theorems proved in the system itself at time t, it can be added into KB with truth value $\langle 1, 1\rangle$ at time $t+1$ if we would like, namely, we treat it as a dynamic system. However, for the theorems acting as definitions of some operators of NAL, they must be included as axioms.

§1.3. Other related works

DL Dynamic logic (DL[1]) is an extension of modal logic and classical logic. Reasoning about properties of composite programs is allowed in DL, and programs are explicitly part of the language.

Queries like $? \to P\ \langle f, c\rangle$, $S \to ?\ \langle f, c\rangle$, $S \to P\ \langle ?\rangle$ in NAL can be understood as programs $\langle X := ?\rangle X \to P\ \langle f, c\rangle$, $\langle X := ?\rangle S \to X\ \langle f, c\rangle$, $\langle X := ?\rangle S \to P\ \langle X\rangle$ in some sense. And programs $X := S, ?\phi, \pi_1; \pi_2, \pi_1 \cup \pi_2, \pi^*$ can also be introduced to NAL.

[1]The standard materials about DL can be found in [D.Harel,D.Kozen,J.Tiuryn00]

KeY The *KeY* project [Beckert,Hähnle,Schmitt07] introduces some interesting update operators to JAVA CARD DL. The goal of *KeY* is to create a formal tool that integrates design, implementation, formal specification, and formal verification of Java Card programs within a commercial platform for object-oriented specifications languages OCL/JML-based software development. Five types of update operators can be found in the *KeY* book [Beckert,Hähnle,Schmitt07]:

function update	$f(t_1, \cdots, t_n) := t$
sequential update	$u_1; u_2$
parallel update	$u_1 \| u_2$
quantified update	$for\ x; \phi; u$
update application	$\{u_1\}u_2$

Some ideas can be borrowed from them to make abstract inference feasible in NAL.

§1.4. Motivation

The standard semantics of Propositional Dynamic Logic(PDL) is the Kripke structure, i.e. the possible world semantics. But what is a possible world? Sometimes it is regarded as a maximal consistent set of propositions, sometimes the recombination of the possible ingredients of the actual world, sometimes a parallel universe which coexists peacefully with ours, and sometimes a possible state. In the formal definition of Kripke semantics of PDL a possible world is a point of a set with no inner structure. However, the possible world of the semantics of First-order Dynamic Logic(FDL) do have a structure, which is the variable assignment. Singular statements are scientific observations made under a certain condition, while universal statements are regularity, law, or principle holding under various observations. The latter type is more helpful, for science is, in some sense, the art of data compression. Since the semantics of NAL is a complicate graph, the domain of quantifiers can be seen as all the points connected with a certain point through edges. But any point or any edge may be added or deleted at any time. Therefore, although statements containing variables stem from the facts of the KB which is based on the actual world, yet we'd better regard them belonging to some abstract world. Then what is a state? A state (may contain contradictory knowledge in some remote location) is probably a snapshot of the KB at a certain time involving no variables. Heraclitus said that you cannot step twice into the same river, for fresh waters are ever flowing in upon you. The point is, whenever a event happens, a state changes, and the states are changed by any event (by update operators in ANS) while we still live in the same possible world only if the variables do not change. In other words, state changes easily by our hands while possible world remains rigid or, maybe, altered randomly or revolutionarily. However, when the variables occurring in the statements are changed, some possible worlds open to us. Abstractly speaking, what is a event of a graph? A point added or deleted, an edge added or deleted, two points merged into one, etc. What is the actual physical world we understand? It is the history of all the states evolving in the KB. What is the possible worlds? They are the variants of the actual world by abstraction, and any change of the variables (by program operators in ANS) results in different possible world. What's the relationship between states and possible worlds? Even we live in the same physical world, we can conceive various possible worlds. An abstraction update operator $S := X$ just like λ in Lambda Calculus should be added to our language. Mathematics is the extension of natural science into the field of possible worlds.

Now we can talk about most of the events, but what is the relationship between them? Edges between statements and the inference rules in NAL reveal it explicitly; and most of the update operators of KeY reveal it to some extent by themselves. For example, two events can occur sequentially, concurrently with no interactions between each other, or they clash, one dominates the other. Besides, sometimes a event happen only if some precondition is satisfied, or another event will happen. And, sometimes we would like to withdraw what we have said or what we have done. So, it's not unnecessary that, a conditional update be introduced $\phi \to u_1/u_2$, as well as a forget update $S := \bot$ and an action update $T :\to S$.

Then how the update affect the program and the truth value of a formula? When some event happen, some state being changed, should the program be rewritten or just updated with an operator to adapt itself to the new environment? Some reduction equations will be given.

Contributions of this paper According to the remarks on NAL, a Non-classical Syllogism (ANS) is given, and some details, for example, the truth functions and most of the reference rules, are quite different from Pei Wang's NAL so as to adapt to the purpose of taking good advantage of Probability and Classical Logic and giving new names to compactly interconnected structures. Combined the ideas of NAL [P.Wang12] and KeY [Beckert,Hähnle,Schmitt07], several update operators are introduced to manage self-monitoring and self control in an explicit manner.

Structure of the paper In section2 we will give the syntax of ANS, which includes the whole language, truth functions and inference rules. In section3 we give the semantics of ANS. In section4 we will give a sequence of reduction equations for eliminating the update operators and show some examples for application. After that, the last section5 make conclusion and discuss the future work.

§2. Syntax of ANS

§2.1. Language

Relative rigid symbols should be distinguished from flexible symbols, the former need to have the same interpretation (as analytical truth) in all execution states, while the latter need to capture state change after update execution, so their interpretation depends on state. Logical variables, built-in operators and some special terms belong to the former, with constant terms the latter.

If Kronecker is believable that God created the natural numbers and all else is the work of man, then $0, 1, +, \times, \leq$ must be rigid symbols. This paper assumes the real number and the operations of real number such as $\leq, +, \times$ are rigid, and the truth value $\langle f, c\rangle$ of a statement can only be changed by programs. Update operators can not change it but can replace it with an variable to be assigned by some program.

The language $\mathscr{L}$ consists of countably infinite constant terms (S) and variables (X) to assure that there are enough fresh terms available to use relative to our KB. Given the formal language $\mathscr{L}$, terms,formulas,programs and updates are mutually defined, for which reason some restrictions should be taken care of to avoid vicious circle of some sort. For simplicity, we will not figure out every detail of the restrictions here.

Terms $Term_{\mathscr{L}}$

$$T ::= \top|\bot|S|X| \in |c| + | \times | \geq |T - P|T \cap P|T \cup P|$$
$$T \times P|\phi \to T/P|\mu X\phi|\nu X\phi|\{T\}|\{\phi\}|\{u\}T$$

Remark: the truth value $\langle f, c\rangle$ and its operators/relations $+, \times, \geq$ should be regarded as special primitive terms fixed in the system! Another special term is $\in$.

Terms $\top$ and $\bot$ refer to the universal concept and empty concept respectively, so intuitively, for any term T, $T \to \top\ \langle 1, 1\rangle$ and $\bot \to T\ \langle 1, 1\rangle$ hold. Compound terms can be created from the observational terms. For example, $T - P|T \cap P|T \cup P|T \times P|\phi \to T/P|\mu X\phi|\nu X\phi|\{T\}|\{\phi\}|\{u\}T$ are the compositions of existing terms. The connectors $-, \cap, \cup, \times, \in\ \{\}$ are like set operations *difference,intersection,union,Cartesian product,in,"the set of"*, while μ, ν are similar to *min* and *max* respectively, but based on different criterion. $\phi \to T/P$ is a conditional term, and ϕ specify the *context*. $\{u\}T$ stands for the modified term which is updated by the action u.

Formulas $Form_{\mathscr{L}}$

$$\phi ::= T \to P\ \langle f, c\rangle|T \leftrightarrow P\ \langle f, c\rangle|\neg\phi|\phi \wedge \psi|\phi \vee \psi|$$
$$\phi \Rightarrow \psi|\phi \Leftrightarrow \psi|\forall X\phi|\exists X\phi|[\pi]\phi|\langle\pi\rangle\phi|\{u\}\phi$$

There is a pair of truth values $\langle f, c\rangle$ following every formula ϕ in ANS just as in NAL.

Programs $Prog_{\mathscr{L}}$

$$\pi ::= X := T|X :=?|?\phi|\pi_1; \pi_2|\pi_1 \cup \pi_2|\pi_1 \cap \pi_2|\pi^*|\{u\}\pi$$

Through programs abstract reasoning and operations are available based on some assumptions.

Updates $Updt_{\mathscr{L}}$

$$u ::= S := X | T := S | T := \bot | T :\to S | u_1; u_2 | u_1 \| u_2 | \phi_{\geq \langle \bar{f}, \bar{c} \rangle} \to u_1 / u_2 | (\forall X \phi) u | \{u_1\} u_2$$

Update operator $\{S := X\}$ is the passage leading from the perceptible world to the abstract world. It is the opposite of program operator $X := T$. Assumptions or hypothesises involving variables can only be formulated through abstraction operator $S := X$. $T := S$ means to give a name S to term T, and $T := \bot$ is the delete operation, when we want to delete a term(point) S, we turn to $S := \bot$, when we want to delete a formula(edge) ϕ, we'd better keep $\{\phi\} := \bot$ at hand. If you want Tom open the door, you need $\{Tom\} \times \{door\} :\to$ *open*, and after that, an edge is added between the term $\{Tom\} \times \{door\}$ and *open*. $\phi_{\geq \langle \bar{f}, \bar{c} \rangle} \to u_1 / u_2$ means if $\phi \langle f, c \rangle$ and $\langle f, c \rangle \geq \langle \bar{f}, \bar{c} \rangle$ then $\{u_1\}$ should be executed, else $\{u_2\}$ be executed. Sometimes we write $\phi \to u_1 / u_2$ for short. All of the actions can be carried out sequentially,concurrently,conditionally or repeatedly.

