ISSN: 2769-545X

Research Article

Advances in Machine Learning & Artificial Intelligence

Intentional First Order Logic for Strong-AI Generation of Robots

Zoran Majkic*

ISRST, Tallahassee, FL, USA

***Corresponding Author**
Zoran Majkic, ISRST, Tallahassee, FL, USA.

Submitted: 2023, Apr 26; **Accepted**: 2023, May 15; **Published**: 2023, May 27

Citation: Majkic, Z. (2023). Intentional First Order Logic for Strong-AI Generation of Robots. *Adv Mach Lear Art Inte, 4*(1), 23-31.

Abstract

Neuro-symbolic AI attempts to integrate neural and symbolic architectures in a manner that addresses strengths and weaknesses of each, in a complementary fashion, in order to support robust strong AI capable of reasoning, learning, and cognitive modeling. We consider the robot's four-levels knowledge structure: The syntax level of particular natural language (Italian, French, etc..), two universal language levels: its semantic logic structure (based on virtual predicates of FOL and logic connectives), and its corresponding conceptual PRP structure level which universally represents the composite mining of FOL formulae grounded on the last robot's neuro system level. Therefore, this paper we consider the intentional First Order Logic as a symbolic architecture of modern robots, able to use natural languages to communicate with humans and to reason about their own knowledge with self-reference and abstraction language property.

Keywords: Strong AI, Intentional FOL, General Semantics

1. Introduction

The central hypothesis of cognitive science is that thinking can best be understood in terms of representational structures in the mind and computational procedures that operate on those structures. Most work in cognitive science assumes that the mind has mental representations analogous to computer data structures, and computational procedures similar to computational algorithms. Mainstream machine learning research on deep artificial neural networks may even be characterized as being behavioristic. In contrast, various sources of evidence from cognitive science suggest that human brains engage in the active development of compositional generative predictive models from their self-generated sensorimotor experiences. Guided by evolutionarily shaped inductive learning and information processing biases, they exhibit the tendency to organize the gathered experiences into event- predictive encodings. Meanwhile, they infer and optimize behavior and attention by means of both epistemic and homeostasis-oriented drives.

Knowledge representation strongly connected to the problem if knowledge processing, reasoning and "drawing inferences", is one of the main topics in AI. By reviewing the knowledge representation techniques that have been used by humans, we will be aware of the *importance of language.* The predominant part of IT industry and user's applications is based on some sublanguage of the standard (extensional) FOL (First Order Logic) with Tarski's semantics based (only) on the truth; my effort is to pass to a more *powerful evolution of the FOL* able to support the meaning of knowledge as well, by replacing the standard FOL and its DB theory and practice in IT business. All this work is summarized and extended to AI applications of many-valued logics in my recent book [1].

Last 15 years of my work in AI was mainly dedicated to development of a new intentional FOL, by integrating Montague's and algebraic Bealer's approaches, with a conservative Tarski's semantics of the standard FOL [2]. Basic result was the publication of the conservative extension of Tarski's semantics to intentional FOL, and two-step intentional semantics, which guaranteed a conservative extension of current RDB, but more than 50-years old technology, toward new IRDB (Intentional RDB) [3,4]. Indeed, in my next Manifesto of IRDB, I hoped also to find interested research groups and funds to begin the realization of IRDB as a new platform (compatible with all previously developed RDB application), able also to support New SQL for Big Data, and ready for other AI improvements [5].

It is dedicated to show how this defined IFOL in can be used for a new generation of intelligent robots, able to communicate with humans with this intentional FOL supporting the meaning of the words and their language compositions [1]. As in we can consider three natural language levels: The *syntax* of a particular natural language (French, English, etc..) its *semantic logic structure* (transformation of parts of the language sentences into the logic predicates and definition of corresponding FOL formulae) and its corresponding *conceptual structure*, which differently from the semantic layer that represents only the logic's semantics, represents the composed meaning of FOL formulae based on the grounding of intentional PRP concepts [6].

Thus, intentional mapping from the free FOL syntax algebra into the algebra of intentional PRP concepts,

$$I : \mathcal{A}_{FOL} \rightarrow \mathcal{A}_{int}$$

provided by IFOL theory, is a part of the semantics-conceptual mapping of natural languages. Note that differently from the particularity of any given natural language of humans, the underlying logical semantics and conceptual levels have universal human knowledge structure, provided by innate human brain structure able to rapidly acquire the ability to use any natural language.

Parsing, tokenizing, spelling correction, part-of-speech tagging, noun and verb phrase chunking are all aspects of natural language processing long handled by symbolic AI, and has to be improved by deep learning approaches. In symbolic AI, discourse representation theory and first-order logic have been used to represent sentence meanings. We consider that the natural language (first level) can be parsed into a logical FOL formula with a number of virtual predicates and logic connectives of the FOL. By such a parsing, we obtain the second, semantic logic, structure corresponding to some FOL formula. However, natural language is grounded in experience. Humans do not always define all words in terms of other words; humans understand many basic words in terms of associations with sensory-motor experiences for example. People must interact physically with their world to grasp the essence of words like "blue," "could," and "left." Abstract words are acquired only in relation to more concretely grounded terms.

Theoretical neuroscience is the attempt to develop mathematical and computational theories and models of the structures and processes of the brains of humans and other animals. If progress in theoretical neuroscience continues, it should become possible to tie psychological to neurological explanations by showing how mental representations such as concepts are constituted by activities in neural populations, and how computational procedures such as spreading activation among concepts are carried out by neural processes. Concepts, which partly correspond to the words in spoken and written language, are an important kind of mental representation.

Alan Turing developed the Turing Test in 1950 in his paper, "Computing Machinery and Intelligence". Originally known as the Imitation Game, the test evaluates if a ma- chines behavior can be distinguished from a human. In this test, there is a person known as the "interrogator" who seeks to identify a difference between computer-generated output and human-generated ones through a series of questions. If the interrogator can- not reliably discern the machines from human subjects, the machine passes the test. However, if the evaluator can identify the human responses correctly, then this eliminates the machine from being categorized as intelligent.

