"First International Round-Table on Abstract Intelligent Agent", 25-27 Jan.1993, Rome, Pub. ENEA

TOGA

A Methodological and Conceptual Pattern for Modeling Abstract Intelligent Agent

Adam Maria Gadomski

ENEA, CRE Casaccia,
INN-RIN, S.P. 88,
00060 S. Maria di Galeria (RM), Italy
E-mail: gadomski_a@eca401.enea.it

Abstract

The necessity of creation of some conceptual patterns for modeling "intelligent" systems which could be accepted by the majority of AI and cognitive science specialists, is discussed.

The definition of "intelligence" as a specific complex property of dynamic system is suggested. The proposal of the functional structure of an abstract intelligent system is analyzed.

It is stressed that modeling an abstract intelligent system requires the assumption of some hierarchical abstract conceptualization frames.

The purpose of this paper is to suggest some patterns for modeling the Abstract Intelligent Agent (AIA) based on the TOGA (Top-down Object-based Goal-oriented Approach) methodology. TOGA is composed of three basic conceptualization systems: The Theory of Abstract Objects (TAO), the Knowledge Conceptualization System (KNOCS), and the Methodological Rules System (MRUS).

The intention of TOGA is the integration of engineering and cognitive paradigms, but .. any integration of independent ideas requires a compromise.

1 Introduction

Development of AI technology is stimulated by new, ever more complex industrial tasks and by the continuous increase of the risk of technological disasters caused by humans, in particular, in such contexts in which:

(a) the amount of information from the domain of interest is so big, or the density of it is so high that the probability of human errors strongly increases,

(b) problem solving requires from human remembering, mental elaboration and proper application of too complex and to large for him professional knowledge.

(c) access to the data is too difficult or requires too much effort,

(d) particular interests and human emotions are stronger than the rational motivations

(e) false decisions and faulty actions lead to dangerous unexpected situations.

(f) physical domain of intervention is not accessible for humans.

Therefore, different computer systems for autonomous execution of tasks, for "fitting" complex machines' functions to the needs and abilities of their human users , and for supporting individual human decision making are required [Kan,89], [IAEA-TECDOC,88], [Mancini,86].

Currently, two approaches at the development of these systems are present in the literature.

The first is classical: in well foresight and defined situations a system (=agent) uses a fixed knowledge which is in forms of algorithms or procedures, and is organized in tree forms. It means, the objects of interventions, tools, and their attributes are initially established. The system can only percept the values (quantitative or qualitative) of invariant process variables and parameters, i.e. these values are the system input data .

The second approach is based on expert systems technology, i.e. it relies on acquisition and processing of qualitative, heuristic knowledge by meta-knowledge included in so called inference engines. Here the input data are also the active part of the executed calculus. In other words the system can acquist and modify rules and algorithms i.e. its own temporal "knowledge".

New and especially promising "intelligent" autonomous and decision support systems are those which are able to utilize both approaches for *supporting, substituting* and also *evaluating* some human mental activities. Here, this type of computer systems is called ICA (Intelligent Computer Agent). For example, such systems should assist plant operators in identifying their mis-conceptions and lack of understanding of plant status, and should play different roles, such as the role of : intelligent executor, advisor, controller, coordinator or tutor [Stock,89].

Basic problems for design and development of ICA are acquisition, conceptualization and representation of knowledge necessary for particular problems [Minski,75], [Winograd,87]. Recent research results, for ex. [Mancini.,86], [Kay,91], [Nwana,91], [Balducelli,93] lead to the conclusion that for identification and specification of a *problem oriented knowledge* for ICA, a functional model of its human users is required.

According to this, we can assume that both models should based on a more general model of an *Abstract Intelligent Agent* (AIA) which realizes goal driven interventions in its abstract environment [Gadomski,89].

Modeling of AIA requires the integrated conceptualization framework and a methodology suitable for [Gadomski,91]:

- the design of an "intelligent" agent which has some properties of goal-oriented human mental activity,
- the specification of some observation, control, management and communication problems of human agents under time constrains,
- the goal-oriented description of physical artificial dynamic systems.

Following the above assumptions, the present paper will discuss:

(1) some basic elements of the conceptualization theory **TOGA** (Top-down Object-based Goal-oriented Approach) which enables representation of the "intelligent" activities of artificial dynamic systems, and

(2) an attempt to the definition of general patterns for the functional modeling of goal-oriented intelligent agents.