Most of the definitions of terms,formulas and updates should be given mutually. For example, we define free variables of a term, formula or an update as follows:

Definition 1. *Free Variables*
$Fv(X) = \{X\}$;
$Fv(\top) = Fv(\bot) = Fv(S) = Fv(\{T\}) = Fv(\{\phi\}) = \varnothing$;
$Fv(\phi \to T/P) = Fv(\phi) \cup Fv(T) \cup Fv(P)$;
$Fv(\mu X \phi) = Fv(\phi) \setminus \{X\}$; *so is* ν.
$Fv(\{u\}T) = Fv(u) \cup Fv(T)$; *so is* $-, \cap, \cup, \times$.
$Fv(T \to P) = Fv(T \leftrightarrow P) = Fv(T) \cup Fv(P)$;
$Fv(\neg \phi) = Fv(\phi)$;
$Fv(\phi \wedge \psi) = Fv(\phi \vee \psi) = Fv(\phi \Rightarrow \psi) = Fv(\phi \Leftrightarrow \psi) = Fv(\phi) \cup Fv(\psi)$;
$Fv((\forall X \phi)\psi) = Fv((\exists X \phi)\psi) = Fv(\phi) \setminus \{X\}$;
$Fv([\pi]\phi) = Fv(\langle \pi \rangle \phi) = Fv(\phi)$;
$Fv(\{u\}\phi) = Fv(u) \cup Fv(\phi)$;
$Fv(S := X) = \{X\}$;
$Fv(T := S) = Fv(T := \bot) = Fv(T :\to S) = Fv(T)$;
$Fv(u_1; u_2) = Fv(u_1 \| u_2) = F(u_1 ! u_2) = Fv(\{u_1\}u_2) = Fv(u_1) \cup Fv(u_2)$;
$Fv(\phi \to u_1/u_2) = Fv(\phi) \cup Fv(u_1) \cup F(u_2)$;
$Fv((\forall X \phi)u) = (Fv(\phi) \cup Fv(u)) \setminus \{X\}$.

A term without free variables is called closed term, and a formula without free variables a sentence.

Axioms The elementary axioms of $+, \times$ of real numbers, for example, the axioms of the real closed field, are supposed to be included in ANS. And most of the theorems as definitions in NAL should be added explicitly as axioms with truth value $\langle 1, 1 \rangle$, but we usually omit $\langle 1, 1 \rangle$ for short. For example, even in NAL, for the operator

{}, the following axioms should be added explicitly as an axiom rather than as a meta-level definition.

$$(T \times P \rightarrow \epsilon) \Leftrightarrow (\{T\} \rightarrow P)$$
$$\forall X((X \rightarrow P) \Rightarrow (T \rightarrow X)) \Leftrightarrow \exists X((X \times P \rightarrow \epsilon) \wedge (\{X\} \leftrightarrow T))$$

However, the above axiom is only about the ideal case $\langle 1,1\rangle$, what about the uncertain cases? The uncertain cases as well as the axioms and reference rules about the program and update operators will be given later.

§2.2. Truth Function

The frequency f and confidence c are based on the evidence w^+, w. The table1 is the same as Pei Wang's[P.Wang12].

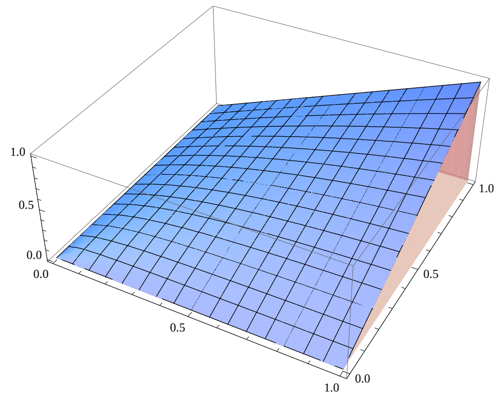
	$\{w^+, w\}$	$\langle f, c\rangle$	$[l, u]$
$\{w^+, w\}$ $(w = w^+ + w^-)$		$w^+ = kfc/(1-c)$ $w = kc/(1-c)$	$w^+ = kl/i$ $w = k(1-i)/i$
$\langle f, c\rangle$	$f = w^+/w$ $c = w/(w+k)$		$f = l/(l-i)$ $c = 1-i$
$[l, u]$ $(i = u - l)$	$l = w^+/(w+k)$ $u = (w^+ + k)/(w+k)$	$l = fc$ $u = 1 - c(1-f)$	

Table 1: Uncertainty Measurements

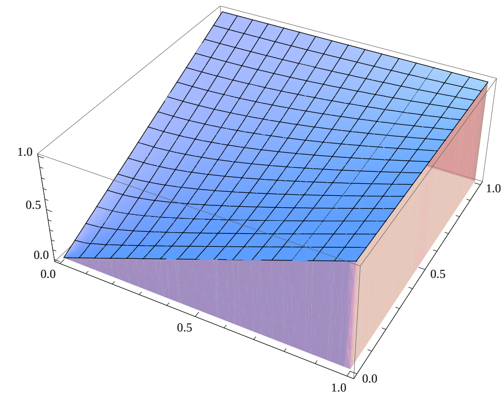
Definition 2 (Truth Function)**.**

$$x \owedge y = \begin{cases} (x^{-1} + y^{-1} - 1)^{-1} & \textit{if } xy \neq 0 \\ 0 & \textit{otherwise} \end{cases}$$
$$x \ovee y = x + y - x \owedge y$$
$$x \multimap y = \begin{cases} (x \owedge y) \times x^{-1} & \textit{if } xy \neq 0 \\ 1 - x & \textit{otherwise} \end{cases}$$
$$x \ominus y = (1 - (x \multimap y)) \times x$$
$$x \boxminus y = x \owedge (1 - y)$$
$$\circledcirc(x) = x^k$$
$$not(x) = 1 - x$$
$$and(x_1, \cdots, x_n) = x_1 \times \cdots \times x_n$$
$$or(x_1, \cdots, x_n) = 1 - [(1 - x_1) \times \cdots \times (1 - x_n)]$$

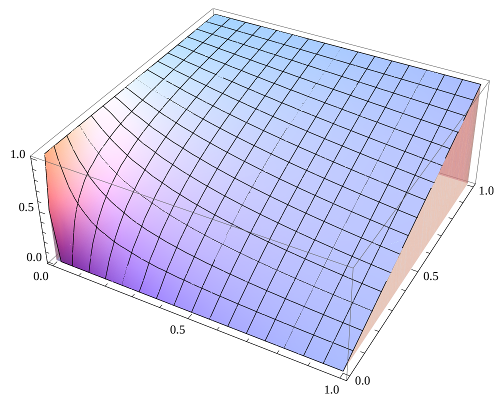
where k is a system parameter.



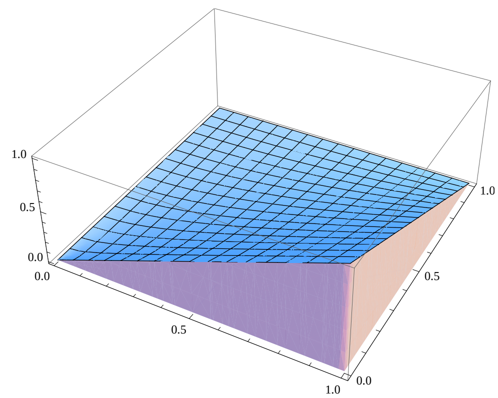
(a) $\owedge$



(b) $\ovee$



(c) $\rightarrowtriangle$



(d) $\ominus$

Methematica code:

$$F[x\_, y\_] := If[xy! = 0, 1/(1/x + 1/y - 1), 0] \tag{$\owedge$}$$
$$Imp[x\_, y\_] := If[xy! = 0, F[x, y]/x, 1 - x] \tag{$\rightarrowtriangle$}$$
$$V[x\_, y\_] := x + y - F[x, y] \tag{$\ovee$}$$
$$M[x\_, y\_] := x - F[x, y] \tag{$\ominus$}$$
$$Plot3D[F[x, y], \{x, 0, 1\}, \{y, 0, 1\}, Filling- > Bottom] \tag{$\owedge$}$$
$$Plot3D[Imp[x, y], \{x, 0, 1\}, \{y, 0, 1\}, Filling- > Bottom] \tag{$\rightarrowtriangle$}$$
$$Plot3D[V[x, y], \{x, 0, 1\}, \{y, 0, 1\}, Filling- > Bottom] \tag{$\ovee$}$$
$$Plot3D[M[x, y], \{x, 0, 1\}, \{y, 0, 1\}, Filling- > Bottom] \tag{$\ominus$}$$

Theorem 1 (Truth Function). *For $x, y \in (0, 1)$*

$$1 \ominus x = 1 \boxminus x = 1 - x \tag{2.1}$$

$$x \ominus y \neq x \boxminus y \tag{2.2}$$
$$x \owedge (1 \ominus x) \neq 0 \tag{2.3}$$
$$x \ovee (1 \ominus x) \neq 1 \tag{2.4}$$
$$1 \ominus (1 \ominus x) = x \tag{2.5}$$
$$1 \ominus (x \owedge (1 - x)) = (1 \ominus x) \ovee (1 \ominus (1 - x)) \tag{2.6}$$
$$1 \ominus (x \ovee (1 - x)) = (1 \ominus x) \owedge (1 \ominus (1 - x)) \tag{2.7}$$

Problem 1. Solve the equation $x \ominus y = x \boxminus y$ for $x, y \in [0, 1]$.
SOLUTION: $x = 0; x = 1; y = 0; y = 1$.

Problem 2. Solve the equation $x \owedge x = x$ (or $x \ovee x = x$) for $x, y \in [0, 1]$.
SOLUTION: $x = 0; x = 1$.

Problem 3. Solve the equation $x \owedge (1 \ominus x) = 0$ (or $x \ovee (1 \ominus x) = 1$) for $x, y \in [0, 1]$.
SOLUTION: $x = 0; x = 1$.

Problem 4. Solve the equation $x \twoheadrightarrow y = (1 \ominus x) \ovee y$ for $x, y \in [0, 1]$.
SOLUTION: $x = 0; x = 1; x = (y^2 + y - 1)/(y - 1); y = 0; y = 1$.

Problem 5. Solve the equation $x \owedge y + x \owedge (1 \ominus y) = x$.
SOLUTION: $x = 0; x = 1; y = 0; y = 1$.

Explanation In some sense Pei Wang's truth function $\langle f, c\rangle$ accepts the frequency explanation of probability but with no need of convergence, where f refers to the frequency and c the confidence of the frequency.

