Logical Characterisation of Concept Transformations from Human into Machine Relying on Predicate Logic

Farshad Badie
Center for Linguistics, Aalborg University,
Aalborg, Denmark. Email: badie@id.aau.dk

Abstract— **Providing more human-like concept learning in machines has always been one of the most significant goals of machine learning paradigms and of human-machine interaction techniques. This article attempts to provide a logical specification of conceptual mappings from humans' minds into machines' knowledge bases. We will focus on the representation of the mappings (transformations) relying on First-Order Predicate Logic. Additionally, the structure of concepts in the common ground between humans and machines will be analysed. It seems quite necessary to pay attention to the philosophy of constructivism and constructivist models of knowing. This research constructs a conceptual ground for expressing and analysing concepts in the common ground between humanistic and informatics sciences and in the context of human-machine interplays.**

Keywords: HCI; Concept; Concept Transformation; Predicate; Hypothesis; Predicate Logic; Machine Learning.

I. Introduction and motivation

In an interaction between human beings (as intentional, aware and intelligent agents) and machines (as unaware and artificial agents), they exchange multiple actions and transactions concerning, e.g., identifications, descriptions, specifications and reasonings. According to [1] and based on our epistemological approach, the multilevel interactions between a trainer (a human being) and an artificial and a metaphorical learner (a machine), could be seen as a radical constructivist account of human cognition and comprehension. Also, these interactions could shape a kind of ontology. Obviously, the human-machine interactions are not agreement-oriented, because an aware agent cannot make an agreement with an unaware agent, but we suppose there is a type of agreement and convention between the human being and herself/himself to forward information about a given domain to the machine and to train the machine about some particular topics and concepts in that domain. In the next section, we will focus on the expression 'concept'. In interactions between human beings and machines, humans can develop their non-evidential and non-axiomatical conceptions of the specified underlying systematic processes in the world.

Training machines based upon personal mental images of reality in the context of human-machine interactions, could provide a proper ground for constructivist machine training. At this point, we take the philosophy of *constructivism* into consideration. Constructivism appears in a variety of guises (e.g., pedagogical, epistemological and complex combinatorial). It has been known as a philosophical theory of learning and as a model of knowing, see [2]-[5]. According to constructivism, a human being is always concerned with the active creation of personal mental representations. As for learning in the framework of constructivism, any agent generates her/his own schemata, see [13]. Relying on our approach, any schema is the product of the trainer's understanding of the world. It conceptually represents the constituents of the trainer's thought about training something. Schemata support the trainer in constructing and in developing her/his concepts (that have been constructed with regard to her/his own realisation of the world). Additionally, they provide strong backbones for the trainer's interpretations and provide proper backgrounds for describing terminologies and world descriptions. The constructivist machine training framework is heuristic, explanatory and developmental for human being's thoughts and reasoning. Actually, any constructivist machine training in the context of human-machine interaction is concerned with heuristic questions focusing on (*i*) 'What/Which is …?', (*ii*) 'How is …?', and (*iii*) 'Why is …?'. The first group of questions focus on the factual, structural, existential and ontological aspects of the world, the second group focus on procedural, methodological and technical aspects of the world and the third group focus on inferential aspects of the world.

This article attempts to construct a conceptual and logical linkage between human's knowledge and machine learning. So, before getting into the details we contemplate the term 'Machine Learning'. Machine Learning is a subfield of Artificial Intelligence and Computer Science. According to [6], a machine learning approach attempts to develop strong algorithms that allow machines to improve [the productivity of] their performances on a given goal [and on an objective function]. In machine learning, the word '*learning*' has been utilised as a *predicate* for the expression '*machine*'. 'Learning' as a binary predicate describes a role that is being performed by the machine. More specifically, machines' *concept learning* approaches try to provide appropriate logical descriptions and specifications for transformed concepts and their interrelationships after having been transformed concerning their relationships with reality. A characteristic feature of most concept learning approaches is the use of background knowledge (e.g., internal knowledge base, ontological description). This feature supports more complicated and specific learning scenarios, because not only a factual (e.g., terminological) description of given examples can be used by the machine, but also structurally rich knowledge representations are taken into account, see [6][7]. In concept learning with background knowledge, with regard to the given set of training examples and background knowledge a machine focuses on hypothesis generation. In this article, we will provide a logical specification of mental mappings from humans into machines. We will focus on representations of transformations from humans' conceptions into machines' knowledge bases relying on First Order [Predicate] Logic (FOL). The results will support figuring out and analysing the most significant components of the logical characterisation of concept transformations. In the second section, we will focus on Concepts and Transforming Concepts. The third section will deal with Concept Transformation Process consisting of Logic of Transformation and the Analysis of Transformation. Section four will summarises the conclusions.

