The Simulator as a Learning Artificial Actor [LAA] Version 1

*

Gerd Doeben-Henisch
gerd@doeben-henisch
in cooperation with the HKVD-FFM Team

August 23, 2020

Abstract

The analysis of the main application scenario revealed that classical logical inference concepts are insufficient for the assistance of human actors during shared planning. It turned out that the simulator has to be understood as a real learning artificial actor which has to gain the required knowledge during the process.

1 Basic Application Scenario

In our project we assume as *basic application scenario*[1] human persons talking with each other in a shared everyday language L_0 with the goal, to collect all their individual knowledge which is related to an agreed question Q and write this shared knowledge down in two documents: (i) a description of an agreed actual static state S and (ii) a description of a set of change rules X, which shall be applied to the actual state S or to some possible successor state S'.

Thus we assume some part S of the real world RW with which the human experts – here called human actors A_{hs} – can interact, the human actors themselves, and texts written in the language L_0 functioning either as state descriptions D_S or as rule descriptions D_X. Written with more details we can state about an *application scenario* AS:

[1] See for this `https://www.uffmm.org/2020/07/26/komega-requirements-no-1-basic-application-scenario/`

$$
\begin{aligned}
AS(x) \quad & iff \quad x = \langle RW, CLCK, D_Q, S, A_{hs}, D_S, D_X, E_{L.0} \rangle && (1)\\
RW \quad & := \quad Real\ World && (2)\\
CLCK \quad & := \quad Clocks\ for\ time\ events && (3)\\
CLCK \quad & \subset \quad RW && (4)\\
clck \in CLCK \quad & : \quad \emptyset \longmapsto T && (5)\\
T \quad & := \quad Time\ marker && (6)\\
D_Q \quad & := \quad Text\ with\ questions\ for\ all\ experts && (7)\\
S \quad & := \quad Static\ state && (8)\\
S \quad & \subset \quad RW && (9)\\
A_{hs} \quad & := \quad Human\ actors && (10)\\
A_{hs} \quad & \subset \quad RW && (11)\\
D_S \quad & := \quad Text\ describing\ S && (12)\\
D_X \quad & := \quad Change\ rules\ describing\ changes\ of\ D_S && (13)\\
E_{L.0} \quad & := \quad Expressions\ of\ everyday\ language\ L_0 && (14)\\
E_{L.0} \quad & \subset \quad RW && (15)\\
D_Q \quad & \subseteq \quad E_{L.0} && (16)\\
D_S \quad & \subset \quad E_{L.0} && (17)\\
D_X \quad & \subset \quad E_{L.0} && (18)
\end{aligned}
$$

Thus the human actors A_{hs} are living in a real world RW and these human actors are either part of a static situation S at a certain point of time $t \in T$ or they *know* about such a state S or they can *imagine* such a state S in a way that the human actors are able to *speak about* this state S or to *write about* this state S in the format of a document D_S with expressions $E_{L.0}$ of the shared everyday language L_0.

While the expressions of the document as such have *no meaning*, there are *facts* F in the state S which shall be *encoded* by the expressions of the document $D_S \subset E_{L.0}$. This is only possible if there exists an explicit mapping between the (real/ known/ imagined) facts F of the (real/ known/ imagined) state S and the real expressions of the document D_S.

From human actors it is known that the brain of a human person not only collects continuously sens data from the environment *external* to the body as well as *internal* to the body. These collected data are all represented as *neural correlates* NN of external events called here collectively the *internal states* IS_{NN} of the system.[2]

[2] For a first introduction into this perspective see e.g. Baars and Gage (2010) [BM10] and Gage an Baars (2018) [GB18]

It is known that these internal states in the brain are organized in a complex way allowing e.g. a distinction between those neural correlates $E_{NN} \subset IS_{NN}$ which represent external expressions E_{RW} and those neural correlates $F_{NN} \subset IS_{NN}$ which represent all kinds of facts F_{RW} from a situation, including even expressions.[3] It is further known that there exists complex mappings inside the brain between these both sets in the format $\mu : E_{NN} \longleftrightarrow F_{NN}$ thereby enabling an *encoding of meaning* by the meaning function μ either from expressions into facts or from facts into expressions. We call the transformation from external events from the real world RW into internal states IS_{NN} general *perception* $perc$ with $perc : RW \longmapsto IS_{NN}$, and vice versa the transformation of internal neural expressions E_{NN} into spoken/ written expressions E_{RW} as either talk $tlk : E_{NN} \longmapsto E_{RW.tlk}$ or write $wrt : E_{NN} \longmapsto E_{RW.wrt}$.