Actually, in the graph, the weight of the edge should better be considered as conditional probability. Pei Wang's truth function for $\wedge$, which equals $f_1 \times f_2$, implicates the independence of the probability distribution, which is inappropriate. Since we would like to give the compactly interconnected structure a new name, the positive effect of one event to another should be paid more attention to, so my truth function for $\wedge$ is a little larger than Pei Wang's, and the truth functions for *implication, difference* agree with the conditional probability, although, rigorously speaking, f can't be regarded as probability, because the background knowledge of the probability distribution may change in a real-time open system, that's just the very reason that my truth function fail to satisfy $x \ominus y = x \boxminus y$. Pei Wang [P.Wang09] have argued that $P_K(\phi|\psi)$ should not be written as $P(\phi|K, \psi)$, because the new knowledge ψ may fall out of the sample space of the distribution P_K. Under the AIKR, new knowledge and problems are out of the reach of Probability Theory. On one hand, Cox theorem shows that from several intuitively reasonable postulations Kolmogorov axioms of probability theory can be proved to be the unique valid calculus for degree of belief, and plausibilities follow the same rules as limiting frequencies. Any agent who violate the Kolmogorov axioms will suffer from sure loss according to Dutch Book argument, so probability theory can't be freely abandoned without pains. On the other hand, Pei Wang [P.Wang11] have argued that new models based on AIKR can't be obtained by minor revisions or extensions of the traditional models. How to compromise the dilemma? Probability theory is a local instrument, and should be treated locally rather than globally according to our bounded rationality, that is why the truth functions are defined as above.

From the solutions to the above problems, it can be seen that the truth functions agree with classical propositional logic for the ideal cases $\langle 1, 1 \rangle$.

Definition 3 (Expectation Function)**.**

$$e(\langle f, c \rangle) := c \times (f - 1/2) + 1/2$$
$$\langle f, c \rangle \leq \langle f', c' \rangle := e(\langle f, c \rangle) \leq e(\langle f', c' \rangle)$$

For two terms S and P, if $\phi[S]$ $\langle f, c \rangle$ and $\phi[P/S]$ $\langle f', c' \rangle$, then we define $S \leq_\phi P := \langle f, c \rangle \leq \langle f', c' \rangle$, where $\phi[S]$ means S occurs in ϕ, and $\phi[P/S]$ means the replacement of S with P in ϕ.

Definition 4. *The truth values of the inference rules are defined as follows:*

Inference	*Function*
F_{int} *intersection*	$f = f_1 \owedge f_2$ $c = and(c_1, c_2)$
F_{uni} *union*	$f = f_1 \ovee f_2$ $c = and(c_1, c_2)$
F_{dif} *difference*	$f = f_1 \ominus f_2$ $c = and(c_1, c_2)$
F_{neg} *negation*	$f = 1 \ominus f_1$ $c = c_1$
F_{trf} *transform*	$f = \circledcirc f_1 c_1$ $c = 0$
F_{cnt} *contraposition*	$f = \circledcirc f_1$ $c = c_1$
F_{ded} *deduction*	$f = f_1 \owedge f_2$ $c = (f_1 \owedge f_2) c_1 c_2$
F_{ana} *analogy*	$f = f_1 \owedge f_2$ $c = f_2 c_1 c_2$
F_{res} *resemblance*	$f = f_1 \owedge f_2$ $c = (f_1 \ovee f_2) c_1 c_2$
F_{abd} *abduction*	$f = f_{max} \multimap f_{min}$ $c = (f_1 \owedge f_2) c_1 c_2$
F_{ind} *induction*	$f = f_{min} \multimap f_{max}$ $c = (f_1 \owedge f_2) c_1 c_2$
$F_{abd'}$ *abduction'*	$f = (f_1 \owedge f_2)(1 - c_{min})^{\frac{1}{k}}$ $c = c_1 c_2$

$F_{ind'}$ *induction'*	$f = (f_1 \oslash f_2)(1 - c_{max})^{\frac{1}{k}}$ $c = c_1 c_2$
F_{exe} *exemplification*	$f = (f_1 \oslash f_2) c_1 c_2$ $c = ((f_1 \oslash f_2) c_1 c_2)/((f_1 \oslash f_2) c_1 c_2 + k)$
F_{com} *comparison*	$f = (f_1 \twoheadrightarrow f_2) \oslash (f_2 \twoheadrightarrow f_1)$ $c = (f_1 \oslash f_2) c_1 c_2$
F_{cnv} *conversion*	$f = f_1 c_1$ $c = (f_1 c_1)/(f_1 c_1 + k)$

Table 2: Truth Function

where $f_{min} = f_i$ $(i \in \{1, 2\})$ *s.t.* $e(\langle f_i, c_i \rangle) = \min\{e(\langle f_1, c_1 \rangle), e(\langle f_2, c_2 \rangle)\}$, *and similarly* $f_{max}, c_{min}, c_{max}$.

Definition 5 (Revision Function and Its Inverse Function)**.** *Given* $\langle f_1, c_1 \rangle$ *and* $\langle f_2, c_2 \rangle$, *revision function* $F_{rev}(\langle f_1, c_1 \rangle, \langle f_2, c_2 \rangle) = \langle f, c \rangle$, *where*

$$f = \begin{cases} 1 & \text{if } \langle f_1, c_1 \rangle = \langle f_2, c_2 \rangle = \langle 1, 1 \rangle \\ [f_1 c_1 (1 - c_2) + f_2 c_2 (1 - c_1)]/[c_1(1 - c_2) + c_2(1 - c_1)] & \text{otherwise} \end{cases}$$

$$c = \begin{cases} 1 & \text{if } \langle f_1, c_1 \rangle = \langle f_2, c_2 \rangle = \langle 1, 1 \rangle \\ [c_1(1 - c_2) + c_2(1 - c_1)]/(1 - c_1 c_2) & \text{otherwise} \end{cases}$$

Conversely, given $\langle f, c \rangle$ *and* $\langle f_1, c_1 \rangle$, *the inverse function of revision function* F_{rev} *can be computed* $F_{ivr}(\langle f, c \rangle, \langle f_1, c_1 \rangle) = \langle f_2, c_2 \rangle$, *where*

$$f_2 = \begin{cases} 1 & \text{if } c = c_1 \\ (cf - cc_1 f - c_1 f_1 + cc_1 f_1)/(c - c_1) & \text{otherwise} \end{cases}$$

$$c_2 = \begin{cases} 1 & \text{if } c_1 = 1/(2 - c) \\ (c - c_1)/(1 - 2c_1 + cc_1) & \text{otherwise} \end{cases}$$

Theorem 2 (Revision on Transformation)**.**

$$F_{rev}(\langle f, c \rangle, F_{trf}(\langle f, c \rangle)) = \langle f, c \rangle$$

§2.3. Inference Rules

Some of the syntax of the reference rules (table3,table4) Pei Wang [P.Wang12] used are kept as they were. However, the truth values of the rules need to be modified, and some other rules need to be added.

	$M \to P \langle f_1, c_1 \rangle$	$P \to M \langle f_1, c_1 \rangle$	$M \leftrightarrow P \langle f_1, c_1 \rangle$
$S \to M \langle f_2, c_2 \rangle$	$S \to P \langle F_{ded} \rangle$ $P \to S \langle F_{exe} \rangle$	$S \to P \langle F_{abd} \rangle$ $(P \to S \langle F_{abd'} \rangle)$ $S \leftrightarrow P \langle F_{com} \rangle$	$S \to P \langle F_{ana} \rangle$
$M \to S \langle f_2, c_2 \rangle$	$S \to P \langle F_{ind} \rangle$ $(P \to S \langle F_{ind'} \rangle)$ $S \leftrightarrow P \langle F_{com} \rangle$	$S \to P \langle F_{exe} \rangle$ $P \to S \langle F_{ded} \rangle$	$P \to S \langle F_{ana} \rangle$
$S \leftrightarrow M \langle f_2, c_2 \rangle$	$S \to P \langle F_{ana} \rangle$	$P \to S \langle F_{ana} \rangle$	$S \leftrightarrow P \langle F_{res} \rangle$

Table 3: Two-premise Syllogistic Rules

$S \to P$	$P \to S$	F_{cnv}
$\phi \Rightarrow \psi$	$\neg\psi \Rightarrow \neg\phi$	F_{cnt}

Table 4: One-premise Syllogistic Rules

Raven Paradox In Classical Logic, the following deduction holds.
$\because A(x) \wedge B(x)$ confirms $\forall x(A(x) \to B(x))$
$\therefore \neg A(x) \wedge \neg B(x)$ confirms $\forall x(\neg B(x) \to \neg A(x))$
$\because \forall x(A(x) \to B(x)) \leftrightarrow \forall x(\neg B(x) \to \neg A(x))$
$\therefore \neg A(x) \wedge \neg B(x)$ confirms $\forall x(A(x) \to B(x))$

If $A(x)$ stands for x *is a Raven*, and $B(x)$ stands for x *is Black*, the above conclusion is what is called "Raven Paradox". As to "Raven Paradox", Pei Wang [P.Wang09] insists that "non-raven" are irrelevant evidence according to Nicod's criterion, and that the frequency f of F_{cnt} equals 0. But in my opinion, if the extension T^E of T is still defined as $\{X \in KB : X \to T\}$ and the intension T^I of T is still defined as $\{X \in KB : T \to X\}$, then inevitably the "non-raven" cases are implicate in the intensional evidence, so the truth value of the contraposition of a statement is increasing with the original statement, and they are equivalent when the frequency is 1. That's why I use the truth function F_{cnt}.

Definition 6 (Deduction)**.** *deduction rule.*

$$\frac{\begin{array}{c} M \to P \langle f_1, c_1 \rangle \\ S \to M \langle f_2, c_2 \rangle \end{array}}{S \to P \langle F_{ded} \rangle}$$

$$\frac{\begin{array}{c} \forall X((X \to M) \Rightarrow (X \to P)) \langle f_1, c_1 \rangle \\ \forall X((X \to S) \Rightarrow (X \to M)) \langle f_2, c_2 \rangle \end{array}}{\forall X((X \to S) \Rightarrow (X \to P)) \langle F_{ded} \rangle}$$

$$\frac{\begin{array}{c}\forall X((P \to X) \Rightarrow (M \to X))\ \langle f_1, c_1\rangle \\ \forall X((M \to X) \Rightarrow (S \to X))\ \langle f_2, c_2\rangle\end{array}}{\forall X((P \to X) \Rightarrow (S \to X))\ \langle F_{ded}\rangle}$$

Definition 7 (Induction)**.** *induction rule.*

$$\frac{\begin{array}{c}M \to P\ \langle f_1, c_1\rangle \\ M \to S\ \langle f_2, c_2\rangle \\ \langle f_2, c_2\rangle \le \langle f_1, c_1\rangle\end{array}}{\forall X((X \to S) \Rightarrow (X \to P))\ \langle F_{ind}\rangle}$$

$$\frac{\begin{array}{c}M \to P\ \langle f_1, c_1\rangle \\ M \to S\ \langle f_2, c_2\rangle \\ \langle f_2, c_2\rangle > \langle f_1, c_1\rangle\end{array}}{\forall X((X \to S) \Rightarrow (X \to P))\ \langle F_{ind'}\rangle}$$

The extensional and intensional cases are similar to definition6.