II. CONCEPTS AND TRANSFORMING CONCEPTS

First, we shall stress the fact that the notion of *concept* is a very sensitive term that must be used with caution, but we assume the use of 'concept' to be comprehensible in this context and in the logical formalisms. In our opinion, a human being's specified realisation of the world finds its real significance with regard to her/his grasp of the various concepts. Concepts support thoughts. Thoughts are also highly dependent on a human's interpretations and realisations of whether a given thing/phenomenon is an instance of a [constructed] concept or not. According to [8], and based on our conceptual approach, a concept is a linkage (relationship) between humans' mental images of reality (for instance, "an image of the *Spring*") and her/his linguistic expressions and statements (for instance, "*Spring* is one of the four conventional temperate seasons, following winter and preceding summer"). Let me represent the described linkage by —*R*—. In descriptive logical approaches, these expressions support the definitions. A *definition* is a kind of equivalence between a term referring to a thing (the thing that is going to be defined) and a description (generally built up using the inductive rules). Also, there is a strong relationship between the mental images and the mental representations of different aspects of the world. In fact, human beings need to logically apply —*R*— in their world descriptions, e.g., in assertions about real-world objects, in assertions about the empirical world, in assertional knowledge representations, in assertions about the ontologies, and in descriptions of terminologies and terminological knowledge. Therefore, human being transform —*R*— into discrete classes of things in order to see its applications. Thus, transformations play a very efficient part in the use of reasons and languages. Actually, transformations allow human to divide a continuously varying world into discrete classes of things, see [9].

At this point we focus on the concept formation process (see [10]) and acknowledge this process as the most fundamental step towards constructivist machine training. By forming concepts, a trainer (who is a human being) sorts her/his specific experiences and empirical studies into general classes [or even rules]. For instance, regarding the fact 'Drinking is a sign of thirst', s(he) represents the classes Drinking and Thirst and the rule 'Drinking → Thirst' in the machine's knowledge base. Consequently, the machine expresses the proposed classes and generates the proposed rule over the background knowledge in machine's knowledge base and with regard to other experiences of the trainer. Moreover, the machine utilises the expressed classes and the generated rules in class-based and rule-based reasoning processes. We have introduced the notion concept construction process in [11][12] and have interpreted it as the super-category of concept formation processes. A concept construction process consists of 'forming concepts' and 'reforming constructed concepts'. The trainer is highly concerned with main characteristics and features of a thing/a phenomenon in order to consider it as an instance of a class. The trainer must employ the examples that can lead her/him to discovering new classes. S(he) searches for [and itemises] the attributes and properties that can be used to distinguish exemplars from non exemplars of various classes. Additionally, s(he) identifies, specifies and relates the generalised examples and compares different examples. The following statements are derived from the above mentioned characteristics of concepts.

The descriptive logical languages and logical techniques transform the relationships between a human's mental images and her/his linguistic expressions into various ideas that are representable in the form of entities (discrete classes of things). The ideas specify the human's definitions (that are supported by linguistic expressions) by employing the logical rules that are (could be) existing between the same classes in the world. Accordingly, an idea is transformed into an hypothesis in order to correspond to a discrete class. As for the fundamental characteristics of concepts, a human being's conception within her/his interactions with a machine is equivalent to her/his act of representing various concepts and linking her/his explanations, and, respectively, definitions, with regard to her/his own mental images.