Finally, because the real world is constantly changing – including the human actors – , one has to assume that the system is capable to adapt to these changes. In this text this is called *learning* λ written as: $\lambda : IS_{NN} \longmapsto IS_{NN}$. Formally this can look like a tautological structure, but because the perceptions and other operations change the internal states constantly this self-referential structure is in itself no tautology.

These assumptions give raise to a simple human actor HA model as follows:

[3] These neural representations of induced facts from outside the brain are *not 1-1*! The brain is transforming the incoming data in a manifold way into internal structures which serves primarily the goal to survive including many pre-fixed sub-goals. To understand the world *outside the brain* is a task on its own. For this the brain has to compute on its own structure to get hints how the world outside the drain probably is. There exists no simple and direct way to conclude from the inside of the brain to the outside!

$$
\begin{aligned}
HA(a) \quad & iff \quad a = \langle RW, IS_{NN}, F_{NN}, E_{NN}, perc, \lambda, \mu, cmp, tlk, wrt \rangle & (19)\\
RW \quad & := \quad Perceivable\ real\ world\ events & (20)\\
IS_{NN} \quad & := \quad Internal\ states & (21)\\
perc \quad & \quad : RW \longmapsto IS_{NN} & (22)\\
F_{NN} \quad & := \quad Neural\ correlates\ of\ external\ facts & (23)\\
F_{NN} \quad & \subset \quad IS_{NN} & (24)\\
E_{NN} \quad & := \quad Neural\ correlates\ of\ ext.\ expressions & (25)\\
E_{NN} \quad & \subset \quad IS_{NN} & (26)\\
E_{NN} \quad & \subset \quad F_{NN} & (27)\\
\lambda \quad & : \quad IS_{NN} \longmapsto IS_{NN} & (28)\\
\mu \quad & : \quad E_{NN} \longleftrightarrow F_{NN} & (29)\\
cmp \quad & : \quad IS_{NN} \times IS_{NN} \longmapsto [0,1] & (30)\\
cmp \quad & := \quad Comparing\ neural\ corelates & (31)\\
tlk \quad & : \quad E_{NN} \longmapsto E_{RW.tlk} & (32)\\
E_{RW.tlk} \quad & \subset \quad RW & (33)\\
wrt \quad & : \quad E_{NN} \longmapsto E_{RW.wrt} & (34)\\
E_{RW.wrt} \quad & \subset \quad RW & (35)
\end{aligned}
$$

This human actor model HA is simple because it makes only a small fraction of all the important aspects of a human actor a subject of discussion. This simple model can be understood as the minimal schema for a *semiotic agent* as discussed in the field of semiotics.[4] What is completely missing at this point are topics like *goal directed behavior*, *preferences*, *desires and emotions*, and much more.

2 Truth Theory

With the assumptions so far we have outlined a minimal situation with human actors writing down texts to describe certain states S and possible changes X of these states in a way that follow-up states can also be candidates for such change rules.

In this context it is important to clarify what it means that an expression is said to be *true*? Thus if some human actors are writing a state description D_S with a certain state S 'in their minds' then it should be possible, that these

[4]See for a good introduction Noeth (1990)[N9̈0]. A more recent edition can be found as a German edition (2000) [N0̈0]. See also Doeben-Henisch (1998) [DH]

human actors can state that the expressions E of D_S are *true statements*.

Defining True Expressions: From our HA-Model it follows that we have at least the following entities: (i) The expressions E_{RW} of the documents $D_{RW.S}$, (ii) the neural correlates of these expressions E_{NN} of $D_{NN.S}$, (iii) the meaning function μ, (iv) the corresponding meaning of these expressions as neural correlates $\mu(D_{NN.S}) = F_{NN.D.S}$, (v) the actual perception of the state S as $perc(S.RW) = IS_{NN.RW.S}$ with $F_{NN.S} \subseteq IS_{NN.RW.S}$.