Differently from the *simulation* of AI by such Turing tests and the Loebner Prize and in accordance with Marvin Minsky, in this paper I argue that a real AI for robots can be obtained by using formal intentional FOL (with defined intentional algebra of intensions of language constructions) for the robots as their symbolic AI component, by defining the sense to ground terms (the words) in an analog way, associating to these words the software processes developed for the robots when they recognize by these algorithms (neural architectures) the color "blue" of visual objects, the position "left" etc... In this way, we would obtain a *neuro-symbolic AI,* which attempts to integrate neural and symbolic architectures in a manner that addresses strengths and weaknesses of each, in a complementary fashion, in order to support robust AI capable of reasoning, learning, and cognitive modeling. To build a robust, knowledge-driven approach to AI we must have the machinery of symbol-manipulation as, in this case, an IFOL. Too much of useful knowledge is abstract to make do without tools that represent and manipulate abstraction, and to date, the only machinery that we know of that can manipulate such abstract knowledge reliably is the apparatus of symbol-manipulation. Abstraction operators provide the IFOL defined as well [1].

Daniel Kahneman describes human thinking as having two components, System 1 and System 2 [7]. System 1 is fast, automatic, intuitive and unconscious. System 2 is slower, systematic, and explicit. System 1 is the kind used for pattern recognition while System 2, in uor case based on IFOL, is far better suited for planning, deduction, and deliberative thinking. In this view, deep learning best models the first kind of thinking while symbolic reasoning best models the second kind and both are needed.

So, for the words (ground linguistic terms), which cannot be "defined by other words", the robots would have some own internal experience of the concrete sense of them. Thus, by using intentional FOL the robots can formalize also the natural language expressions "I see the blue color "by a predicate" see (I, blue color)" where the sense of the ground term "I" (*Self*) for a robot is the name of the main working coordination program which activate all other algorithms (neuro-symbolic AI subprograms) like visual recognition of color of the object in focus. But also, the auto-conscience sentence like "I know that I see the blue color" by using abstracting operators " $\lessdot\_\gtrdot$ " of intentional FOL, expressed by the predicate "know (I,$\lessdot$ see(I, blue color)$\gtrdot$)", etc...

Consequently, we argue that by using this intentional FOL, the robots can develop their own knowledge about their experiences and communicate by a natural language with humans. Therefore, we would be able to develop the interactive robots, which learn and understand spoken language via multisensory grounding and internal robotic embodiment. The *grounding* of the intentional concepts I PRP theory of intentional logic was not considered in my recent book from the fact that this book was only restricted on the symbolic AI aspects (IFOL); so by this paper we extend the logic theory developed with concrete grounding of its intentional concepts in order to obtain a strong AI for robots [1]. Therefore, in next Section we will introduce IFOL and its intentional / extensional semantics [1].

1.1. Algebra for Composition of Meanings in IFOL

Contemporary use of the term "intension" derives from the traditional logical doctrine that an idea has both an extension and an intension. Although there is divergence in formulation, it is accepted that the extension of an idea consists of the subjects to which the idea applies, and the intension consists of the attributes implied by the idea. In con- temporary philosophy, it is linguistic expressions (here it is a logic formula), rather than concepts, that are said to have intensions and extensions. The intension is the concept expressed by an expression of intentional algebra $\mathcal{A}_{int}$, and the extension is the set of items to which the expression applies. This usage resembles use of Frege's use of "Bedeutung" and "Sinn" [8].

Intentional entities (or concepts) are such things as Propositions, Relations and Properties (PRP). What make them "intentional" is that they violate the principle of extensionality the principle that extensional equivalence implies identity. All (or most) of these intentional entities have been classified at one time or another as kinds of universals [9].

In a predicate logics, (virtual) predicates expresses classes (properties and relations), and sentences express propositions. Note that classes (intentional entities) are *reified*, i.e., they belong to the same domain as individual objects (particulars). This endows the intentional logics with a great deal of uniformity, making it possible to manipulate classes and individual objects in the same language. In particular, when viewed as an individual object, a class can be a member of another class.

Definition 1
Virtual Predicates
Virtual predicate obtained from an open formula $\phi \in \mathcal{L}$ is denoted by $\phi(x_1, ..., x_m)$ where (x_1, ..., x_m) is a particular fixed sequence of the set of all free variables in ϕ. This definition contains the precise method of establishing the ordering of variables in this tuple: such a method that will be adopted here is the ordering of appearance, from left to right, of free variables in ϕ. This method of composing the tuple of free variables is unique and canonical way of definition of the virtual predicate from a given open formula.

The virtual predicates are useful also to replace the general FOL quantifier on variables $(\exists x)$ by specific quantifiers $\exists_i$ of the FOL syntax algebra $\mathcal{A}_{FOL}$, where $i \geq 1$ is the position of variable x inside a virtual predicate. For example, the standard FOL formula $(\exists x_k)\phi(x_i, x_j, x_k, x_l, x_m)$ will be mapped into intentional concept $\exists_3\phi(\boldsymbol{x}) \in$ AFOL where x is the list (tuple) of variables (x_i, x_j, x_k, x_l, x_m). Virtual predicates are atoms used to build the semantic logic structures of logic-semantics level of any given natural language.

Let us define the FOL syntax algebra A_{FOL}. For example, the FOL formula $\phi(x_i, x_j, x_k, x_l, x_m) \wedge \psi(x_l, y_i, x_j, y_j)$ will be replaced by a specific virtual predicate $\phi(x_i, x_j, x_k, x_l, x_m) \wedge_S \psi\ (x_l, y_i, x_j, y_j)$, with the set of joined variables (their positions in the first and second virtual predicate, respectively) S = {(4, 1), (2, 3)}, so that its extension is expressed by an algebraic expression
where R_1, R_2 are the extensions for a given Tarski's interpretation I_T of the virtual predicate ϕ, ψ relatively, and the binary operator $\bowtie_S$ is the natural join of these two relations. In this example, the resulting relation will have the following ordering of attributes: $(x_i, x_j, x_k, x_l, x_m, y_i, y_j)$. In the case when S is empty (i.e. its cardinality $|S| = 0$) then the resulting relation is the Cartesian product of R_1 and R_2. For the existential quantification, the FOL formula $(\exists x_k)\phi(x_i, x_j, x_k, x_l, x_m)$ will be replaced in AFOL by a specific virtual predicate $(\exists_3)\phi(x_i, x_j, x_k, x_l, x_m)$. For logic negation operator we will use the standard symbol ¬. Based on the new set of logical connectives introduced above, where the standard FOL operators $\wedge$ and $\exists$ are substituted by a set of specialized operators $\{\wedge_S\}_{S \in \mathcal{P}(\mathbb{N}^2)}$ and $\{\exists n\}_{n \in \mathbb{N}}$ as explained above, we can define the following free syntax algebra for the FOL.