Definition 8 (Abduction)**.** *abduction rule.*

$$\frac{\begin{array}{c}P \to M\ \langle f_1, c_1\rangle \\ S \to M\ \langle f_2, c_2\rangle \\ \langle f_1, c_1\rangle \le \langle f_2, c_2\rangle\end{array}}{\forall X((P \to X) \Rightarrow (S \to X))\ \langle F_{abd}\rangle}$$

$$\frac{\begin{array}{c}P \to M\ \langle f_1, c_1\rangle \\ S \to M\ \langle f_2, c_2\rangle \\ \langle f_1, c_1\rangle > \langle f_2, c_2\rangle\end{array}}{\forall X((P \to X) \Rightarrow (S \to X))\ \langle F_{abd'}\rangle}$$

The extensional and intensional cases are similar to definition6.

Remark: This rule complies with the following intuition.

$$\frac{\begin{array}{c}\phi \Rightarrow \psi \text{ is plausible (high } \langle f_1, c_1\rangle) \\ \psi \text{ is likely to be true (high } f_2) \text{ although with little evidence (low } c_2)\end{array}}{\phi \text{ is more likely to be true (high } f)}$$

Definition 9 (Extension *vs.* Intension)**.** *Most of the time it is hard to distinguish extension from intension, both of them merge in our cognitive process.*

$$\frac{\begin{array}{c}\forall X((X \to S) \Rightarrow (X \to P))\ \langle f_1, c_1\rangle \\ \forall X((P \to X) \Rightarrow (S \to X))\ \langle f_2, c_2\rangle\end{array}}{S \to P\ \langle F_{rev}\rangle}$$

The extension case $\forall X((X \to S) \Rightarrow (X \to P))$ *can be rewritten as* $S \twoheadrightarrow P$, *and the intension case* $\forall X((P \to X) \Rightarrow (S \to X))$ *can be rewritten as* $S \rightarrowtail P$, *as a result, the above rule can be rewritten as:*

$$\frac{S \twoheadrightarrow P \ \langle f_1, c_1 \rangle \quad S \rightarrowtail P \ \langle f_2, c_2 \rangle}{S \rightarrow P \ \langle F_{rev} \rangle}$$

Thanks to induction, we get the extensional direction $S \twoheadrightarrow P$; thanks to abduction, we get the intensional direction $S \rightarrowtail P$; thanks to the revision of the extensional direction and intensional direction we get the simple inheritance relationship $S \rightarrow P$.

Sometimes deduction,induction or abduction inference are executed extensionally and intensionally simultaneously.

$$\frac{M \twoheadrightarrow P \ \langle f_1, c_1 \rangle \quad S \rightarrowtail M \ \langle f_2, c_2 \rangle}{S \rightarrow P \ \langle F_{ded} \rangle} \qquad \frac{M \twoheadrightarrow P \ \langle f_1, c_1 \rangle \quad M \rightarrowtail S \ \langle f_2, c_2 \rangle}{S \rightarrow P \ \langle F_{ind} \rangle} \qquad \frac{P \twoheadrightarrow M \ \langle f_1, c_1 \rangle \quad S \rightarrowtail M \ \langle f_2, c_2 \rangle}{S \rightarrow P \ \langle F_{abd} \rangle}$$

However, if we intend to distinguish them, it is not impossible when one direction is clear.

$$\frac{S \twoheadrightarrow P \ \langle f_1, c_1 \rangle \quad S \rightarrow P \ \langle f_2, c_2 \rangle}{S \rightarrowtail P \ \langle F_{ivr} \rangle} \qquad \frac{S \rightarrowtail P \ \langle f_1, c_1 \rangle \quad S \rightarrow P \ \langle f_2, c_2 \rangle}{S \twoheadrightarrow P \ \langle F_{ivr} \rangle}$$

The other inference rules, such as analogy, resemblance, exemplification, comparison, conversion, contraposition are similarly treated like the above ones. But the followings are quite different from Pei Wang's.

Definition 10 (Extension to Intension *vs.* Intension to Extension)**.** *Given the subset relationship of S and P, sometimes we could not help guessing the properties of S with the help of the properties of P, and sometimes vice versa.*

$$\frac{S \rightarrowtail P \ \langle f, c \rangle}{S \twoheadrightarrow P \ \langle F_{trf} \rangle} \qquad \frac{S \twoheadrightarrow P \ \langle f, c \rangle}{S \rightarrowtail P \ \langle F_{trf} \rangle}$$

The reason we use the function F_{trf} is as follows: when we are pretty sure about the extension $\langle 1, 1 \rangle$ then we expect/guess/hypothesize its intension is $F_{trf}(\langle 1, 1 \rangle) = \langle 1, 0 \rangle$; when we know nothing about the extension $\langle -, 0 \rangle$, we are totally innocent of its intension $F_{trf}(\langle 0, 0 \rangle) = \langle 0, 0 \rangle$; when we know something about its extension, our knowledge of its intension should be discounted (by some system parameter k). And the function f should be monotonously increasing. When we want to estimate the normal relation $\rightarrow$ with the revision function F_{rev} through its extension $\twoheadrightarrow$ and its hypothetical intension $\rightarrowtail$ by F_{trf}, the truth value should not be increased by our guess, that is what theorem2 says.

Definition 11 (Composition Rule)**.** *The composition rules are different from Pei Wang's because the truth functions of intersection,union and difference are different. And rules of $\times, \{\}, \in$ are added.*

$$\frac{\begin{array}{c} M \to T_1 \ \langle f_1, c_1 \rangle \\ M \to T_2 \ \langle f_2, c_2 \rangle \end{array}}{M \to T_1 \cap T_2 \ \langle F_{int} \rangle} \qquad \frac{\begin{array}{c} M \to T_1 \ \langle f_1, c_1 \rangle \\ M \to T_2 \ \langle f_2, c_2 \rangle \end{array}}{M \to T_1 \cup T_2 \ \langle F_{uni} \rangle}$$

$$\frac{\begin{array}{c} T_1 \to M \ \langle f_1, c_1 \rangle \\ T_2 \to M \ \langle f_2, c_2 \rangle \end{array}}{T_1 \cap T_2 \to M \ \langle F_{uni} \rangle} \qquad \frac{\begin{array}{c} T_1 \to M \ \langle f_1, c_1 \rangle \\ T_2 \to M \ \langle f_2, c_2 \rangle \end{array}}{T_1 \cup T_2 \to M \ \langle F_{int} \rangle}$$

$$\frac{\begin{array}{c} M \to T_1 \ \langle f_1, c_1 \rangle \\ M \to T_2 \ \langle f_2, c_2 \rangle \end{array}}{M \to T_1 - T_2 \ \langle F_{dif} \rangle} \qquad \frac{\begin{array}{c} T_1 \to M \ \langle f_1, c_1 \rangle \\ T_2 \to M \ \langle f_2, c_2 \rangle \end{array}}{T_1 - T_2 \to M \ \langle F_{dif} \rangle}$$

$$\frac{\begin{array}{c} M \to T_1 \cup T_2 \ \langle f_1, c_1 \rangle \\ M \to T_1 \ \langle f_2, c_2 \rangle \end{array}}{M \to T_2 \ \langle F_{dif} \rangle} \qquad \frac{\begin{array}{c} T_1 \cap T_2 \to M \ \langle f_1, c_1 \rangle \\ T_1 \to M \ \langle f_2, c_2 \rangle \end{array}}{T_2 \to M \ \langle F_{dif} \rangle}$$

$$\frac{\begin{array}{c} M \to T_1 - T_2 \ \langle f_1, c_1 \rangle \\ M \to T_2 \ \langle f_2, c_2 \rangle \end{array}}{M \to T_1 \ \langle F_{uni} \rangle} \qquad \frac{\begin{array}{c} T_1 - T_2 \to M \ \langle f_1, c_1 \rangle \\ T_2 \to M \ \langle f_2, c_2 \rangle \end{array}}{T_1 \to M \ \langle F_{uni} \rangle}$$

$$\frac{\phi \ \langle f, c \rangle}{\neg\phi \ \langle F_{neg} \rangle} \qquad \frac{\neg\phi \ \langle f, c \rangle}{\phi \ \langle F_{neg} \rangle}$$

$$\frac{\{T\} \to P \ \langle f, c \rangle}{T \times P \to \in \ \langle f, c \rangle} \qquad \frac{T \times P \to \in \ \langle f, c \rangle}{\{T\} \to P \ \langle f, c \rangle}$$

$$\frac{\forall X (X \to P \Rightarrow T \to X) \ \langle f, c \rangle}{\{T\}^{-1} \times P \to \in \ \langle f, c \rangle} \qquad \frac{T \times P \to \in \ \langle f, c \rangle}{\forall X (X \to P \Rightarrow \{T\} \to X) \ \langle f, c \rangle}$$

$$\frac{\{\{T\}^{-1}\}}{T} \qquad \frac{T}{\{\{T\}^{-1}\}} \qquad \frac{\{\{T\}\}^{-1}}{T} \qquad \frac{T}{\{\{T\}\}^{-1}}$$

$$\frac{\{\phi\} \to \{\psi\} \ \langle f, c \rangle}{\phi \Rightarrow \psi \ \langle f, c \rangle} \qquad \frac{\phi \Rightarrow \psi \ \langle f, c \rangle}{\{\phi\} \to \{\psi\} \ \langle f, c \rangle}$$

$$\frac{\begin{array}{c} S_1 \to P_1 \ \langle f_1, c_1 \rangle \\ S_2 \to P_2 \ \langle f_2, c_2 \rangle \end{array}}{S_1 \times S_2 \to P_1 \times P_2 \ \langle F_{int} \rangle}$$

$$Y \times \{X : \phi[X] \ \langle f, c \rangle \ \& \ \langle f, c \rangle > \langle 0, 0 \rangle\} \to \in \ \langle f', c' \rangle \Leftrightarrow \phi[Y] \ \langle f', c' \rangle$$

The extensional and intensional cases are similar to definition6, as well as the propositional level $\wedge, \vee$ *are treated like* $\cap$ *and* $\cup$.