III. CONCEPT TRANSFORMATION PROCESS

"As accounted from above, a concept is a relation, and in fact, a binary predicate between humans' mental images of the world and their linguistic expressions [and, thus, definitions]". Obviously, the definitions always attempt to provide appropriate descriptions for the mental images. Subsequently, the existing interrelationships and dependencies between mental images and the provided descriptions support idea generation. At this point we focus on the analysis of idea transformation from humans' minds into machines' knowledge bases. Suppose that the trainer has considered *n* objects. For instance, the set of *n* objects is equal to $\{sofa_1, glass_2, plate_3, \ldots, brown_n\}$, and we shall draw your attention to the logical description of the transformation process.

A. Logic of Transformation

[1] The trainer assigns her/his ideas to the objects and focuses on *idea assertion*. For instance, s(he) assigns her/his first idea to the first object. So s(he) constructs $Idea_1(object_1)$. For instance, s(he) constructs *Furniture*(*sofa*) to express the fact that sofa is a furniture (or sofa is a member of the class Furniture). Similarly, s(he) assigns the second [and, respectively, the third, fourth, … , and *n*th] ideas to the second [, third, fourth, … , and *n*th] objects. Therefore, there are totally *n* assignments like: $Idea_1(object_1)$, $Idea_2(object_2)$, … , $Idea_n(object_n)$. This conclusion represents a linear model. Considering $i \in [1,n]$ and relying on FOL, $Idea_i$ represents an unary predicate and $object_i$ represents a constant symbol (as an instance of the unary predicate $Idea_i$).

[2] The trainer makes a relation between her/his achievements. Employing FOL, there exists a $Relation[\ Idea_1(object_1)\ ,\ Idea_2(object_2)\ ,\ ...\ ,\ Idea_n(object_n)\ ]$. For instance, s(he) can relate the assertions (the world descriptions) *Furniture*(*sofa*) and *Colour*(*brown*) to each other. Then, *Relation*[*Furniture*(*sofa*) , *Colour*(*brown*)] is capable of representing different types of relationships between *sofa* and *brown* with regard to their labels in the trainer's mind. Actually, the proposed world descriptions can actively develop her/his knowledge. Also, the relationship between the world descriptions can establish various expressions in her/his mind. Let me conclude that these relationships construct more specified ideas based upon the proposed world descriptions. Relying on FOL and considering $p,\ q \in [1,n]$, $Relation[\ Idea_p(object_p),\ Idea_q(object_q)\ ]$ represents a binary predicate between two unary predicates (between $Idea_p$ and $Idea_q$). This relation is also valid between $object_p$ and $object_q$ as the instances of $Idea_p$ and $Idea_q$. In this step, the trainer has produced a linear relational model, see Figure 1.

[3] The approached linear relational model is based on FOL. But it could also be represented in the form of a *j*-by-*i* matrix like *I*, where $i, j \in [1,n]$. This step represents the most significant assumption of the transformation. We shall stress the fact that we have represented the linear description $Relation[\ Idea_1(object_1)\ ,\ Idea_2(object_2)\ ,\ \dots\ ,\ Idea_n(object_n)\ ]$ in the form of a *j*-by-*i* matrix in order to allow the required linear transformation (that reflects the ideas) to be represented in a well-structured format. Additionally, a matrix can appropriately be used in establishing a transformation. Here, we have a matrix (relational model), see Figure 2.

[4] This step focuses on reflection. The idea assertion $Idea_1(object_1)$ (located in the first row and the first column of the matrix *b*) gets reflected in $Predicate_1(constant_1)$ (located in the first row and the first column of matrix *c* that is the product of the transformation) and $Idea_n(object_n)$ (located in the *j*th row and the *i*th column of the matrix *b*) gets reflected in $Predicate_n(constant_n)$ (located in the *j*th row and the *i*th column of matrix *c*). Thus, all cells in the relational model *b* collectively are reflected in an equivalent relational model (matrix), see Figure 3.