Definition 2
FOL Sintax Algebra
Let $\mathcal{A}_{FOL} = (\mathcal{L}, \doteq, \top, \{\wedge_S\}_{S \in \mathcal{P}(\mathbb{N}^2)}, \neg, \{\exists n\}_{n \in \mathbb{N}})$ be an extended free syntax algebra for the First-Order logic with density $\doteq$, with the set *L* of first-order logic formulae with the set of variables in *V*, with *T* denoting the tautology formula (the contradiction formula is denoted by $\perp \equiv \neg\top$).

We begin with the informal theory that universals (properties (unary relations), relations, and propositions in PRP theory) are genuine entities that bear fundamental logical relations to one another [10]. To study properties, relations and propositions, one de- fines a family of set-theoretical structures, one defines the intentional algebra, a family of set-theoretical structures most of which are built up from arbitrary objects and fundamental logical operations (conjunction, negation, existential generalization, etc..) on them.

Definition 3
Intensional Logic PRP Domain D
In intensionl logic the concepts (properties, relations and propositions) are denotations for open and closed logic sentences, thus elements of the structured domain $D = D_{-1} + D_I$, (here + is a disjoint union) where

- A subdomain D_{-1} is made of particulars (individuals).
- The rest $D_I = D_0 + D_1$…..... + D_n……….. is made of universals (concepts): D_0 for propositions with a distinct concept Truth $\in D_0$, D_1 for properties (unary concepts) and D_n; n ≥ 2; for n-ary

concept.

The concepts in D_I are denoted by *u,v,..*, while the values (individuals) in D_{-1} by *a,b,…..* The empty tuple < > of the nullary relation $r_\emptyset$ (i.e. the unique tuple of 0-ary relation) is an individual in D_{-1}, with $D^0 =_{def} \{<>\}$. Thus, we have that $\{f,t\} = \mathcal{P}(\mathcal{D}^0) \subseteq \mathcal{P}(D_{-1})$, where by *f* and *t* we denote the empty set ∅ and set {< >} respectively.

The intensional interpretation is a mapping between the set *L* of formulae of the FOL and intensional entities in $\mathcal{D}$, $I : \mathcal{L} \to \mathcal{D}$, is a kind of "conceptualization", such that an open-sentence (virtual predicate) $\phi(x_1,\ldots, x_k)$ with a tuple of all free variables $(x_1,\ldots, x_k)$ is mapped into a k-ary concept, that is, an intensional entity $u = I(\phi(x_1, ..., x_k)) \in D_k$, and (closed) sentence ψ into a proposition (i.e., logic concept) $v = I(\psi) \in D_0$ with $I(\top) = Truth \in D_0$ for the FOL tautology $\top \in \mathcal{L}$ (the falsity in the FOL is a logic formula $\neg\top \in \mathcal{L}$). A language constant *c* is mapped into a particular $a \in D_{-1}$ (intension of c) if it is a proper name, otherwise in a correspondent concept *u* in D_I. Thus, in any application of intensional FOL, this intensional interpretation that determines the meaning (sense) of the knowledge expressed by logic formulae is uniquely determined (prefixed) (for example, by a grounding on robot's neuro system processes, explained in next section). However, the extensions of the concepts (with this prefixed meaning) vary from a context (possible world, expressed by an extensionalizzation function) to another context in a similar way as for different Tarski's interpretations of the FOL.

Definition 4
Extensions and Extensionalization Functions
Let $\mathfrak{R} = \bigcup_{k\in\mathbb{N}} \mathcal{P}(\mathcal{D}^k) = \sum_{k\in\mathbb{N}} \mathcal{P}(D^k)$ be the set of all k-ary relations, where $k \in \mathbb{N} = \{0, 1, 2, ...\}$. Notice that $\{f,t\} = \mathcal{P}(\mathcal{D}^0) \subseteq \mathfrak{R}$, that is $f,t \in \mathfrak{R}$ and hence the truth-values are extensions in $\mathfrak{R}$.

We define the function $f_{<>} : \mathfrak{R} \to \mathfrak{R}$, such that for any $R \in \mathfrak{R}$,

$$f_{<>}(R) =_{def} \{<>\} \text{ if } R \neq \emptyset; \ \ \emptyset \text{ otherwise} \tag{1}$$

The extensions of the intensional entities (concepts) are given by the set Ɛ of extension alization functions $h : \mathcal{D} \to D_{-1} + \mathfrak{R}$, such that

$$h = h_{-1} + h_0 + \sum_{i\geq 1} h_i : \sum_{i\geq -1} D_i \longrightarrow D_{-1} + \{f,t\} + \sum_{i\geq 1} \mathcal{P}(D^i) \tag{2}$$

Where $h_{-1} : D_{-1} \to D_{-1}$, for the particulars while $h_0 : D_0 \to \{f,t\} = \mathcal{P}(\mathcal{D}^0)$ as signs the truth values in *{f,t}* to all propositions with the constant assignment $h_0(Truth)$ = t = {< >} , and for each $i \geq 1$, $h_i : D_i \to \mathcal{P}(D^i)$ assigns a relation to each concept.

Consequently, intensions can be seen as names (labels) of atomic or composite concepts, while the extensions correspond to various rules that these concepts play in different worlds.

The intensional entities for the same logic formula, for example $x_2 + 3 = x_1^2 - 4$, which can be denoted by $\phi(x_2, x_1)$ or $\phi(x_1, x_2)$, from above we need to differentiate their concepts by $I(\phi(x_2,x_1)) \neq I(\phi(x_1,x_2))$ because otherwise we would obtain erroneously that $h(I(\phi(x_2, x_1))) = h(I(\phi(x_1, x_2)))$. Thus, in intensional logic the ordering in the tuple of variables x in a given open formula ϕ is very important, and explains why we introduced in FOL the virtual predicates in Definition 1.