§2.4. Deduction

Definition 12 (Deductive Consequence and Theorems)**.** *A deduction is a sequence of formulas, each of which is either an axiom or a hypothesis or derived from existing formulas by the inference rules. A statement with truth value* $\langle 1,1\rangle$ *is a theorem of ANS. What else are theorems? For arbitrary* $\langle f,c\rangle$*, if* $\bigwedge\Gamma\langle f,c\rangle \vdash \phi\ \langle 1,-\rangle$ *always holds, then we say* ϕ *is a deductive consequence of* Γ*, sometimes we write* $\Gamma\vdash\phi$ *for short. Since ANS is a dynamic system, we assume the deduction theorem holds in ANS, which means that, if we get* $\phi\vdash\psi$ *at time* t*, then* $\phi\Rightarrow\psi\ \langle 1,1\rangle$ *is added to ANS at time* $t+1$*. So after "for any* $\langle f,c\rangle, \bigwedge\Gamma\langle f,c\rangle \vdash_t \phi\ \langle 1,-\rangle$*" proved,* $\bigwedge\Gamma\to\phi\ \langle 1,1\rangle$ *is added at time* $t+1$*, and we get theorem* $\vdash_{t+1} \bigwedge\Gamma\to\phi$*.*

Remark: A real-time open system can be considered as a sequence of linear systems along the time axis. Time is discretely and potential infinitely flowing. We do not care about $\vdash_\infty$, although it is expected that the classical theorems are included in $Th(\vdash_\infty)$.

Remark: For certain $\langle f,c\rangle$, we get $\bigwedge\Gamma\langle f,c\rangle \vdash \phi\ \langle 1,c'\rangle$, which means "we believe/doubt that ϕ is true to some extent c' if we believe Γ is about f true to extent c". However, if "for arbitrary $\langle f,c\rangle, \bigwedge\Gamma\langle f,c\rangle \vdash \phi\ \langle 1,-\rangle$ is proved, we know that "whether we believe it or not, ϕ is true", which means that $\bigwedge\Gamma\Rightarrow\phi$ might be called *a priori synthetical truth*. If "for crisp value $\langle 1,-\rangle(\langle 0,-\rangle)$[2], $\bigwedge\Gamma\langle 1,-\rangle(\langle 0,-\rangle)\vdash\phi\ \langle 1,-\rangle$ is proved, we know that "whether the premise is true or false, ϕ is true", which means that $\bigwedge\Gamma\ \langle 1,-\rangle(\langle 0,-\rangle)\Rightarrow\phi$ might be called *an analytical truth*.

Remark: If $\Gamma\vdash\phi\ \langle f,c\rangle \Rightarrow \psi\ \langle f,g(c)\rangle$ and $\Gamma\vdash\psi\ \langle f,c\rangle\Rightarrow\phi\ \langle f,g'(c)\rangle$ always holds for any $\langle f,c\rangle$, where $g(1)=g'(1)=1$ and $g(g')$ is monotone, then we write $\Gamma\vdash\phi\Leftrightarrow\psi$ for short in the following.

Theorem 3.

we have a weak law of excluded middle, for any term T, M

$$T\to M\ \langle 1,1\rangle(\langle 0,1\rangle)\vdash T\to M\cup(\top - M)\ \langle 1,1\rangle$$

for any formula ϕ,ψ

$$\phi\Rightarrow\psi\ \langle 1,c\rangle(\langle 0,c\rangle)\vdash\phi\Rightarrow\psi\vee\neg\psi\ \langle 1,c^2\rangle$$

and weak law of contradiction

$$M\to T\ \langle 1,1\rangle(\langle 0,1\rangle)\vdash M\cap(\top - M)\to T\ \langle 1,1\rangle$$
$$\phi\to\psi\ \langle 1,c\rangle(\langle 0,c\rangle)\vdash\phi\wedge\neg\phi\Rightarrow\psi\ \langle 1,c^2\rangle$$

double negation

$$\vdash\phi\Leftrightarrow\neg\neg\phi$$

[2] $\langle 1,-\rangle(\langle 0,-\rangle)$ means $\langle 1,-\rangle$ or $\langle 0,-\rangle$

Proof. All of the above theorems follow from theorem1 and the inference rules. For double negation, the function g, g' equals ID. □

Remark1: Thanks to definition12 and the above theorem, the weak law of excluded middle and contradiction is true in ANS in some sense. They are of great importance to our reasoning and cognition and will be used in the definition18 of choice rule.

Remark2: Since the weak law of contradiction is in very special form (only for $\langle 1, - \rangle$ *or* $\langle 0, - \rangle$), and it needs a relevant premise, so it is actually not such a theorem that can be used everywhere, as a result, ANS do not have to face the irrelevant deduction problem.

Definition 13 (translate function)**.** *$*$ is a translate function.*
for any term T, S containing only $-, \cup, \cap$

$$T^* = p_T$$
$$(T - S)^* = \neg p_S$$
$$\cup^* = \vee$$
$$\cap^* = \wedge$$
$$\rightarrow^* = \Rightarrow$$

where p_T, p_S are special unused terms representing atomic formulas, which can be substituted by any formula ϕ.

Theorem 4. *If $\Gamma \vdash \phi$ then $\Gamma^* \vdash \phi^*$.*

Proof. The rigorous proof should be given via the principle of mathematical induction of the structure of the formulas. Here it is eliminated. Instead, some intuition will be given.

Intuitively, the terms of the theorems can be substituted by other terms, specifically, the special term $\{\phi\}$ firstly. For example,

$$T \rightarrow M \vdash T \rightarrow M \cup (\top - M)$$
$$\Downarrow$$
$$\{\phi\} \rightarrow \{\psi\} \vdash \{\phi\} \rightarrow \{\psi\} \cup \{\neg\psi\}$$
$$\Downarrow$$
$$\phi \Rightarrow \psi \vdash \phi \Rightarrow \psi \vee \neg\psi$$

□

Actually, an inverse function $\circ$ of the translate function $*$ in definition13 can be defined as follows.

Definition 14. *for any atomic formula* ϕ

$$\begin{aligned}\phi^\circ &= \{\phi\} \\ (\neg\phi)^\circ &= \top - \{\phi\} \\ \vee^\circ &= \cup \\ \wedge^\circ &= \cap \\ \Rightarrow^\circ &= \rightarrow\end{aligned}$$

However, the proposition corresponding to theorem4 do not hold, so we add the following rule to ANS.

Definition 15 (Termize Rule)**.** *The term case of the formula theorem.*

$$\frac{\phi \langle f, c\rangle}{\phi^\circ \langle f, c\rangle}$$

Specially, for the atomic formula ϕ*.*

$$\frac{\phi \langle 1, 1\rangle}{\{\phi\} \leftrightarrow \top \langle 1, 1\rangle}$$

The termize rule is introduced to take care of the search problem $X :=?$, to make sure that the search of an answer to a question can be executed on the level of terms.

Definition 16 (Syntactical Complexity)**.** *The syntactical complexity of an atomic term (i.e., word) is* 1*. The syntactical complexity of a compound term or a statement is* 1 *plus the sum of the syntactical complexity of its components.*

Remarks: A very complicated compound term T, even containing variables, can be renamed to be a atomic term S, is it rational to be counted 1 when we compute the syntactical complexity of it? Yes! The reason is that, whatever exists in the KB is rational! The rename operator $T := S$ can only be used in $\phi \langle \bar{f}, \bar{c}\rangle \rightarrow u_1/u_2$, namely, only when the compound term T is very compact can it be renamed.

Definition 17 (Syntactical Simplicity)**.** *If the syntactic complexity of a term or a statement is* n*, then its syntactic simplicity is* $s = 1/n^r$*, where* $r > 0$ *is a system parameter.*

Definition 18 (Choice Rule)**.** *If the syntactic simplicity of a statement* $\phi_i[S]$ *is* s*, then its choice value is* $ch(\phi_i\langle f, c\rangle) := s \times e'$*, where* $\phi_j = \phi_i[T/S]$ *for some term* T *and*

$$e' = \begin{cases} c \times (f - 1/n) + 1/n & \text{if} \vdash \bigvee_{i=1}^{n} \phi_i \langle 1, -\rangle \\ e(\langle f, c\rangle) & \text{otherwise} \end{cases}$$

When there are two different statements ϕ, ψ *both of which are candidate answers to a question, the one with higher choice value* ch *is preferred.*

For two formulas $\phi \langle f, c\rangle$ *and* $\psi \langle f', c'\rangle$*, we define* $\{\phi \langle f, c\rangle\} \leq \{\psi \langle f', c'\rangle\} := ch(\phi \langle f, c\rangle) \leq ch(\psi \langle f', c'\rangle)$.

Remark: For special cases, for example, whether the sun will rise or not tomorrow, Pei Wang's expectation function $e(\langle f, c\rangle)$ corresponds to Carnap's λ-continuum, however, if the possibilities of an event are more than two, Carnap's λ-continuum is flexible enough to handle with it, while Pei Wang's expectation function is not very reasonable, so we expand it to e'. In other words, when we know the exact partition of the event, we use the function of Carnap, if we are totally innocent, it defaults that the event is partitioned into two parts(yes/no).

§2.5. Axioms and Reference Rules for Programs

$$\langle X \coloneqq T\rangle\phi \ \langle f, c\rangle \Leftrightarrow \phi[T/X] \ \langle f, c\rangle$$

where $\phi[T/X]$ means the substitution of T for X in ϕ.