[5] The relational model *c* represents a relationship between $Predicate_1(constant_1)$, $Predicate_2(constant_2)$, …, and $Predicate_n(constant_n)$. Therefore, we have a description like $Relation[\ Predicate_1(constant_1)\ ,\ \dots\ ,\ Predicate_n(constant_n)\ ]$. Consequently, there are *n* assignments from the [unary] $Predicate_1$ into $constant_1$, from $Predicate_2$ into $constant_2$, …, and finally from $Predicate_n$ into $constant_n$. These assignments have been related with each other by means of *n*-ary *Relation*. Based on FOL, the effect of *n*-ary *Relation* is equivalent to $Predicate[\ Predicate_1(constant_1)\ ,\ Predicate_2(constant_2)\ ,\ \dots\ ,\ Predicate_n(constant_n)\ ]$. Note that the outer predicate is *n*-ary and works on *n* internal unary predicates. Then, the trainer has produced a linear relational model, see Figure 4.

[6] This step focuses on generating the relational hypothesis model. Actually, the effect of the first unary predicate on the first constant symbol generates the first hypothesis (or $Hypothesis_1$), the effect of the second unary predicate on the second constant symbol generates the second hypothesis (or $Hypothesis_2$), …, and the effect of the *n*th unary predicate on the *n*th constant symbol generates the *n*th hypothesis (or $Hypothesis_n$). Subsequently the outer *n*-ary predicate relates $Hypothesis_1$, $Hypothesis_2$, …, $Hypothesis_n$. Therefore, there is a relationship between all generated hypotheses. Thus, we have $Predicate[\ Hypothesis_1\ ,\ Hypothesis_2\ ,\ \dots\ ,\ Hypothesis_n\ ]$. Therefore, we have a relational hypothesis model, see Figure 5.

[7] Finally, there is a set like $\{\ Hypothesis_1\ ,\ Hypothesis_2\ ,\ \dots\ ,\ Hypothesis_n\ \}$ that represents the generated hypotheses for the machine.

B. Analysis of Transformation

"Suppose that (*i*) I_n denotes the *n*-component linear relational model $[Idea_1(object_1)\ ,\ Idea_2(object_2)\ ,\ \dots\ ,\ Idea_n(object_n)]$, (*ii*) P_n denotes the *n*-component linear relational model $[Predicate_1(constant_1)\ ,\ Predicate_2(constant_2)\ ,\ \dots\ ,\ Predicate_n(constant_n)]$, and (*iii*) H_n denotes the *n*-component linear relational model $[Hypothesis_1\ ,\ Hypothesis_2\ ,\ \dots\ ,\ Hypothesis_n]$". First, we focus on the forward direction from human to machine. There are *reflection* functions like $\boldsymbol{R}_i$ from human being's ideas into predicates. Let me represent the set of $\boldsymbol{R}_i$ by $\boldsymbol{R}$. So, $\boldsymbol{R}: I_n \rightarrow P_n$. Then, $\boldsymbol{R}$ represents the transformed ideas into predicates. Semantically, the reflection functions $\boldsymbol{R}$ satisfy the *n*-component model $[Hypothesis_1\ ,\ Hypothesis_2\ ,\ \dots\ ,\ Hypothesis_n]$ (i.e., provide proper models that attempt to satisfy the hypotheses). Then, there is a model like $\boldsymbol{R} \vDash H_n$. Therefore, the reflection functions $\boldsymbol{R}$ semantically satisfy the set of hypotheses in the machine (Result 1). At this point, we focus on the backward direction from machine to human. There are various *conformation* functions like $\boldsymbol{C}$ such that $H_n \vDash \boldsymbol{C}$. Semantically, any conformation function gets satisfied by a hypothesis like $Hypothesis_i$ belonging to the *n*-component relational model $[Hypothesis_1\ ,\ Hypothesis_2\ ,\ \dots\ ,\ Hypothesis_n]$. Note that $\boldsymbol{C}$ denotes the set of $\boldsymbol{C}_i$. So, $\boldsymbol{C}$ represents the transformed predicates into ideas and formally, $\boldsymbol{C}: P_n \rightarrow I_n$ (Result 2). According to the results 1 and 2 we have:

$$(\boldsymbol{R}: I_n \rightarrow P_n) \models H_n \models (\boldsymbol{C}: P_n \rightarrow I_n).$$

$$\text{Then:} \quad I_n \rightarrow P_n \models H_n \models P_n \rightarrow I_n.$$

In fact, the reflection transformations from ideas into predicates satisfy the hypotheses. And the hypotheses satisfy the inverse reflection transformations (or conformation transformations) from predicates into ideas.