A possible *definition* of a *true expression* with regard to some aspect of the real world could then be, that a human actor trained in the language L_0 can translate this expression into a neural construct which can with the learned meaning function μ be mapped into possible neural constructs representing facts $F_{NN.D.S}$ and *compare these* with those neural correlates $F_{NN.RW.S}$ gained through perception from the state S_{RW} in the real world, written as $cmp(F_{NN.D.S}, F_{NN.RW.S}) = x, x \in [0, 1]$. If there exists a *sufficient similar correspondence* $x > 0.95$ between the both sets $F_{NN.D.S}, F_{NN.RW.S}$ then one can say that the given expression from the document $D_{RW.S}$ is *true* with regard to the perceived situation S_{RW}, otherwise not.[5]

The Phenomenological Counterpart: In this text it is assumed that the *subjective-introspective* counterpart of such a neural mapping process – often called the *phenomenological* perspective – is the *subjective feeling* of an *intuitive evidence* that the learned meaning is matching an actual perception. This then is enables the feeling of being *convinced* that some intended fact is really what it should be.

Possible Risks: One sees easily that this definition has a lot of presuppositions which not automatically are given in any case. Thus two different actors can have learned until the time of the writing and reading of this expression *different meaning functions*, or the actual perceptions are different. Thus there must exist minimal *patterns of interaction* between human actors how they can *clarify* a possible *agreement* or not. We call such a pattern of interactions to clarify the possible agreement with regard to the 'truth of expressions' the *truth game*. Without such a truth game it is not possible to establish a social con-

[5] The value '0.95' used here as a kind of a 'threshold' for being 'similar enough' is completely speculative. To know more exact values one had to investigate these internal neural processes in more detail. With the actual available methods and tools in Neuroscience this seems to be still impossible.

vention called *true expression*.[6]

Until now it has been assumed that the state S is given as a real state S_{RW} which can actually been perceived. Very often or even in most cases this is not the case. There are also the cases (i) that we *remember* a state S_{mem} or (ii) we *imagine* a state S_{img}. What can we say about *true expressions* in these cases?

State Descriptions From Memory: If we write a state description $D_{RW.S}$ based on our *memories* of a state $S_{NN.mem}$ then – in principle – there can be other neural correlates $S_{NN.X}$ which are the intended meaning $\mu(D_{NN.S}) = F_{NN.S}$ of the written text and these neural correlates $F_{NN.S}$ can be *compared* with the remembered neural correlates $S_{NN.mem}$ of such a state. This is not too different from the perception case. This comparison can yield *sufficiently similar agreements* in a way which causes again this *subjective feeling* of evidence that the intended meaning of this text is in agreement with a state of the world S_{RW} which we can remember as $S_{NN.mem}$, or not.

To use such an *explanation* for the decision of the truth of an expression in case of an *absent* state S requires an additional assumption about the human actor model HA. This assumption deals with the distinction between (i) *actual perceptions* P_{NN}, (ii) *abstracted actual perceptions*$P_{NN.cat}$ and (iii) *stored abstracted perceptions* $P_{NN.cat.mem}$. In this text it is assumed that all actual perceptions are *individual, concrete neural structures* which can be seen as *possible instances* of abstracted actual perceptions $P_{NN.cat}$ which can be stored as $P_{NN.cat.mem}$. The difference between the abstracted perceptions $P_{NN.cat}$ and the stored abstracted perceptions $P_{NN.cat.mem}$ is that the stored abstracted perceptions can associate multiple additional abstracted perceptions $P_{NN.cat}$ in a network-like structured thereby enhancing the category with many properties. The abstracted perception $P_{NN.cat}$ functions in this context like an *interface* between different concrete perceptions P_{NN} and a large set of stored abstracted perceptions $P_{NN.cat.mem}$. A complete theory of the memory of the human actor is until today missing. In this text only such assumptions will be made which are necessary to give a minimal sufficient explanation of the observable behavior.

With these considerations we are getting the following (sketchy) extension of the HA model:

[6] At this point one has to include the large topic of *language games* induced by the later Ludwig Wittgenstein (1953) [Wit53] as well the concept of *speech acts* as introduced by J.L.Austin (1955)[Aus55] with beginnings in the years before 1955.

$$
\begin{aligned}
HA(a) &\quad iff \quad a = \langle ..., P_{NN}, P_{NN.cat}, P_{NN.cat.mem}, abstr, cncpt \rangle && (36)\\
... &\quad\quad ... \\
P_{NN} &\quad \subset \quad IS_{NN} && (37)\\
P_{NN.cat} &\quad \subset \quad IS_{NN} && (38)\\
P_{NN.cat.mem} &\quad \subset \quad IS_{NN} && (39)\\
abstr &\quad : \quad P_{NN} \times IS_{NN} \longmapsto P_{NN.cat} && (40)\\
cncpt &\quad : \quad P_{NN} \times P_{NN.cat} \times IS_{NN} \longmapsto P_{NN.cat.mem} && (41)
\end{aligned}
$$

As one can see in this case too things can go wrong if certain assumptions are not valid. The most prominent aspect is the *similarity of memories* between different human actors. This poses in this case too the necessity for a requirement to provide accepted conventions how different human actors can decide, whether their memories are sufficiently similar.