A formal specification frameworks for the representation of the problem knowledge, problem management knowledge, and knowledge acquisition are indispensable tools for ICA designers (knowledge engineers). This meta-knowledge is necessary for:

- *acquisition* and selection of a problem oriented knowledge,
- *allocation* of this knowledge to an ICA system,
- *definition* of a new ICA- human user *communication knowledge* and ICA new cognitive interface functions**,**
- *standardization* of the ICA life-cycle documentations.

The paper is based on the author’s methodological research [Gadomski, 88a, 90a, 90b], and experiences referred to the study and design of different types of computer systems for decision making and plant operators support [Businaro,88], [Gadomski, 88b], [Balducelli,90].

2 Conceptual Framework: the TOGA Theory

TOGA is a system of *conceptual frameworks* and the methodology of complex problem specifications . It has been generalized to the form of the foundations of the theory of Abstract Intelligent Agent. At present, the formal structure of TOGA theory is yet under development. The current effort is focused on two issues:

- theoretical which includes the development of a formalized model of intelligent reasoning processes,
- applicative which concerns the construction of a software prototype of an autonomous intelligent reasoning module.

TOGA theory is composed of three elements:

1. the **Theory of Abstract Objects** (TAO)**,** which is the primary domain independent conceptualization system;
2. the **Knowledge Conceptualization System** (KNOCS**)**, which constitutes the second level conceptualization system, i.e. the axiomatic assumptions and definitions related to: conceptualization of the real world, realization of an *abstract intelligent agent* (AIA) and its goal-oriented activity;
3. the **Methodological Rules System** (MRUS) for the specification of complex problems.

2.1 Elements of the Theory of Abstract Objects

TAO is a mathematical construction but it is also a conceptualization system which can be used for representation of perceived impressions and symbolic imaginations involved in human goal-oriented activity.

The theory refers to *abstract objects,* according to the Oxford Dictionary ,

abstract means "separated from matter", a conceptual product obtained by neglecting some properties of analyzed thing .

Following this, by an *abstract object* in the TAO theory is intended *a conceptual representation of any object or a property abstracted from its physical realization, or if it is a mental construction, abstracted from some its imaginary properties.*

For intelligent computer agents, TAO is assumed as a primary conceptual context of symbols and images recognized physically in the ICA environments.

TAO is based on the network concept and on generally known mathematical theories, such as set theory, functional analysis, and graph theory.

Below, the basic concepts of TAO are formally presented.

Any theory can be considered as a frames system which enables structuralization and operation over a certain class of sets. In the case of TAO, any numerable set called *primitive set.* is its domain .

TAO is a frames system which enables structuring of primitives in the forms of:

- *Objects*, specified by their *names, attributes’ names, values*, and *value domains*;
- Relational isolated networks of objects, called

'world-of-objects' (w-o-o) which:

* can be arbitrarily divided into *systems* and their *environments*,
* can be aggregated in *universes* of objects linked by *r-connections*.

The TOGA's abstract object frame is not defined by the reference to the real world nor to a programming environment, but it is a conceptual scheme which may represent formally any concept that could be described by : object-name, attribute names and their values.

The abstract objects are the elements of every conceptualization of any (physical and mental) domain of activity of an intelligent system. In this sense, *process*, *relation*, *change* can also be considered as objects in the adequate worlds of objects[1].

Definitions of *Object* and *Object Frame*

For an AIA, an object may be everything representable in terms of the set theory, as a ordered couple

(On, A),

where: *On* is a primitive called *object name*,

A is a subset of primitives in role of *attribute names*,

and *On* is not an element of *A.*

Object name and *attribute name* are descriptors.

Object frame (o-frame) is a metadescriptor and is defined as the ordered couple of descriptors :

(*object name*, set of numbered *attributes names*).

Every attribute name is an element of *A,* and it is connected with an *attribute value*, *w*. The ordered couple *(a,w)* is called an *attribute*.

Every attribute value is a point in an attribute space, *W*.

The following classes of possible attribute spaces are taken under consideration:

WN - set of non ordered text-expressions,
WO - set of ordered text-expressions,
WF - set of mathematical functions,
WA - set of areas in defined numerical space,
WR - numerical space.

Roughly speaking, an attribute may have qualitative or quantitative value domain .

In order to present any abstract theory we must accept a formal representation of introduced concepts. TAO is represented in two representation symbology , mathematical and graphical.