Definition 5
Let us define the extensional relational algebra for the FOL by,

$$\mathcal{A}_{\mathfrak{R}} = (\mathfrak{R}, R_=, \{<>\}, \{\bowtie_S\}_{S\in\mathcal{P}(\mathbb{N}^2)}, \sim, \{\pi_{-n}\}_{n\in\mathbb{N}}),$$

Where $\{<>\} \in \mathfrak{R}$ is the algebraic value correspondent to the logic truth $R_=$ is the binary relation for extensionally equal elements, with the following operators:

1.Binary operator $\bowtie_S: \mathfrak{R}\times\mathfrak{R} \to \mathfrak{R}$, such that for any two relations $R_1, R_2 \in \mathfrak{R}$, the is equal to the relation obtained by natural join of these two relations if S is a non-empty set of pairs of joined columns of respective relations (where the first argument is the column index of the relation R_1 while the second argument is the column index of the joined column of the relation R_2); otherwise it is equal to the cartesian product $R_1 \times R_2$.
2.Unary operator $\sim$: $\mathfrak{R} \to \mathfrak{R}$, such that for any k-ary (with k ≥ 1) relation $R \in \mathcal{P}(\mathcal{D}^k) \subset \mathfrak{R}$ we have that $\sim(R) = \mathcal{D}^k \backslash R \in \mathcal{P}(\mathcal{D}^k)$, *where* '\' is the substraction of relations. For $u \in \{f,t\} = \mathcal{P}(\mathcal{D}^0) \subseteq \mathfrak{R}$, $\sim(u) = \mathcal{D}^0 \backslash u$.
3.Unary operator $\pi_{-n} : \mathfrak{R} \to \mathfrak{R}$, such that for any k-ary (with k ≥ 1) relation $R \in \mathcal{P}(\mathcal{D}^k) \subset \mathfrak{R}$ we have that $\pi_{-n}(R)$ is equal to the relation obtained by elimination of the n-th column of the relation R if 1 ≤ n ≤ k and k ≥ 2; equal to, from (1), $f_{<>}(R)$ if n = k = 1; otherwise it is equal to R.
We will use the symbol '=' for the extensional identity for relations in $\mathfrak{R}$.

The intensional semantics of the logic language with the set of formulae *L* can be represented by the mapping

$$\mathcal{L} \longrightarrow_I \mathcal{D} \Longrightarrow_{h\in\mathcal{E}} \mathfrak{R},$$

Where $\longrightarrow_I$ is a fixed intensional interpretation $I : \mathcal{L} \to \mathcal{D}$ with image $im(I) \subset \mathcal{D}$, and $\Longrightarrow_{h\in\mathcal{E}}$ is the set of all extensionalization functions $h : im(I) \to D_{-1} + \mathfrak{R}$ in $\mathcal{E}$

So, we can define only the minimal intensional algebra (with minimal number of operators) Aint of concepts, able to support the homomorphic extension $h : \mathcal{A}_{int} \to \mathcal{A}_{\mathfrak{R}}$ of the extensionalization function $h : \mathcal{D} \to D_{-1} + \mathfrak{R}$.

Definition 6
Basic Intensional FOL Algebra
Intensional FOL algebra is a structure

$$\mathcal{A}_{int} = (\mathcal{D}, Id, Truth, \{conj_S\}_{S\in\mathcal{P}(\mathbb{N}^2)}, neg, \{exists_n\}_{n\in\mathbb{N}}),$$

with binary operations $conj_S : D_I \times D_I \to D_I$,

$neg : D_I \to D_I$, unary operation $exists_n : D_I \to D_I$, and unary operations , such that for any extensionalization function $h \in \mathcal{E}$, and $u \in D_k, v \in D_j, k, j \geq 0$,

1. $h(Id) = R_{=}$ *and* $h(Truth) = \{<>\}$, *for* $Id = I(\doteq (x, y))$ *and* $Truth = I(\top)$.
2. $h(conj_S(u, v)) = h(u) \bowtie_S h(v)$, where $\bowtie_S$ is the natural join operation and $conj_S(u, v) \in D_m$ where $m = k + j - |S|$if for any every pair $(i_1, i_2) \in S$ it holds that
 $1 \leq i_1 \leq k$, $1 \leq i_2 \leq j$ *(otherwise* $conj_S(u, v) \in D_{k+j}$*).*
3. $h(neg(u)) = \sim (h(u)) = \mathcal{D}^k \backslash (h(u))$ (the complement of k-ary relation *h* (*u*)in D^k), if k ≥ 1, where neg (u) ∈ D_k. For $u_0 \in D_0$, $h(neg(u_0)) = \sim (h(u_0)) = \mathcal{D}^0 \backslash (h(u_0))$.
4. $h(exists_n(u)) = \pi_{-n}(h(u))$, where π_{-n} is the projection operation which eliminates n-th column of a relation and existsn(u) D_{k-1} if 1≤ n ≤ k (otherwise $exists_n$ is the identity function).

Notice that for

$u, v \in D_0$, so that $h(u), h(v) \in \{f, t\}$,
$h(neg(u)) = \mathcal{D}^0 \backslash (h(u)) = \{<>\} \backslash (h(u)) \in \{f, t\}$, and
$h(conj_\emptyset(u, v)) = h(u) \bowtie_\emptyset h(v) \in \{f, t\}$.

We define a derived operation $(\mathcal{P}(D_i) \backslash \emptyset) \to D_i,\ i \geq 0$, such that, for any
$B = \{u_1, ..., u_n\} \in \mathcal{P}(D_i)$ and $S = \{(l, l) \mid 1 \leq l \leq i\}$we have that

$$union(\{u_1, ..., u_n\}) = \begin{cases} u_1, & \text{if } n = 1 \\ neg(conj_S(neg(u_1), conj_S(neg(u_2), ..., neg(u_n))...), & \text{otherwise} \end{cases} \quad (3)$$

Than we obtain that for n ≥ 2:

$$h(union(B)) = h(neg(conj_S(neg(u_1), conj_S(neg(u_2), ..., neg(u_n))...)$$
$$= \mathcal{D}^i \backslash ((\mathcal{D}^i \backslash h(u_1)) \bowtie_S ... \bowtie_S (\mathcal{D}^i \backslash h(u_n))) = \mathcal{D}^i \backslash ((\mathcal{D}^i \backslash h(u_1)) \bigcap ... \bigcap (\mathcal{D}^i \backslash h(u_n)))$$
$$= \bigcup \{h(u_j) \mid 1 \leq j \leq n\}, \text{ that is,}$$

$$h(union(B)) = \bigcup \{h(u) \mid u \in B\} \quad (4)$$

Note that it is valid also for the propositions in $u_1, u_2 \in D_0$, so that ***h(union(****u*₁,*u*₂**))** = $h(u_1) \bigcup h(n_2) \in \{f, t\}$ where f is empty set ∅ while *t* is a singleton set {< >} with empty tuple < >, and hence the join $\{<>\} \bowtie \emptyset = \emptyset$ and $\{<>\} \bowtie \{<>\} = \{<>\}$.