IV. CONCLUSIONS

Training machines based upon personal mental images of the world in the context of human-machine interactions shapes the skeleton of constructivist human-machine interactions. Schemata in constructivist training frameworks could demonstrate the trainer's realisations of the world. They conceptually represent the constituents of the trainer's thoughts for training concepts. Schemata support the trainer in developing her/his constructed concepts (that have been constructed with regard to her/his own realisation of the world). In this article we have provided a logical and epistemological specification of concepts and we have seen the linkages between human's mental images and her/his linguistic expressions as the origins of manifestation of concepts. Accordingly, we have logically specified the mental mappings from human into machine and we have focused on logical representations of transformations from human beings' conceptions into machines' knowledge bases relying on First-Order Predicate Logic. We have identified the transformations from humans' mind into machines' knowledge bases by 'reflection transformations' and we have labeled the inverse cases by 'conformation transformations' in order to analyse the proposed logical descriptions. The reflection transformations from ideas into predicates satisfy the hypotheses. And the hypotheses satisfy the conformation transformations from predicates into ideas. In future research, we will employ the results in formal semantic analysis of concept transformations from minds into knowledge bases and in specifying their conceptualisations.

REFERENCES

[1] G. McIntyre Boyd, "Conversation Theory". Handbook of Research on Educational Communications and Technology, 2004.

[2] T. Husen, and T. N. Postlethwaite, "Constructivism in Education". The International Encyclopaedia of Education, Supplement Vol.1. Oxford/New York: Pergamon Press, 1989, 162–163.

[3] R. J. Spiro, P. J. Feltovich, M. J. Jacobson, and R. L. Coulson, "Cognitive Flexibility, Constructivism, and Hypertext. Random Access Instruction for Advanced Knowledge Acquisition in Ill-structured Domains", Educational Technology, 1991, 24-33.

[4] P. McGawand Peterson, "Constructivism And Learning", International Encyclopaedia of Education, Oxford: Elsevier, 2007.

[5] D. C. Phillips, "The Good, the Bad, And the Ugly: The Many Faces of Constructivism", Educational Researcher 24 - No 7, 1995, 5–12.

[6] T. Mitchell, "Machine learning", in Machine Learning, Kluwer Academic Publishers, 1997.

[7] N. Lavrac and S. Dzeroski, "Inductive Logic Programming: Techniques and Applications". Artificial Intelligence. Ellis Horwood (Simon & Schuster), 1994.

[8] H. Götzsche, "Deviational Syntactic Structures", Bloomsbury Academic: London / New Delhi / New York / Sydney, 2013.

[9] B. M. Lake, "Towards More Human-Like Concept Learning in Machines: Compositionality, Causality, and Learning-To-Learn", Massachusetts Institute of Technology, 2014.

[10] W. Parker, "Concept Formation" in http://teachinghistory.org, [accessed March 2016].

[11] F. Badie, "A Semantic Basis for Meaning Construction in Constructivist Interactions", CELDA-15 Proceedings, Ireland, 2015, 369-373.

[12] F. Badie, "Towards A Semantics-Based Framework For Meaning Construction in Constructivist Interactions:, ICERI2015 Proceedings, Spain, 2015, 7995-8002.

[13] Schema in http://plato.stanford.edu/entries/schema/, [accessed March 2016].

$Idea_1(object_1)$	…	$Idea_n(object_n)$

Fig. 1: Linear Relational Model

$Idea_1(object_1)$	…	$Idea_i(object_i)$
…		
$Idea_j(object_j)$	…	$Idea_n(object_n)$

Fig. 2: Relational Model

$Predicate_1(constant_1)$	…	$Predicate_i(constant_i)$
…		
$Predicate_j(constant_j)$	…	$Predicate_n(constant_n)$

Fig. 3: Relational Model

$Predicate_1(constant_1)$	…	$Predicate_n(constant_n)$

Fig. 4: Linear Relational Model

$Hypothesis_1$	…	$Hypothesis_n$

Fig. 5: Linear Relational Model

 ISBN: 978-1-61208-468-8