In the mathematical notation we can represent an object Q as follows:

Q : *On* [*(a1: w1,W1,**W1**)*, *(a2: w2,W2,**W2**)*, ...],

where *(ai: wi,Wi,**Wi**)* for i=1,2,... are attribute representations,

Wi denotes a domain of variability of the i-th attribute value in the attribute space ***Wi,*** and

On denotes an object name **.**

Let *{Q}* be a set of objects $O1[A_1]$, $O2[A_2]$,..$ON[A_N]$ and

if exists non empty subset of attributes *AS* such as

$$AS = A_1 \cap A_2 \cap \ldots \cap A_N,$$

then the Cartesian product of all *AS* components, represented as

$$Was_1 \times \ldots \times Was_m,$$

for i=1, ...m, will be called a *'common space'* of *{Q}*.

Of course, the properties of a distinguished object space depend on the assumed nature of AS set attributes.

Definition of *o-relation*

Let *{Q}* be an object set and *Q1, Q2* (- *{Q}* then, if *a(Q1)* is an attribute of the object *Q1* and *b(Q2)* is an attribute of *Q2* then the following expression:

r(Q1,Q2) : rn [a(Q1), b(Q2)],

where *rn* denotes a primitive,and *(Q1,Q2)* an ordered couple, will be called *o-relation* between *Q1* and *Q2* , and *rn* will be called *o-relation name*.

Definition of *system*

An ordered couple *({Q},{r})* denotes a *system*, *Sn,* iff

- all *Q* (- *{Q}* have a common space,
- for every *(Qa,Qb)* , where *Qa,Qb* (- *{Q}* , exists *r[Qa,Qb]* (- *{r}* , and
- for every

(r[Qa,Qb] (- {r}) -> (Qa v *Qb) (- {Q}.*

Definition of *structure*

The set *{r}* will be called *structure* of the system *Sn: ({Q},{r}).*

Definition of *world of objects*

If *({Q},{r})* is a system and *Qa ,Qb* (- *{Q}* for every *r[Qa,Qb]* (- *{r}*

[1] Some notational conventions**:** - " term " denotes an intuitively used notion of the word 'term',

then this system is isolated and will be called
a world of objects, *w-o-o* , denoted Ω .

A graphical representations of *system* is presented in the Fig.1.

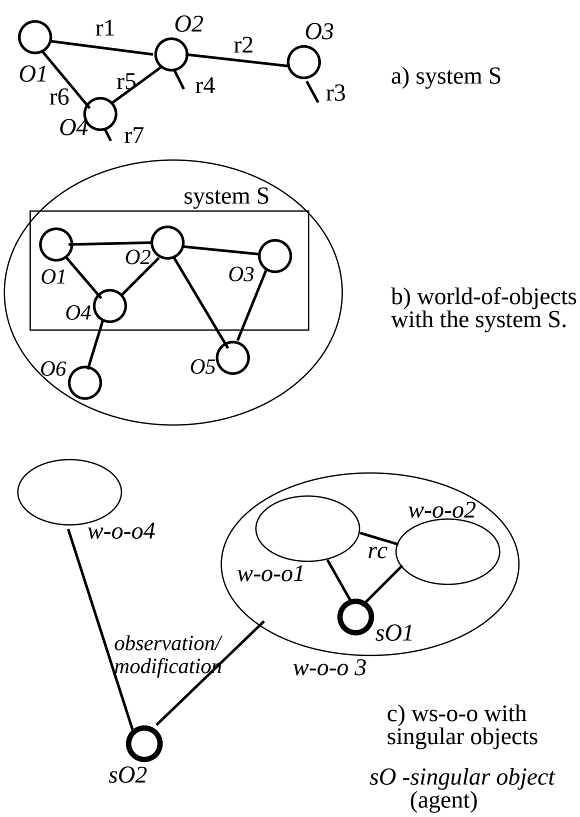

Fig. 1 Examples of basic structures in TAO.

A w-o-o *({Q},{R})* is complete if for every attribute *a* of every object *Q* (- *{Q}* exists such *b{Q'}* and *Q'* (- *{Q}*
then

$$R [a(Q), b(Q')] \quad (- \ \{R\}.$$

For example, two objects $O_1[a]$, $O_2[a]$ with the relation $r [a]$ represent a minimal complete w-o-o.

In graphical representation two objects are in a relation iff they have a common points in the representation space, for ex. it is illustrated in the Fig.2.

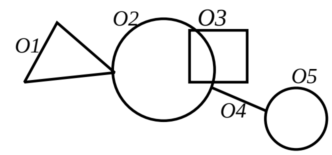

Fig.2. An example of the graphical representation of objects with a reciprocal relations.