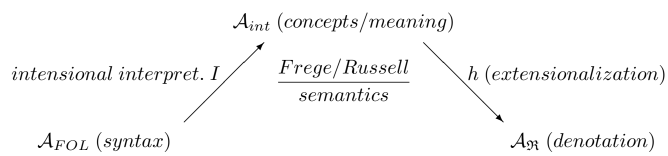
Thus, we define the following homomorphic extension

$$I : \mathcal{A}_{FOL} \to \mathcal{A}_{int}$$

of the intensional interpretation $I : \mathcal{L} \to \mathcal{D}$ for the formulae in syntax algebra A_{FOL} from Definition 2:

1. The logic formula $\phi(x_i, x_j, x_k, x_l, x_m) \wedge_S \psi(x_l, y_i, x_j, y_j)$ will be intensionally interpreted by the concept $u_1 \in D_7$, obtained by the algebraic expression $conj_S(u, v)$ where $u = I(\phi(x_i, x_j, x_k, x_l, x_m)) \in D_5, v = I(\psi(x_l, y_i, x_j, y_j)) \in D_4$ are the concepts of the virtual predicates ϕ, ψ, relatively, and S={(4,1), (2,3)}. Consequently, we have that for any two formulae ϕ, ψ □ L and a particular operator conjS uniquely determined by tuples of free variables in these two formulae,$I(\phi \wedge_S \psi) = conj_S(I(\phi), I(\psi))$.
2. The logic formula $\neg\phi(x_i, x_j, x_k, x_l, x_m)$ will be intensionally interpreted by the concept $u_1 \in D_5$, obtained by the algebraic expression $neg(u)$ where *u* is the concept of virtual predicate ϕ, $u = I(\phi(x_i, x_j, x_k, x_l, x_m)) \in D_5$. Consequently, we have that for any formulae $\phi \in \mathcal{L}$, $I(\neg\phi) = neg(I(\phi))$.
3. The logic formula $(\exists_3)\phi(x_i, x_j, x_k, x_l, x_m)$ will be intensionally interpreted by the concept $u_1 \in D_4$, obtained by the algebraic expression $exists_3(u)$ where $u = I(\phi(x_i, x_j, x_k, x_l, x_m)) \in D_5$ is the concept of virtual predicate ϕ. Consequently, we have that for any formulae $\phi \in \mathcal{L}$ and a particular operator $exists_n$ uniquely determined by the position of the existentially quantified variable in the tuple of free variables in ϕ (otherwise *n* = 0 if this quantified variable is not a free variable in ϕ), $I((\exists_n)\phi) = exists_n(I(\phi))$.

So, we obtain the following two-steps interpretation of FOL based on two homomorphism's, intentional *I*, and extensional *h*:



(5)

We can enrich the expressivity of such a minimal FOL intentionality by new modal operators, or in different way provided in what follows. As, for example, in Bealer's intentional FOL, where he introduced the intentional abstraction operator, which will be considered in rest of this section, as a significant enrichment of the intentional FOL considered above.

In reflective languages, reification data is causally connected to the related reified aspect such that a modification to one of them affects the other. Therefore, the reification data is always a faithful representation of the related reified aspect. *Reification data* is often said to be made a *first -class object.* In programming language design, a first-class citizen (also type, object, entity, or value) in a given programming language is an entity, which supports all the operations generally available to other entities. These operations typically include being passed as an argument, returned from a function, modified, and assigned to a variable. The concept of first and second-class objects was introduced by Christopher Strachey in the 1960s when he contrasted real numbers (first-class) and procedures (second-class) in ALGOL.

In FOL, we have the variables as arguments inside the predicates, and terms, which can be assigned to variables, are first-class objects while the predicates are the second-class objects. When we transform a virtual predicate into a term, by using intentional abstraction operator, we transform a logic formula into the first-class object to be used inside other predicates as first-class objects. Thus, abstracted terms in the intentional FOL are just such abstracted terms as reification of logic formulae. For example, the sentence "Marco thinks *that Zoran runs*", expressed by $(Marco, \lessdot runs(Zoran) \gtrdot)$ by using binary predicate *thinks* and unary predicate *runs* where the ground atom *runs* (*Zoran*) is reified into the predicate *thinks*

If ϕ(x) is a formula (virtual predicate) with a list (a tuple) of free variables in x = (x_1, ..., x_n) (with ordering from-left-to-right of their appearance in ϕ), and α is its subset of distinct variables, then $\lessdot \phi(\mathbf{x}) \gtrdot_{\alpha}^{\beta}$ is a term, where β is the remaining set of free variables in x. The externally quantifiable variables are the free variables not in α. When n = 0,$\lessdot \phi \gtrdot$ is a term which denotes a proposition, for n ≥ 1 it denotes a n-ary concept.

Definition 7
Intentional Abstraction Convention
From the fact that we can use any permutation of the variables in a given virtual predicate, we introduce the convention that

$$\lessdot \phi(\boldsymbol{x}) \gtrdot_{\alpha}^{\beta} \quad is\ a\ term\ obtained\ from\ virtual\ predicate\ \ \phi(\boldsymbol{x}) \qquad (6)$$

if α is not empty such that $\alpha \bigcup \beta$ is the set of all variables in the list (tuple of variables) $x = (x_1, \ldots\ldots x_n)$ of the virtual predicate (an open logic formula) ϕ, *and* $\alpha \bigcap \beta = \emptyset$, so that $|\alpha| + |\beta| = |\boldsymbol{x}| = n$. Only the variables in β (which are the only free variables of this term) can be quantified. If β is empty then $\lessdot \phi(\boldsymbol{x}) \gtrdot_{\alpha}$ is a ground term. If ϕ is a sentence and hence both α and β are empty, we write simply $\lessdot \phi \gtrdot$ for this ground term. More about this general definition of abstract terms can be find in [1]. In this paper, we will use the simplest cases of ground terms $\lessdot \phi \gtrdot$, where ϕ is a sentence.