State Descriptions by Imagination: Finally there is the case of an *imagined state* $S_{NN.img}$. In this text we assume that an *imagined state* $S_{NN.img}$ is the result of a generative process where human actors take an *initial state* $S_{NN.img.0}$ as starting point and then show by which actions one can change the initial state in a way which finally leads to the imagined state $S_{NN.final}$. The initial state $S_{NN.0}$ has to be some *known* state or an *actual perceived* state and the applied changes X shall be changes which are known to be possible or they are candidates for to been shown to be possible. Such an *imagined state* $S_{NN.final}$ can then be classified as being *true* – or not – in the same way like the actual perceived state or the remembered state. But because the meaning of the imagined state can become quite complex it can last some time until the *verification* of the truth of such a description can be finalized. Imagined states $S_{NN.img}$ have also to be stored in the memory but they are not caused by direct perceptions but by *inner activities* of the system itself, in this case in the format of a *generative process*. On the *phenomenological* level this can be understood as *thinking ahead*. Thinking ahead can help to identify *possible states in a possible future* which perhaps can be more advantageous than others which can happen if we as human actors do not prevent the occurrence of the less advantageous states.

And, indeed, the case of state descriptions by imagination is the most interesting and most demanding case. And it is this case which is the central idea of the whole project of supporting human actors by sharing their knowledge to become able to imagine together possible advantageous future states for their life.

3 Generating Imagined States of the Future

How the brain is working while generating new imagined states which suffice all the above mentioned requirements is still not really known. What one can do and what will be done in this project is to define an application scenario by *defining the allowed observable behavior* and then asking back which kind of processes are necessary to enable the generation of imagined future states.

As described above the minimal setting consist of a group of human actors A_{hs} which take as triggering event some written question D_Q and try then to describe a first initial state S_0 in a document $D_{S.0}$. Then they describe in another document D_X a set of known or imagined change rules. Then they will apply the change rules again and again until they agree to stop the process.

In this text it is assumed that a *follow up state* S' of a *given state* S will be generated by either (i) *deleting* some expressions E^- of the given state S if converted to S' or (ii) by extending the state S by some new expressions E^+ for the construction of S' by*creating* new expressions, written as $S' = S - E^- \cup E^+$.

For this to happen it is assumed in this text that a *change rule* $\xi \in X$ has the following format: *IF*-part and *THEN*-part, written as $X \subseteq X_{if} \times X_{then}$. While an element of the then-part has always the format $X_{then} \subseteq E^- \times E^+$ an element of the if-part is a set of expressions which have to be *valid*. One can generalize this case in the following way: If we assume that the expressions of the state S are all *true* with regard to the assumed part of the world then it has to be clarified for each expression ϕ in the if-part of a change rule whether this expression is *true* in the assumed state S or not. The statement that the expression ϕ is *true* can be clarified in different ways. Based on

1. the *intuitive evidence* of a human actor, written as $S \models_{hs} \phi$
2. the *syntactical evidence* of formal logic, written as $S \vdash \phi$
3. the *semantic evidence* of formal logic, written as $S \models \phi$
4. the *computed evidence* of a Turing machine, written as $S \models_{tm} \phi$

3.1 Intuitive Evidence: $S \models_{hs} \phi$

Relying on the intuitive evidence of human actors is as good or as bad as this intuition can be explained. Because the real machinery of human thinking is still mostly hidden in the unconscious space of neural processing, we can only try to approximate these mechanisms, by introducing rules and/ or methods of measurements to enable some transparency between us. But as we know these cultural techniques are very unstable, often fuzzy, and the emergence of so-called *fake facts* demonstrates that a real convincing rationality is still far away.