Specially, to change the truth value of a statement $\langle X \coloneqq T\rangle\phi \ \langle X\rangle \Leftrightarrow \phi \ \langle T\rangle$

Convention:$\phi \ \langle f, c\rangle$ in this subsection means the truth value F of ϕ satisfies $F \geq \langle f, c\rangle$.

$$\begin{aligned}
&[X \coloneqq ?]\phi \ \langle f, c\rangle \Leftrightarrow \forall X\phi \ \langle f, c\rangle \\
&[\pi](\phi \Rightarrow \psi) \ \langle f_1, c_1\rangle \wedge [\pi]\phi \ \langle f_2, c_2\rangle \Rightarrow [\pi]\psi \ \langle F_{ded}\rangle \\
&[\pi](\phi \vee \psi) \ \langle f_1, c_1\rangle \wedge [\pi]\phi \ \langle f_2, c_2\rangle \Rightarrow [\pi]\psi \ \langle F_{dif}\rangle \\
&[\pi]\phi \ \langle f_1, c_1\rangle \wedge [\pi]\phi \ \langle f_2, c_2\rangle \Rightarrow [\pi](\phi \wedge \psi) \ \langle F_{int}\rangle \\
&[\pi_1; \pi_2]\phi \ \langle f, c\rangle \Leftrightarrow [\pi_1][\pi_2]\phi \ \langle f, c\rangle \\
&[\pi_1 \cup \pi_2]\phi \ \langle f, c\rangle \Leftrightarrow [\pi_1]\phi \ \langle f, c\rangle \wedge [\pi_2]\phi \ \langle f, c\rangle \\
&[\pi_1 \cap \pi_2]\phi \ \langle f, c\rangle \Leftrightarrow [\pi_1; \pi_2]\phi \ \langle f, c\rangle \wedge [\pi_2; \pi_1]\phi \ \langle f, c\rangle \\
&[\pi^*]\phi \ \langle f, c\rangle \Leftrightarrow \phi \ \langle f, c\rangle \wedge [\pi][\pi^*]\phi \ \langle f, c\rangle \\
&[?\phi \ \langle f, c\rangle]\psi \ \langle f_1, c_1\rangle \Leftrightarrow (\phi \ \langle f, c\rangle \Rightarrow \psi \ \langle f_1, c_1\rangle) \\
&[\pi^*](\phi \ \langle f, c\rangle \Rightarrow [\pi]\phi \ \langle f, c\rangle) \Rightarrow (\phi \ \langle f, c\rangle \Rightarrow [\pi^*]\phi \ \langle f, c\rangle)
\end{aligned}$$

$$\frac{\langle X \coloneqq ?\rangle\phi \ \langle X\rangle}{\nu X\phi \ \langle X\rangle}$$

$$\frac{\langle X \coloneqq ?\rangle\phi}{\nu X\phi}$$

Where ν is based on $\leq$ of $\langle f, c\rangle$ defined in definition3.

To answer an question $? \Rightarrow \phi$, we search the candidate solutions of it on the level of terms via the termize rule(Definition15).

$$\frac{\langle X :=?\rangle(X \to \{\phi\})}{\nu X(X \to \{\phi\})}$$

Where ν is based on $\le$ of $\{\phi\langle f, c\rangle\}$ defined in definition18, so $\nu X(X \to \{\phi\})$ should not be confused with $\nu X\phi$ because they are based on different $\le$, the same notation ν is used here only for simplicity. In fact, this is the formalization of *choice rule*18.

$$\frac{\phi \Rightarrow [\pi^n]\psi\ \langle f, c\rangle \qquad n \in \omega}{\phi \Rightarrow [\pi^*]\psi\ \langle f, c\rangle}$$

Remark: Most of the programs are similar to those in DL, yet not all of them can be eliminated by means of syntactic rewriting. The most important difference is $\langle X :=?\rangle$, which is defined to determine how to choose the answers to questions. So $\langle X :=?\rangle\phi \Leftrightarrow \neg[X :=?]\neg\phi$ do not hold in ANS, yet for the other programs π, $\langle\pi\rangle\phi \Leftrightarrow \neg[\pi]\neg\phi$ still holds. Besides, $\langle X :=?\rangle$ can still be compared with $\exists X\phi$. Whenever $\exists X\phi$, the term which most probably satisfies ϕ in our KB should be picked out.

§2.6. Axioms and Reference Rules for Updates

There is no free variables in the KB initially, if we want to execute abstract reasoning, we need to turn to $S := X$ for help.

$$\{S := X\}\phi \Leftrightarrow \phi[X/S] \qquad \text{(abstraction)}$$

where $\phi[X/S]$ means the substitution of X for S in ϕ. $S := X$ is the opposite of $X := S$, and we have $\vdash \langle X := S\rangle\{S := X\}\phi \Leftrightarrow \phi$.

$$\frac{\{T := S\}\phi\ \langle f, c\rangle}{\begin{array}{c}\phi[S/T]\ \langle f, c\rangle \\ T \leftrightarrow S\ \langle 1, 1\rangle\end{array}} \qquad \text{(name)}$$

where $\phi[S/T]$ means the substitution of S for T in ϕ whenever T occurs in ϕ. $T := S$ can be regarded as a name operation.

Remark: The following is invalid. We can't change the truth value by update operators. Truth value can only be changed by programs, because real numbers are rigid terms.

$$\frac{\{T := S\}\phi\ \langle T\rangle}{\phi\ \langle S\rangle}$$

$$\frac{\{T \coloneqq \bot\}\phi \ \langle f, c\rangle}{\begin{array}{c} T \leftrightarrow \bot \ \langle 1, 1\rangle \\ \{\phi\} \leftrightarrow \bot \ \langle 1, 1\rangle \end{array}} \quad \text{(delete)}$$

whenever T occurs in ϕ.

$\bot$ can be seen as a refuse container. Whenever an edge leading to $\bot$, the edge (ϕ) together with the point (T) can be abandoned.

From $\{T \coloneqq S\}$ we have $T \leftrightarrow S$, which means

$$\frac{\{T \coloneqq S\}M \leftrightarrow M}{T \leftrightarrow S}$$

However, from $T :\rightarrow S$ we only get $T \rightarrow S$. So it can be seen as a weak name operation. When $\{T \coloneqq S\}\phi$ executed, all of the T in ϕ should be replaced by S, while $\{T :\rightarrow S\}\phi$ executed, ϕ should be updated if $T \rightarrow S$ is a sub-formula of ϕ.

$$\frac{\{T :\rightarrow S\}\phi \ \langle f, c\rangle}{\begin{array}{c} T \rightarrow S \ \langle 1, 1\rangle \\ \phi[T \rightarrow S] \ \langle -, -\rangle \end{array}} \quad \text{(action)}$$

Where $\phi[T \rightarrow S] \ \langle -, -\rangle$ means the truth value of ϕ is updated by $T \rightarrow S \ \langle 1, 1\rangle$.

$$\frac{\begin{array}{c} \phi \ \langle f, c\rangle \\ \langle f, c\rangle \geq \langle \bar{f}, \bar{c}\rangle \\ \{\phi_{\geq\langle \bar{f}, \bar{c}\rangle} \rightarrow u_1/u_2\}\psi \end{array}}{\{u_1\}\psi} \ \text{(condition)}$$

$$\frac{\begin{array}{c} \phi \ \langle f, c\rangle \\ \langle f, c\rangle < \langle \bar{f}, \bar{c}\rangle \\ \{\phi_{\geq\langle \bar{f}, \bar{c}\rangle} \rightarrow u_1/u_2\}\psi \end{array}}{\{u_2\}\psi} \ \text{(condition)}$$

The abstraction, name, delete, action and condition rules are given for the formulas, in fact, the above rules for terms can be similarly treated. The compound updates can be eliminated by means of syntactic rewriting, but the complicate cases will be handled in section4, after section3 makes their semantics more clear.

§3. Semantics of ANS

Formal semantics of a language should be formally defined in its meta-language. The experience-grounded semantics can't be formally figured out explicitly, from the view of its whole lifespan, NAL can be seen as a total syntactic system, so is ANS, both of whose meta-language are hard to figure out. Perhaps its semantics is very hard to define because it is not *one* logic system, given the initial KB and the exact time stamp of the input, there is one logic system, and the KB and the input can be seen as axioms indexed by the time stamp at that exact moment. However, when talking about semantics we can't count the time flow in. However, to distinguish a semantic level from the syntactic level, it is not impossible. But this paper cannot solve this problem and can only put forward a semi-formal semantics, according to which both of the logic part $S \to P$ and the "truth value" part $\langle f, c \rangle$ are assumed to be syntactic, and the semantics of ANS is considered to consist two inseparable part. The first part is a Kripke Structure which mainly takes care of the programs and updates, and the second part is $\mathfrak{R}$ which takes care of the truth functions, so $\mathfrak{R}$ is responsible for the operations of the frequency f and the confidence c. All of the truth functions can be easily explained in $\mathfrak{R}$, and all of the reference rules, as well as the operators, are determined by the truth functions. But what is the meaning of $\vDash^t S \to P \langle f, c \rangle$? It means we have reason to believe that S is about f *part of* P to extent c at time t, which cannot be explicitly represented in a formal meta-language, so we assume it can be formalized by $S \subset^t P \langle f, c \rangle$. What is the meaning of *part of*? It may approximately be seen as *subset* relation, but the degree of *subset* is determined by the truth function other than the elements of the set.

Kripke Structures A Kripke Structure $\mathcal{K}$ in this paper is a quadruple (D, M, W, ρ) consisting of a non-empty domain D, a set S of states, a set W of variable assignments, and a program relation ρ such that:

- D is a non-empty domain rich enough to contain KB_0^*, where KB_0^* is the closure of KB_0 under all of the *term-formation* operators, and KB_0 is the initial Knowledge Base consisting all of the terms and relations between terms. KB_t may evolve with time flow but D remains invariant, which means D contains all of the terms that are conceivable and all of the terms that are forgotten, and $\bigcup_{t=0}^{\infty} KB_t \subset D$.

- M is a set of states, consisting of all interpretations I that $I(S) \in D$. Based on the assumption of D, any term S can be interpreted as itself just like in the Henkin canonical model.

- W is a set of variable assignment(possible worlds) σ that $\sigma(X) \in D$.

$$\sigma(S/X)(Y) = \begin{cases} S & \text{if } X = Y \\ \sigma(Y) & \text{otherwise} \end{cases}$$

- ρ is a program relation that, for all $\sigma, \sigma' \in W$ and for any program $\pi \in Prog_{\mathscr{L}}$, $(\sigma, \sigma') \in \rho(\pi)$ iff π started in σ and terminates in σ'.

Given the formal language $\mathscr{L}$, semantics of terms,formulas,programs and updates also need to be mutually defined.