1.2. Four-levels Robot's Brain Structure

Let us consider a model of robot for understanding language about space and movement in realistic situations, as finding video clips that match a spatial language description such as "People walking through the kitchen and then going to the dining room" and following natural language commands such as "Go down the hall towards the fireplace in the living room [11,12]".

Video retrieval is a compelling application: in the United States alone, there are an estimated 35 million surveillance cameras installed, which record four billion hours of video per week. Analyzing and understanding the content of video data remains a challenging problem. A spatial language interface to video data can help people naturally and flexibly find what they are looking for in video collections. Studying language used to give directions could enable a robot to understand natural language directions. People talk to robots even if they do not have microphones installed, and it makes sense to build systems that understand what they say. A robot that understands natural language is easy for anyone to use without special training. By using the deductive properties of the IFOL, the robot can make logic deductions as well about the facts that it visually recognized and to obtain its own auto epistemic deductions about obtained knowledge, as shortly explained in introduction, by using intentional abstractions in Definition 7.

Consequently, I will focus on a narrow subset of a natural language, grounding that language in data collected from a real world. This strategy has two benefits. First, it decreases the scope of the language understanding problem, making it more tractable. Second, by choosing a semantically deep core domain, it offers an opportunity to explore the connection between linguistic and non-linguistic concepts.

The linguistic structure extracted from spatial language expressions and many of the features in the model for spatial relations are based on the theories of Jackendoff, Landau and Jackendoff and Talmy [6,13,14]. For example, the implementation of the mining of "across" in is obtained by an algorithm (of robot's AI neuro-system) for computing the axes a figure imposes on a ground, and set of features which quantify "roughly perpendicular", using a machine learning algorithm to fine-tune the distinctions by training on labeled data [14]. Regier built a system that assigns labels such as "through" to move showing a figure relative to a ground object [15]. Bailey developed a model for learning the meanings of verbs of manipulation such as "push" and "shove" [16]. Kelleher and Costello built models for the meanings of static spatial prepositions such as "in front of" and "above" [17]. Siskind created a system for defining meanings for words such as "up" and "down." The framework reasons about formal temporal relations between primitive force-dynamic properties such as "supports" and "touches" and uses changes in these properties to define meanings for verbs [18]. His framework focuses on word-level event recognition and features, etc...

Reasoning about movement and space is a fundamental competence of humans and many animals. Humans use spatial language to tell stories and give directions, abstracting away the details of a complex event into a few words such as "across the kitchen." A system that understands spatial language could be directly useful to people by finding video that matches spatial language descriptions, or giving natural language directions. We will consider a robot, which retrieves video clips that match a natural language description using a probabilistic graphical model that maps between natural language and paths in the environment [11].

In this particular environment, spatial relations are modeled as probabilistic distributions for recognizing words paired with scenes. The distributions are trained from labeled examples using a set of geometric features that capture the semantics of spatial prepositions. The distribution modeled is the probability of a particular spatial relation given a trajectory and an object in the environment. This distribution corresponds to the probability that a spatial relation such as "across" or "to" describes a particular trajectory and landmark. The input to the model is the geometry of the path and landmark object; the output is a probability that the spatial relation can be used to describe this scene. These distributions are trained using labeled path examples, and in robot's brain correspond to its AI neuro system. The system learns distributions for spatial relations, for example, by using a naive Bayes probabilistic model.

So, now we can focus to the integration of such robot's AI neuro system with its AI symbolic system based on three natural language cognitive levels: The *syntax* of a particular natural language (French, English, etc..) its *semantic logic structure* (trans- formation of parts of the language sentences into the logic predicates and definition of corresponding FOL formulae) and its corresponding *conceptual structure,* which differently from the semantic layer that represents only the logic's semantics, represents the composed meaning of FOL formulae.

In this example, we focus on spatial language search of people's motion trajectories, which are automatically extracted from video recorded by stationary overhead cameras. The system takes as input a natural language query, a database of surveillance video from a particular environment and the locations of non-moving objects in the environment. When the robot performs video retrieval by its AI neuro system, clips are returned in order according to the joint probability of the query and the clip. Thus, for each video clip in given database, this robot's neuro system computes the probability that considered clip satisfies a natural language query, parsed into logic FOL formula (second natural language semantic level) and consequently into intentional algebra A_{int} term with intentional concepts which labels are grounded by robot's neuro system processes (algorithms). Let *NL* be a given natural language. If we denote the set of finite nonempty lists of as given natural language words by NL_{list}, then this parsing can be represented by a partial mapping

$$pars : \mathcal{NL}_{list} \rightarrow \mathcal{L} \qquad (7)$$

where *L* is the set of logic formulae of intensional FOL.

We suppose that the concepts in the conceptual structure expressed by the intensional algebra $\mathcal{A}_{int}$ of atomic concepts $u \in \mathcal{D}$, and their corresponding logic atoms expressed by virtual predicates $\phi(x) \in \mathcal{L}$ of FOL are the part of innate robot's knowledge, such that for robot's innate and unique intentional interpretation $I : \mathcal{L} \rightarrow \mathcal{D}$, $u = I(\phi(\mathbf{x}))$. Moreover, we suppose that robot has a parser capability to transform the sentences of particular natural language into the formulae of FOL with innate set of the atoms expressed by virtual predicates.

In this example we consider the predicates of IFOL as the verbs (V) of natural language, as follows

$$Find(x_1, x_2, x_3, x_4)$$

where the time-variable x_1 (with values "in past", "in present", "in future") indicates the time of execution of this recognition-action, the variable x2 is used for the subject who executes this action (robot in this case), the variable x3 is used for the object given to be eventually recognized (in this case a video clip) and x4 for the statement (users query) that has to be satisfied by this object, and virtual predicate

$$Walk(x_1, x_2, x_3, x_4, x_5)$$

where the time-variable x_1 (with values "in past", "in present", "in future") indicates the time of execution of this action, variable x_2 for the figure (F) that moves ("person", "cat", etc..), x_3 for the initial position of walking figure (defined by the spatial relation (SR) "*from*", for example "from the table") , x_4 for the intermediate positions during movement of the figure (defined by (SR) "*through*", for example "through the corridor"), and x_5 for the final position of figure (defined by (SR) "*to*", for example "to the door").