3.2 Syntactical Logical Evidence: $S \vdash \phi$

In Europe one can find since the classical Greek Philosophy different trials of philosophers, logicians, mathematicians and others to clarify human thinking by developing different kinds of *Models of Logical Thinking*. The most radical and today mostly used model of logical thinking is that propagated by *formal logic*.[7] The general idea of the so called *syntactical inference* takes a set of given expressions assumed as being *true* – in our case the expressions of the state S – and checks then if it is possible to transform the expressions of S according to some agreed *change rules* X in a way, that one reaches after finitely many operations an expression e which looks like the expression ϕ. In this case it is said that the expression ϕ could be *derived* or *deduced* or *inferred* in a purely logical manner without relying on any kind of *meaning*. This is usually written as $S \vdash \phi$.

The most simple case would be given if the asked expression ϕ is already an element of S, thus $\phi \in S$. To investigate more complicated expressions it would be necessary to *translate* the everyday expressions of S into canonical formal expressions fitting the syntactical rules of formal logic. Although it is not too difficult to transform a substantial part of everyday language expressions $E_0 \subseteq L_0$ into formal logical expressions $E_L \subseteq L_L$[8] we do not (!) want in our project to replace the everyday language by some formal logic language. It is known today that the everyday language is an extremely rich and well developed encoding system. The strategy here therefore is to use the everyday language directly.

With this different goal in mind we have therefore to leave formal logic for a moment and have to look further.

3.3 Semantical Logical Evidence: $M \models \phi$

Shortly after the establishment of the syntactical approach to formal logic[9] basic papers to the idea of a *semantic* version of logical deduction have been published by Tarski (about 1923 - 1938)[10].

[7] For the history of Logic until the new formal Logic see Kneale & Kneale (1962) [KK62]

[8] See the wonderful book of Peter Hinst (1974) [Hin74] demonstrating this with expressions of the German language.

[9] Somehow at the time of the 2nd edition of the Principia Mathematica (1925-7) [WR]

[10] See the collection of papers from Tarski edited by Woodger (1956) [Tar] . See also my lengthy discussion of the position of Tarski with regard to the ongoing project: Doeben-Henisch (2020) [DH20]

At a first glance things look a bit different. In the expression $M \models \phi$ we can rediscover the expression ϕ which stands in question, but left from the inference symbol $\models$ there is an 'M' indicating something which is called a *model* which shall *satisfy* or *fulfill* the expression at the right. Tarski himself uses as an example for such a relation the case that ϕ is the expression $x + y = z$ and as an example for a model he uses a set of triples of natural numbers like $\langle 2, 3, 5\rangle$ and the like. For someone trained in formal expressions the case is clear: ϕ is an expression with so-called *variables* which can be replaced according to some rules by 'appropriate' *constants* and the model is a set of constants ordered as triples which can replace the variable by keeping the order. Therefore $2+3 = 5$ are expressions generated by replacement and this is taken as argument that the model satisfies the expression ϕ.

If one looks to the example from a more general point of view then the triples of numbers represent a set of expressions which are assumed each to be true (like our set of expressions of the state S). The expression ϕ on the other side is not an expression which can be said to be true on account of the variables as part of the expressions. Thus to check whether ϕ is in some sense 'true' ϕ has to be converted into an expression which can be true in the sense that it can be decided with regard to some real situation (in our case a finite set of expressions assumed to be true). Thus in our application case we never are interested in *all possible instances* of the schema ϕ but only whether there exist at least one case in the assumed state S. For instance if the state S would have the expression $\langle 2, 3, 5\rangle$ as a member of S and ϕ would look like $x+y = z$, then one could infer that the model S would satisfy the expression $2+3 = 5$ by replacing the variable by the constants $\langle 2, 3, 5\rangle$. Thus in cases of finite models (what is the standard case in our application scenarios) one can satisfy either variable-free-expressions directly or one can satisfy a finite set of expressions generated by replacing the variables according to the available constants.

But, as already mentioned in the case of the syntactical inference concept of formal logic, the semantic inference concept presupposes formalized expressions too. If we want to work directly with everyday language expressions the semantic formal approach seems not to be a good invitation.

3.4 Computed Turing Evidence: $S \models_{tm} \phi$

No Logic: Using a purely syntactical or semantical logical inference concept has been ruled out above because in this text the usual *formalization* of the natural language L_0 will not be used. Formalization limits the scope of the problem, especially it excludes that dimension which is in this case the most important one: *meaning*.