Semantics of Terms

Definition 19 (Sematntics of Terms). *Given a Kripke structure $\mathcal{K} = (D, M, W, \rho)$, for every state $I \in M$ and any variable assignment $\sigma \in W$, the valuation function $V_{I,\sigma}$ for terms is inductively defined by:*

$V_{I,\sigma}(\bot) = \bot$
$V_{I,\sigma}(\top) = \top$
$V_{I,\sigma}(X) = \sigma(X)$
$V_{I,\sigma}(S) = I(S)$
$V_{I,\sigma}(\{\phi\}) = \{\phi\}$, *the other set operaters should be treated similarly.*

$$V_{I,\sigma}(\phi\ \langle f,c\rangle \to T_1/T_2) = \begin{cases} V_{I,\sigma}(T_1) & \textit{if } I,\sigma \vDash \phi\ \langle f,c\rangle \textit{ and } \langle f,c\rangle \geq \langle \bar{f},\bar{c}\rangle \\ V_{I,\sigma}(T_2) & \textit{otherwise} \end{cases}$$

where $\bar{f}, \bar{c}$ are system parameters

$$V_{I,\sigma}(\nu X\phi) = \begin{cases} I(S) & \textit{if } I,\sigma(S/X) \vDash \phi\langle f,c\rangle\ \&\ \langle f,c\rangle = \max\{\langle f,c\rangle : I,\sigma(T/X) \vDash \phi\langle f,c\rangle\} \\ \uparrow & \textit{otherwise} \end{cases}$$

μ is similar.

$V_{I,\sigma}(\{u\}T) = V_{I',\sigma}(u)(T)$ *with* $I' = V_{I,\sigma}(u)(I)$.

Semantics of Formulas Since $\mathfrak{R}$ speaks for itself, and the necessary and sufficient condition for the formulas are hard to figure out, we'll only give some examples and can't help but pass over it.

Definition 20 (Semantics of Formulas). *Given a Kripke structure $\mathcal{K} = (D, M, W, \rho)$, $\pi \in Prog_{\mathscr{L}}$, for every state $I \in M$ and any variable assignment $\sigma \in W$, the satisfiability relation $\vDash$ for formulas is inductively defined by:*

- $I, \sigma \vDash^t S \to P\ \langle f,c\rangle \quad iff \quad S \subset^t P\ \langle f,c\rangle$

- $I, \sigma \vDash^t \{\phi\} \to \{\psi\} \langle f, c\rangle \quad iff \quad I, \sigma \vDash^t \phi \Rightarrow \psi \langle f, c\rangle$
- $I, \sigma \vDash^{t+1} \neg\phi \langle 1-f, c\rangle \quad if \quad I, \sigma \vDash^t \phi \langle f, c\rangle$
- $I, \sigma \vDash^{t+1} \phi \langle 1-f, c\rangle \quad if \quad I, \sigma \vDash^t \neg\phi \langle f, c\rangle$
- $I, \sigma \vDash^{t+1} T \to P \cup M \langle -, -\rangle \quad if \quad I, \sigma \vDash^t T \to P \langle -, -\rangle$ *and* $I, \sigma \vDash^t T \to M \langle -, -\rangle$.
- $I, \sigma \vDash^{t+1} \phi \wedge \psi \langle -, -\rangle \quad if \quad I, \sigma \vDash^t \phi \langle -, -\rangle$ *and* $I, \sigma \vDash^t \psi \langle -, -\rangle$.
- $I, \sigma \vDash^t [\pi]\phi \langle -, -\rangle \quad iff$ *for every* $\sigma' \in W$, $(\sigma, \sigma') \in \rho(\pi) \Rightarrow I, \sigma' \vDash^t \phi \langle -, -\rangle$.
- $I, \sigma^t \vDash \{u\}\phi \langle -, -\rangle \quad iff \quad I', \sigma \vDash^t \phi \langle -, -\rangle$ *with* $I' = V_{I,\sigma}(u)(I)$.
- ……

A general Tarski semantic flavour can be tasted if we ignore the time t in some of the above examples, but not all of the formulas can be explicitly analyzed to be given the semantic necessary and sufficient condition.

Semantics of Programs

Definition 21 (Semantics of Programs)**.** *Given a Kripke structure* $\mathcal{K} = (D, M, W, \rho)$, $\pi \in Prog_{\mathscr{L}}$, *for every state* $I \in M$ *and any variable assignment* $\sigma \in W$, *the program relation* ρ *for programs is inductively defined by:*

- $\rho(X := T) = \{(\sigma, \sigma(V_{I,\sigma}(T)/X) : \sigma \in W\}$
- $\rho(X :=?) = \{(\sigma, \sigma(S/X)) : S \in KB_t\}$
- $\rho(?\phi) = \{(\sigma, \sigma) : I, \sigma \vDash \phi\}$
- $\rho(\pi_1; \pi_2) = \rho(\pi_1) \circ \rho(\pi_2) =_{df} \{(\sigma_1, \sigma_2) : \exists\sigma\, ((\sigma_1, \sigma) \in \rho(\pi_1)\ \&\ (\sigma, \sigma_2) \in \rho(\pi_2))\}$
- $\rho(\pi_1 \cup \pi_2) = \rho(\pi_1) \cup \rho(\pi_2)$
- $\rho(\pi_1 \cap \pi_2) = \rho(\pi_1; p_2) \cap \rho(\pi_2; p_1)$
- $\rho(\pi^*) = \bigcup_{i \geq 0} \rho(\pi^n)$ *with* $\rho(\pi^0) = Id$ *and* $\rho(\pi^{n+1}) = \rho(\pi^n) \circ \rho(\pi)$.

Semantics of Updates A valuation of an update is a mapping from a term to another term $T \mapsto T'$ or as a set U of ordered pair (T, T').

Definition 22. *A modification* I' *of* I *can be defined as follows:*

$$I'(T) = \begin{cases} T' & \text{if } (T, T') \in U \\ I(T) & \text{otherwise.} \end{cases}$$

or it can be defined as follows:
for any (partial) function $f, g : A \to B$, $f \oplus g =_{df} (f \upharpoonright (dom(f) \setminus dom(g))) \cup g$, *then* $I' = I \oplus U$.

Definition 23 (Semantics of Updates). *Given a Kripke structure $\mathcal{K} = (D, M, W, \rho)$, $\pi \in Prog_{\mathscr{L}}$, for every state $I \in M$ and any variable assignment $\sigma \in W$, the valuation function $V_{I,\sigma}(u)$ for updates is inductively defined by:*

- $V_{I,\sigma}(S \coloneqq X) = \{(I(S), \sigma(X))\}$
- $V_{I,\sigma}(T \coloneqq S) = \{(I(T), I(S))\} \cup \{(\varnothing, \{T \leftrightarrow S\})\}$
- $V_{I,\sigma}(T \coloneqq \bot) = \{(I(T), I(\bot))\} \cup \{(\varnothing, \{T \leftrightarrow \bot\})\}$
- $V_{I,\sigma}(T :\rightarrow S) = \{(\varnothing, \{T \rightarrow S\})\}$
- $V_{I,\sigma}(u_1; u_2) = V_{I,\sigma}(u_1) \oplus V_{I',\sigma}(u_2)$ *with* $I' = V_{I,\sigma}(u_1)(I) =_{df} I \oplus V_{I,\sigma}(u_1)$
- $V_{I,\sigma}(u_1 \| u_2) = V_{I,\sigma}(u_1) \oplus V_{I,\sigma}(u_2)$
- $$V_{I,\sigma}(\phi_{\geq\langle \bar{f},\bar{c}\rangle} \rightarrow u_1/u_2) = \begin{cases} V_{I,\sigma}(u_1) & \text{if } I, \sigma \vDash \phi \; \langle f, c\rangle \text{ and } \langle f, c\rangle \geq \langle \bar{f}, \bar{c}\rangle \\ V_{I,\sigma}(u_2) & \text{otherwise} \end{cases}$$
- $$V_{I,\sigma}((\forall X \phi) u) = \begin{cases} \bigcup\{f(S) : I, \sigma(S/X) \vDash \phi\} & \text{if there exists } S,\ I, \sigma(S/X) \vDash \phi \\ I & \text{otherwise} \end{cases}$$

  with $f(S)$ *recursively defined as* $f(S) = V_{I,\sigma(S/X)}(u) \oplus \bigcup\{f(S') : S' \leq_{\phi} S \;\&\; I, \sigma(S'/X) \vDash \phi \; \langle f', c'\rangle\}$*, assuming* $\{S : I, \sigma(S/X) \vDash \phi\}$ *is well ordered by* $\leq_{\phi}$ *corresponding to definition3.*

- $V_{I,\sigma}(\{u_1\} u_2) = V_{I',\sigma}(u_2)$ *with* $I' = V_{I,\sigma}(u_1)(I)$

§4. ANS Reduction

§4.1. Axioms and Reference Rules for Updates(Continued)

Definition 24.

$$u_1 \equiv u_2 \;\; iff \text{ for any } I \in M, \sigma \in W, V_{I,\sigma}(u_1) = V_{I,\sigma}(u_2)$$
$$u_1 \equiv' u_2 \;\; iff \text{ for any } I \in M, \sigma \in W, I \oplus V_{I,\sigma}(u_1) = I \oplus V_{I,\sigma}(u_2)$$

Theorem 5. *For $\alpha \in Term_{\mathscr{L}} \cup Form_{\mathscr{L}} \cup Updt_{\mathscr{L}}$, we have,*
$u \equiv' u' \Rightarrow \{u\}\alpha \equiv \{u'\}\alpha$
$\{u_2\}(\{u_1\}\alpha) \equiv \{u_1; u_2\}\alpha$
$u_1; (u_2; u_3) \equiv (u_1; u_2); u_3$
$u_1 \| (u_2 \| u_3) \equiv (u_1 \| u_2) \| u_3$
$\phi \to (u_1 \| u_2)/u_3 \equiv (\phi \to u_1/u_3) \| (\phi \to u_2/u_3)$
$\phi \to u_1/(u_2 \| u_3) \equiv (\phi \to u_1/u_2) \| (\phi \to u_1/u_3)$
$(\forall X\phi)(\psi \to u_1/u_2) \equiv \psi \to ((\forall X\phi)u_1)/((\forall X\phi)u_2)$ *for* $X \notin Fv(\psi)$

There are lots of propositions like these that can be proved valid by checking their semantics. We will write down a sequence of valid formulas in the proof of the following theorems without details. The theorems themselves can be used as reduction axioms. Actually, the following theorems need to be put together as a package to be proved mutually rather than to be proved separately, but we break them into several sub-theorems.