The robot takes as input a natural language query, a database of surveillance video from a particular environment and the locations of non-moving objects in the environment. It parses the query into a semantic structure called a spatial description clause (SDC) [12]. An SDC consists of a figure (F), a verb (V), a spatial relation (SR), and a landmark (L). The system extracts SDCs automatically using a conditional random field chunker. Let us consider the example illustrated in Figure in of a natural language query $nq \in \mathcal{NL}_{list}$, defined by a sentence:

"The person walked *from* the couches in the room *to* the dining room table" which is composed by two SDC with the first one
• (F) = "the person"
• (V) = "walked"
• (SR) = "from"
• (L) = "the couches in the room"
and the second SDC,
• (SR) = "to"
• (L) = "the dining room table"

Remark: Note that all SDC components different from (V), are particulars in D−1 in PRP domain, provided by Definition 3. The sense (mining) of the components (F) and (L) are grounded by the machine-learning video-recognition processes of the robot that is by its neuro systems. The sense of the (SR) components is grounded by the meaning of the spatial relations, provided by different author's methods, mentioned previously, and implemented by particular robots processes.

What we need in next is to extend this grounding also to the virtual predicates of the FOL open formulae in *L*.

Consequently, from these Spatial Description clauses, for the (V) of the past-time verb (V) "to walk", the semantic logic structure recognized by robot is the sentence $\phi = pars(nq) \in \mathcal{L}$, obtained from (7) so that, based on the virtual predicate to Walk, the sentence ϕ is

Walk (in past, person, from the couches in the room, NULL, to the dining room table) (8)

Note that the inverse parsing of such logic sentence ϕ to natural language sentence is directly obtained, so that the robot can translate its semantic logic structures into natural language to communicate by voice to the people.

We consider that each grammatically *plural* word name "videoclips", robot can de- fine by generalization by creating the virtual unary predicate *videoclips(y)*, such that its intentional *concept* $u_2 = I(videoclips(y)) \in D_1$ in PRP domain, whose meaning is grounded by robots pattern-recognition process fixed by a machine learning method. In a similar way, each unary concept of visual objects can be created by robot by a machine learning method for enough big set of this type of objects.

So, each grammatically *singular* word name, like "John's video clip" is a particular (element of D_{-1}) in PRP domain, whose meaning is grounded by the internal robot's image of this particular video clip, *recognized* as such by robots pattern-recognition process. Thus, for a given extensionalization function h in (2), and fixed robot's intentional mapping I, from the diagram (5), we obtain that the set C, of video clips in a given database of videoclips presented to this robot, is equal to

$$C = h(I(videoclips(y))) \qquad (9)$$

Consequently, the human command in natural language $nc \in \mathcal{NL}_{list}$ to this robot,
"Find videoclip such that ϕ in the given set of videoclips"

(where ϕ has to be substituted by the sentence above) is parsed by robot into its second level (semantic logic structure) by virtual predicate Find of the verb "to find" (*in present*) and a variable y of type "videoclip" (objects of research) and substituting "that ϕ" by abstracted term $\lessdot\phi\gtrdot$, and by substituting "in the given set of" with the logic conjunction connective $\wedge_S$ of the IFOL expressed, from (7), by the following formula *ψ(y) = pars(nc)*

$$Find(inpresent, me, y, \lessdot\phi\gtrdot) \wedge_S videoclips(y) \qquad (10)$$

where S = (2, 1) for joined variables in two virtual predicates.

The meaning of the unary concept $u_1 = I(Find(in\ present, me, y, \lessdot\phi\gtrdot))$, corresponding to the natural language subexpression "Find (me),videoclip such that ϕ" of the command above, is represented by its AI neuro system process of probabilistic recognition of video clips satisfying the natural language query ϕ (In fact, u_2 is just equal to the name of this process of probabilistic recognition) [12].

However, during execution of this process, the robot is able also to deduce the truth of the autoepistemic sentence, for a given assignment of variables $g : \mathcal{V} \to \mathcal{D}$, with

$$g(x_1) = in\ present \text{ and } g(x_2) = me,$$

$$Know(x_1, x_2, \lessdot Find(inpresent, me, y, \lessdot\phi\gtrdot)\gtrdot_y)/g \qquad (11)$$

of the virtual predicate Know(x_1, x_2, x_3), where the time-variable x_1 (with values "in past", "in present", "in future") indicates the time of execution of this action, the variable x_2 is used for the subject of this knowledge and x3 is used for an abstracted term expression this particular knowledge). Thus, by using deductive properties of the true sentences of FOL, this autoepistemic sentence about its state of self-knowledge, the robot would be able to communicate to humans this sentence, traduces in natural language as "I (me) know that I am (me) finding videoclip such that ϕ"

From the fact that robot defined the type of the variable y to be "videoclip", by traduction of the FOL deduced formula above into the natural language, this variable will be traduced in natural language by "videoclip". In the same way, during the execution of the human command above, expressed by the FOL formula ψ(y) in (10), with composed concept $u_3 = I(\psi(y)) \in D_1$, that is, by using the homomorphic property of intentional interpretation I,

$$u_3 = u_1 \bowtie_S u_2 \qquad (12)$$

the robot can deduce also the true epistemic sentence, for a given assignment of variables $g : V \to D$, with $g(x_1)$ = in present and $g(x_2)$ = me,

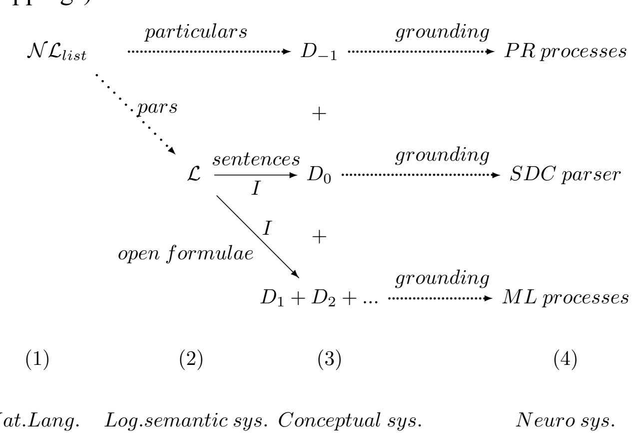
$$Know(x_1, x_2, \lessdot Find(in\ present, me, y, \lessdot\phi\gtrdot) \wedge_S videoclips(y)\gtrdot_y)/g \qquad (13)$$

and hence the robot would be able to communicate to humans this sentence, traduces in natural language as "I (me) know that I am (me) finding *videoclip* such that ϕ in the set of videoclips"