Meta-Language and Meaning: The assumed application scenario in this text assumes two kinds of written texts D_S, D_X as given and assumes further, that one can generate another third text $D_{S'}$ by *applying* the second text D_X onto the first text D_S. In the simple case one takes one of the individual expressions ϕ_i from the text D_X with the format $\phi_i = \phi_{if} + \phi_{then}$ and then one has to proceed as follows:

1. Show that it holds: $S \models_{tm} \phi_{i.if}$
2. If this is the case, do: $S' = S \cup \phi_{i.then} = S - E^- \cup E^+$

Thus we have two different kinds of operations. In the ϕ_{then}-case one has to *delete* complete expressions or *create* ones. This implies that the ϕ_{then}-expression is a *meta-language* talking *about* the (*object-*)expressions of the documented state S. The *naming* of the object-expressions is in the simple case realized by using the object-expressions in the ϕ_{then}-expression as their own names.

In the ϕ_{if}-case one has also the perspective of a meta-language talking *about* the expressions of the documented state S and the expressions ϕ_i in the ϕ_{if}-part. Additionally it is implicitly assumed that there exists a *relation* between these different expressions which has to be clarified. And in an everyday discourse this relation is primarily determined by the *meaning relation* μ of the participating human actors which convert the expressions of the change rules in D_X into *meaning substrates* $F_{NN.X}$ represented by neural structures and then these meaning substrates $F_{NN.X}$ will be *matched against* the meaning substrates $F_{NN.S}$ of the expressions in the state $D_{NN.S}$. Thus a possible *validation relationship* $\models$ has to be seen and computed as $F_{NN.S} \models_{S.X} F_{NN.X}$.

Assisting Irrational Humans by Rationalizing Machines: These transformations from expressions E into meaning structures F and then matching processes between different meaning structures F, F' are until today mostly *hidden* in the widely unconscious machinery of the brain. To make these processes somehow *feasible* in a *transparent, rational* manner would require to establish formal processes which mimic meaning and matching processes in a way that an *artificial actor* could act synchronously with a human actor.

This implies that an expression ϕ_{if} which has no direct counterpart in the state S or which cannot be shown to be satisfiable by substituting variables by appropriate constants has to be processed in a more sophisticated way. If an *artificial* actor AA should do the job *assisting* the *human* actors HA then the artificial actor must possess (i) the capability to activate an *artificial* meaning function μ_{aa} to convert the expressions of the state S as well as the expressions of the if-parts of the change rules X into their *meaning correlates* F_S as well as $F_{X.if}$. Then, in a next step, the artificial actor (ii) has to *validate* both meaning correlates against each other like $F_S \models_{aa} F_{X.if}$.

These both processes – transforming into the meaning as well as validating meaning against meaning – are processes, i.e. a sequence of states which are generated by applying certain operations to generate successors for given states.

Actually it is not known whether and how one could organize such processes within the syntactical or semantical inference concepts of formal logic. But it is known that there exists a formal concept called *Turing machine* which accepts symbolic input and allows the definition of a great variety of operation-sequences which can be adaptive and which can – in principle – compute all such kinds of processes needed for meaning clarifications and validations between meanings.[11] Thus although it is not known from the beginning how the meaning structures will look like during the course of time, the Turing machine structures allows an adaptive handling of such structures. This is *bad* in the beginning, if in the beginning possibly no working structures are available, but it is *good* because those structures can be built up in a dynamic way.[12]

Learning to Learn: That Turing machines TM – especially in the version of the *Universal Turing Machine* UTM – can in principle learn nearly everything is known from the beginning. But that we even until today have no example of a *really learning* TM hints to the implicit difficulty how to determine *general learning* successfully. As in the general case of biological life where the biological structures allow in principle any kind of learning and intelligence we have to recognize that it needed about 3.5 Billion (10^9) years until the advent of human actors. The capabilities of human actors are until now the most impressive structures which exist in the whole known Universe, but they are nevertheless – as we experience every day – very limited, slow, mostly not very precise, to mention only a few points. Thus these hard facts of nature should keep us aware of the fact, that the construction of real and effective learning devices seems not to be a *soft case*.[13]

[11]The today called *Turing machine* concept goes back to Alan Matthew Turing (1936-7) [Tur 7] which introduced 1936 a formal concept in a paper, where he showed that one can get the famous undecidability theorems of Goedel from 1931 [Goe31]by using an ideal device which can only manipulate symbols on a sheet of (infinite) paper, which was remarkably different to the Goedel-Numbering which Goedel was using.