Definition 25 (indom)**.** *The two value* $(\langle 1,1\rangle or\langle 0,0\rangle)$ *term indom can be inductively defined in the language of ANS, but here we will not give the details. Its general meaning is* $\{M\} \times \{u\} \to indom\ iff$ *the term of* M *and the term on the left of* $:=$ *or* $:\to$ *in* u *coincide with each other.*

Applying Updates to Terms and Formulas

Theorem 6. *The parallel update before terms and formulas can be eliminated.*

Proof.

$$\{u_1 \| u_2\}T \leftrightarrow ((\{T\} \times \{u_2\} \to indom)_{\geq\langle \bar{f},\bar{c}\rangle} \to \{u_2\}T/\{u_1\}T) \qquad \text{(parallel)}$$

Updates for formulas and updates for terms are similar. □

Applying Updates to Programs How do an update operator affect a programs? Can a program be updated other than rewritten to adapt to the changing environment?
Define $\pi_1 \equiv \pi_2\ iff\ \langle\pi_1\rangle\phi \Leftrightarrow \langle\pi_2\rangle\phi$ for any ϕ.

Theorem 7. *Updates before all of the programs except $X :=?$ can be eliminated.*

Proof. $\{u\}\langle X := T\rangle \equiv \langle X := \{u\}T\rangle$
$\{u\}\langle ?\phi\rangle \equiv \langle ?\{u\}\phi\rangle$
$\{u\}\langle \pi_1 \bigtriangledown \pi_2\rangle \equiv \langle \{u\}\pi_1 * \{u\}\pi_2\rangle, \bigtriangledown \in \{;, \cup, \cap\}$
$\{u\}\langle \pi^*\rangle \equiv \langle (\{u\}\pi)^*\rangle$ □

Applying Updates to Updates

Theorem 8. *The update application operator* $\{u_1\}u_2$*, the quantified update operator* $(\forall X\phi)u$ *and the sequential update operator* $u_1; u_2$ *can be eliminated.*

Proof.

$$
\begin{aligned}
& u_1; u_2 \equiv u_1 \| \{u_1\}u_2 && \text{(sequential)}\\
& (\forall X\phi)u \equiv u[\nu X\phi/X] \| \cdots \| u[\mu X\phi/X] && \text{(quantified)}\\
& \{u\}S \coloneqq X \equiv S \coloneqq \{u\}X \\
& \{u\}T \coloneqq S \equiv T \coloneqq \{u\}S \\
& \{u\}T \coloneqq \bot \equiv T \coloneqq \{u\}\bot \\
& \{u\}T :\rightarrow S \equiv T :\rightarrow \{u\}S \\
& \{u\}(u_1 \| u_2) \equiv \{u\}u_1 \| \{u\}u_2 \\
& \{u\}(\phi \rightarrow u_1/u_2) \equiv \{u\}\phi \rightarrow \{u\}u_1/\{u\}u_2 \\
& \{u\}((\forall X\phi)u') \equiv (\forall X\{u\}\phi)\{u\}u' \ for\ X \notin Fv(u)
\end{aligned}
$$

□

All of the above valid formulas in theorem6, theorem7, theorem8 can be used as reduction axioms to eliminate update operators, and every occurrence of the compound update operators can be eliminated, just like in [P.Rümmer06].

§4.2. Applications

Several examples are given as follows. From them it can be seen how the update operators help us to cognitively understand the respective problems.

Example1(Pattern Classification): What is a table? It consists three or four legs(ϕ_1); it is hard enough to sustain a man(ϕ_2); it is made of wood or metal(ϕ_3); it is $\cdots(\phi_4, \cdots)$

We put together all of the properties of table $\bigwedge_i \phi_i(table)$. Now we observe something S has property ψ. How to tell if S is a table or not?

$$\{\forall X(\{S \coloneqq X\}\psi \Leftrightarrow \{table \coloneqq X\}\bigwedge_i \phi_i) \rightarrow (S :\rightarrow table)/(S \coloneqq S)\}M \leftrightarrow M$$

Example2(Theory): What is a group? After we find some of the operations (ϕ) of + over $\mathbb{Z}$, for example, *identity, associativity, invertibility*, are very interesting, maybe we would like to give it a weak name $\{\phi\} :\rightarrow group$, after observing that some operations (ϕ) of + over $\mathbb{Z}$ and some operations (ψ) of × over $\mathbb{Z}\setminus\{0\}$ are quite similar, we abstract every constant terms ($\overrightarrow{S}$) of them $\{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\phi, \{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\psi$, then we find they are absolutely the same ting $\forall\overrightarrow{X}(\{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\phi \Leftrightarrow \{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\psi)$ $\langle 1,1\rangle$, we give it a strong name $\{\{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\phi\} \coloneqq group$. After that, we found a theory, namely, *group theory*, i.e. $\{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\phi$. By $\{\overrightarrow{S} \coloneqq \overrightarrow{X}\}\phi$ we can deduce some statement γ,

$\{\overrightarrow{S} := \overrightarrow{X}\}\phi \Rightarrow \gamma \, \langle 1,1 \rangle$, then we know $\{\psi\} \to \{\gamma\}$ and $\psi \Rightarrow \gamma$. Someday when we recognize the operations (δ) of + over Matrix is also a group, $\{\delta\} \to group$, then we know $\delta \Rightarrow \gamma$ immediately.

Example3**(Categoricity):** A Cognitive Interpretation of Non-standard Models. What is cardinal number? Is the abstract mathematical concept generated from experience? Can we capture it uniquely? According to Russell, the cardinality of X is the class $[X]$ of all sets that are equinumerous with X. In ANS, what is 3? Russell's definition can be understood as that, between all the terms X which consist of only three parts as well as all the formulas Y which character three edges and the cardinal number 3, there is an arrow $\to$, i.e. $X \cup Y \to 3$. But 3 is a weak name given by $:\to$. When there is a new term M inputted, is $M \to 3 \, \langle 1,1 \rangle$ or not? We can't tell! Because there is no unique semantic explanation for 3! We can't exclude that M has another property (ψ) that is same as some common property of $X \cup Y$. This is the "Caesar problem". Just like PA admits non-standard models because of Compactness theorem. In our daily communication, once we find the agent mistakes ψ with 3 in our mind, we can force him to distinguish them by teaching him another term N which consists of three parts but without property ψ. However, we can't excluded the potential "Caesars" once and for all. So it is hard, if not impossible, to figure out a rigorous name $\{\phi\}$ for 3 like in the above example to make sure we can make an explicit definition by $\{\phi\} := 3$. Cardinal numbers had better to be defined from an meta-level, in other words, implicitly defined by axioms.

Example4**(Context/Localization):** Are whales fish? It all depends. In biology, they are not. In daily life, average people take for granted that they are fish. But how to distinguish the context *biology* from daily life? First of all, what is biology? It is a science composed of a set of statements(Γ), $\{\Gamma\} :\to biology$, then we have $\{\Gamma\} \to biology \, \langle 1,1 \rangle$ and $biology \to \{\Gamma\} \, \langle 1, 1/(1+k) \rangle$ by *conversion*. In daily life, we talk about something $\phi[whale]$ about whale. When we want to emphasize the context of biology, what should we do? Can we find a way to talk about $whale_{bio}$ instead of $whale$?

$$\{whale := X\}\phi$$
$$\Downarrow$$
$$\langle X := (\Gamma \langle 1, 1/(1+k) \rangle \to whale_{bio}/whale) \rangle \{whale := X\}\phi$$
$$\Downarrow$$
$$\phi[(\Gamma \langle 1, 1/(1+k) \rangle \to whale_{bio}/whale \rangle)/whale]$$

What's the meaning of $\Gamma \langle 1, 1/(1+k) \rangle \to whale_{bio}/whale \rangle$ in ϕ?

$$\psi[\phi \, \langle f, c \rangle \to T_1/T_2]$$
$$\Updownarrow$$
$$(\phi \, \langle f, c \rangle \wedge \langle f, c \rangle \geq \langle \bar{f}, \bar{c} \rangle \Rightarrow \psi[T_1]) \wedge (\neg(\phi \, \langle f, c \rangle \wedge (\langle f, c \rangle \geq \langle \bar{f}, \bar{c} \rangle)) \Rightarrow \psi[T_2])$$

At last, if the context is strong, namely, $\langle 1, 1/(1+k) \rangle \geq \langle \bar{f}, \bar{c} \rangle$ then we talk about $whale_{bio}$ in biology rather than $whale$ in daily life.

If the context can be depicted by a formula ϕ directly rather than indexed by some term like *biology*, and the truth value of the context $\phi\langle f, c\rangle$ is not very high, but we still want to emphasize the context, what should we do? We can change the truth value high for the moment and change it back later.

$$\langle X := \langle 1, 1\rangle\rangle\{F := X\}\phi\langle F\rangle$$

To save the resources, most of the time, the extension and intension reference should not be executed unless necessary. Once we want to consider the extension or intension aspect of the case, we localize a big enough context, within which we make the distinction.

§5. Conclusion and Further Research

A variation (ANS) of Pei Wang's NAL is given. Some interesting update operators are introduced to make Pei Wang's self-monitor and self-control idea possible. Further work need to be done. For example,

- Formalize other self-monitor and self-control ideas in Pei Wang[P.Wang12].
- Study the interactions between the programs and the updates. Especially, how can a compound program be abstracted to a macro? How to found a theory through programs and then apply it on the level of KB? Namely, how to achieve more intelligence by self-programming?
- Compliment the reference rules, and make the semantics complete. Define and prove the soundness of ANS.
- Programming ANS.

Acknowledgments Thanks Pei Wang for teaching me every detail of his NAL for half an year and for lots of invaluable discussions.

References

[Beckert,Hähnle,Schmitt07] *Verification of Object-Oriented Software.* Springer-Verlag Berlin Heidelberg, 2007.

[D.Harel,D.Kozen,J.Tiuryn00] *Dynamic Logic.* Foundations of Computing. The MIT Press, Cambridge, Massachusetts, 2000.

[P.Rümmer06] *Sequential, Parallel, and Quantified Updates of First-Order Structures.* Springer-Verlag Berlin Heidelberg, 2006.

[P.Wang09] *Formalization of Evidence: A Comparative Study.* Journal of Arti cial General Intelligence, 2009.

[P.Wang11] *The Assumptions on Knowledge and Resources in Models of Rationality.* International Journal of Machine Consciousness. Vol.3, No.1, 2011.

[P.Wang12] *Non-Axiomatic Logic: A Model of Intelligent Reasoning.* World Scientific Publishing. to appear.