Note that the subset of videoclips extracted by robot from a given set of videoclips $C = h(u_2)$ in (9), defines the current extensionalization function h, in the way that this
subset is

$$E = h(u_3) = h(u_1) \bowtie_S h(u_2) = h(u_1) \bowtie_S C = h(u_1) \subseteq C \qquad (14)$$

Thus, for the grounding of spatial language for video search, the robot's internal knowledge structure is divided into four levels, in ordering: natural language, semantic logic structure, conceptual structure and neuro structure, as represented by the following diagram (only two continuous arrows (intentional mapping $I : \mathcal{L} \to D_I$ where $D_I = D_0 + D_1 + \ldots$ are the universals in PRP domain theory) represent the total mappings, while other (dots) are partial mappings)



(15)

It is easy to see that the conceptual system, based on PRP domain *D* composed by particulars in D_{-1} and universals (concepts) in DI = $D_0 + D_1 + D_2 + \ldots$ of the IFOL, is the level of grounding of the natural language of the robot to its neuro system composed by the following processes:

1. PR (Pattern Recognition) processes of recognition of the particulars. For example, for SDC components (F) “the person”, (L) “the couches in the room” and “the dining room table”, etc.
2. SDC (Spatial Description Clauses) parser used for the sentences, for example, for a natural language query $nq \in \mathcal{NL}_{list}$ that is, logical proposition (sentence) $\phi = pars(nq) \in \mathcal{L}$ in (8), which is labeled by its intentional proposition label $I(\phi) \in D0$. Thus, the grounding of nq is obtained by linking its intentional proposition *I(pars(nq))* in PRP to the SDC parser process (part of robot’s neuro system).
3. ML (Machine Learning) processes, like that used for the recognition of different types of classes (like the set of video clips). For example, for the language plural world “video clips” in $\mathcal{NL}_{list}$, such that $pars(\text{"}videoclips\text{"}) = videoclips(y) \in \mathcal{L}$ with its intentional unary concept $u_2 = I(videoclips(y)) \in D_1$ which is grounded to robot’s ML process for the “videoclips”.

Note that, while the top line in the diagram (15) is the ordinary component of the natural language grounding developed by robot’s neuro system, the two lines bellow is the new robots knowledge structure of the added *symbolic AI system* based on the Intensional First Order Logic and its grounding to robot’s processes (its neuro AI system), by which the robot is able to provide logic deductive operations and autoepistemic self- reasoning about its current knowledge states and communicate it to humans by using natural languages.

2. Conclusion

Computation is defined purely formally or syntactically, whereas minds have actual mental or semantic contents, and we cannot get from syntactical to the semantic just by having the syntactical operations and nothing else Machine learning is a sub-field of artificial intelligence. Classical (non-deep) machine learning models require more human intervention to segment data into categories (i.e. through feature learning). Deep learning is also a sub-field of machine learning, which attempts to imitate the interconnectedness of the human brain using neural networks. Its artificial neural networks are made up of layers of models, which identify patterns within a given dataset. Deep learning can handle complex problems well, like speech recognition, pattern recognition, image recognition, contextual recommendations, fact checking, etc.

However, with this integrated four-level robot’s knowledge system presented in diagram (15), where the last level represents the robot’s neuro system containing the deep learning as well, we obtain that also the semantic theory of robot’s intentional FOL is a procedural one, according to which sense is an abstract, pre-linguistic procedure de- tailing what operations to apply to what procedural constituents to arrive at the product (if any) of the procedure.

References

1. Majkic, Z. (2022). Intensional First-Order Logic: From AI to New SQL Big Data. Walter de Gruyter GmbH & Co KG.
2. Bealer, G. (1982). “Quality and concept,” Oxford University Press.
3. Majkić, Z. (2013). Conservative intensional extension of Tarski's semantics. Advances in Artificial Intelligence, 2013, 3-3.
4. Majkic, Z. (2011). Intensionality and two-steps interpretations. arXiv preprint arXiv:1103.0967.
5. Majkic, Z. (2014). Intensional RDB manifesto: a unifying New SQL model for flexible Big Data. arXiv preprint arXiv:1403.0017.
6. Jackendoff, R. S. (1985). Semantics and cognition (Vol. 8). MIT press.
7. Kahneman, D. (2011). Thinking, Fast and Slow, Farrar, Straus and Giroux, ISBN 978-0374275631. Reviewed by Freeman Dyson in New York Review of Books, 40-44.
8. Frege, G. (1892). Uber sinn und bedeutung. Zeitschrift für Philosophie und philosophische Kritik, 100(1), 25-50.
9. Bealer, G. (1993). Universals. The Journal of Philosophy, 90(1), 5-32.
10. Bealer, G. (1979). Theories of properties, relations, and propositions. The Journal of Philosophy, 76(11), 634-648.
11. Kollar, T., Tellex, S., Roy, D., & Roy, N. (2010, March). Toward understanding natural language directions. In 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 259-266). IEEE.
12. Tellex, S., Kollar, T., Shaw, G., Roy, N., & Roy, D. (2010, November). Grounding spatial language for video search. In International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (pp. 1-8).
13. Landau, B., & Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition?. Behavioral and brain sciences, 16(2), 255-265.
14. Talmy, L. (2005). The fundamental system of spatial schemas in language. From perception to meaning: Image schemas in cognitive linguistics, 29, 199-234.
15. Regier, T. P. (1992). The acquisition of lexical semantics for spatial terms: A connectionist model of perceptual categorization. University of California, Berkeley.
16. Bailey, D. R. (1997). When push comes to shove: A computational model of the role of motor control in the acquisition of action verbs. University of California, Berkeley.
17. Kelleher, J. D., & Costello, F. J. (2009). Applying computational models of spatial prepositions to visually situated dialog. Computational Linguistics, 35(2), 271-306.
18. Siskind, J. M. (2001). Grounding the lexical semantics of verbs in visual perception using force dynamics and event logic. Journal of artificial intelligence research, 15, 31-90.