[12]Turing himself has considered the case of a *learning Turing Machine* very early. One example is a technical report from 1949 where he discusses the case of a learning Turing Machine explicitly, see: [MM69]

[13]The history of Artificial Intelligence [AI] (cf. Nilsson (2010)[Nil10]) shows how hard the way was and is to get good results and a modern textbook about AI (cf. Russell & Norvig (2010)[RN10]) shows besides impressive results a vast empty space of missing solutions.

4 Looking Ahead

With these considerations it has become clear that the implementation of a *simulator* Σ understood as an *artificial actor* AA which shall assist human actors in their planning and investigations of possible future states has to be planned as a process where the simulator in the beginning can only be of limited help. But during the process of assistance it is planned to enhance the capabilities of the simulator step by step making the simulator a *true learning simulator* which can hopefully assist the human actors more and more.

References

[Aus55] J.L. Austin. *HOW TO DO THINGS WITH WORDS. The William James Lectures delivered at Harvard University in 1955*. Oxford University Press, 1 edition, 1955. The content of this book started in 1939 and has been enriched and enlarged all the years until the Harvard Lectures in 1955. This lecture has been published posthumous 1962 based on the lecture notes of Austin.

[BM10] Bernard J. Bars and Gage Nicole M. *COGNITION, BRAIN, AND CONSCIOUSNESS. Introduction to Cognitive Neuroscience*. Elsevier - Academic Press, Amsterdam - Boston - Heilberg et al., 2 edition, 2010.

[DH] Gerd Doeben-Henisch. Semiotic machines. In E.W.B. Hess-Lüttich et al., editors, *Signs & Space – Raum & Zeichen. An International Conference on the Semiotics of Space and Culture*. Gunter Narr Verlag. pp.313–327.

[DH20] Gerd Doeben-Henisch. Reviewing tarski's semantic and model concept. 85 years later .... *eJournal uffmm.org, ISSN 2567-6458*, pages 1–11, 2020. https://www.uffmm.org/wp-content/uploads/2020/08/review-tarski-semantics-models-v1-printed.pdf.

[GB18] Nicole M. Gage and Bernard J. Baars. *Fundamentals of Cognitive Neuroscience: A Beginner's Guid*. Academic Press - Elsevier, London - Oxford - San Diego - Cambridge (MA), 2 edition, 2018.

[Goe31] Kurt Goedel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, i. *Monatshefte fuer Mathematik und Physik*, 38:173–98, 1931.

[Hin74] Peter Hinst. *Logische Prodädeutik. Eine Einführung in die deduktive Methode und Logische Sprachanalyse*. Wilhelm Fink, 1 edition, 1974.

[KK62] William Kneale and Martha Kneale. *The Development of Logic*. Clarendon Press, Oxford (UK), 1 edition, 1962. Reprinted 1986 with corrections.

[MM69] Bernard Meltzer and Donald Michie, editors. *Intelligent Machinery*, number 5 in Machine Intelligence, Edinburg, 1969. Edinburgh University Press. The contribution 'Intelligent Machinery' from A.M.Turing is a report from 1949.

[N9̈0] Winfried Nöth. *Handbook of Semiotics*. Indiana University Press, Bolomington - Indianapolis, 1 edition, 1990. Enlarged and completely rewritten edition of the 'Handbuch der Semiotik' (1985).

[N0̈0] Winfried Nöth. *Handbuch der Semiotik*. J.B.Metzler, Stuttgart - Weimar, 2nd edition, 2000. Completely rewritten 2nd edition of the 'Handbook of Semiotics' (1990).

[Nil10] Nils J. Nilsson, editor. *The Quest for Artificial Intelligence. A History of Ideas and Achievements*. Cambridge University Press, New York, 2010.

[RN10] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Inc., Upper Saddle River, 3 edition, 2010.

[Tar] Alfred Tarski. *Logic, Semantics, Metamathematics. Papers from 1923 – 1938*. Oxford University Press. This book contains the major contributions of Tarski to logic before the World War II.

[Tur 7] Alan M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2):230–265, 1936-7.

[Wit53] Ludwig Wittgenstein. *Philosophical Investigations*. Macmillan, New York, 1953.

[WR] Alfred North Whitehead and Bertrand Russell. Principia Mathematica. PM has been first published 1910-13. The 2nd edition has been published 1925-7. This ist the 8th impression of the 2nd edition.