Definition of *r-connection* and *point of view*

Two ws-o-o can be linked by common primitives, which are different roles (i.e. object name, attribute name or relation) in different ws-o-o. It is called *r-connection.*

For ex. if the word 'title' is a primitive then in one w-o-o it can be the attribute name of an object named 'book' but in another w-o-o 'title' may be an object name.

For ex . $\Omega 1 \ \rho(P: a, O) \ \Omega 2$ means that $\Omega 1$ and $\Omega 2$ have a r-connection, where *p* - is a primitive and *a* denotes its attribute role in $\Omega 1$, and *O* denotes its object role in $\Omega 2$.

One of the new ideas included in TAO is the definition of the class of *singular objects,* and the formalization of the concept of the *point-of-view* referred to an object.

Point-of-view (p-o-v) is a complete representation of an object into a pre-chosen w-o-o. The representation of a selected object X *from the p-o-v of* another object Y includes only these attributes of X which are linked with Y by their common relations.

Operations on ws-o-o: *abstraction* and *specification*

The *singular objects* are particular objects which can observe create or modify other objects inside world-of-objects.
Singular objects can be internal or external, relatively to the ws-o-o which are the domains of their activity.
Of course, they can be considered "normal" objects in another universe of objects.
Operations on ws-o-o and on objects universes are unique attributes of the external singular object in TAO.
One of the TAO operations is an *abstraction operation, {Âb},* which transforms systems or ws-o-o.
Any w-o-o can be the base for other *descendent* ws-o-o obtained by *abstraction*.
Abstraction does not have reciprocal operation, i.e. $\hat{A}b^{-1}$ does not exist.

$$Y = \hat{A}b \ X .$$

Abstraction operation reduces information X about primary objects.
Every abstract object in D can be obtained as a product of operations on lower, higher and parallel ws-o-o.

Definition of one w-o-o determines its relative "ortogonal" abstraction hierarchies.

We can distinguish two types of abstract objects: *descendent abstract objects (dao)* and *absolute abstract objects (aao)*. Contrary to *aao*, every *dao* has a link with its ancestor .
Another operation is *specification,* it increases information about objects and relations.
Specification enables decomposition of elements of w-o-o on subelements.

Abstract operation is every operation performed on abstract objects , therefore *specification* is also an *abstract operation*.
The problem of "abstraction" has the rich and different representations in the literature, see [Balducelli,93], [Giunchiglia,92]. For ex. the theory of abstraction proposed by Giunchiglia and Walsh can be applied to the top-down representation of the TAO concepts.

2.2. Knowledge Conceptualization System

The Knowledge Conceptualization System (KNOCS) is the *second conceptualization system* which is based on axioms and definitions related to the conceptualization of the real world from the perspective of a real AIA.
Primary KNOCS assumption is that every product of the human reasoning activity can be conceptualized in the framework of TAO.

KNOCS is the system of axioms and definitions for the description of interactions between intelligent agent and the real world. It enables the conceptualization of different physical systems such as industrial plants, robots, human operators or organizations.
If we assume that the real world is only the source of physical signals then from the point of view of their receivers, *information* can be the subjective attribute of these signals.
Let us call some receivers *agents.*

If we assume that the same signal can be a source/carrier of different informations for two different agents then we will easy accept that some agent's internal properties are responsible for information creation. Let us call this abstract functional object *perceptor* $\underline{\Pi}$, and these function *perception.*

Let π is an information for a system *S*, if it does not change invariants of the system *S* (for ex. system physical structures) in such case, the behavior of the system *S* is a *information sensitive*.

We should remember that up to now, information, perceptors, and perception are only considered functional abstract concepts.
Subsymbolic information are such physical properties of signals which modify variable elements of a system, for ex . input data for neural network.

In KNOCS, the perception function is represented by the perception operator Π, and :

$$\Pi: \quad s \quad \text{-------}> \pi$$

where *s* is a signal.
Now π can directly activate agent's external response or can be memorized as a *concept,* i.e. as a new active property of agent physical memory.
All operations on these concepts which are caused
by other concepts structured before, are performed by an abstract agent.
In the similar way, some sequences of concepts (recognized as abstract agent output) activate functions of physical agent (human or computer).
From the perspective of an absolute observer, every π must be expressed in terms of one of agent's conceptualization systems.

TOGA includes three connected conceptualization layers illustrated on the Fig. 3 :
- TAO , as primary conceptualization tool,

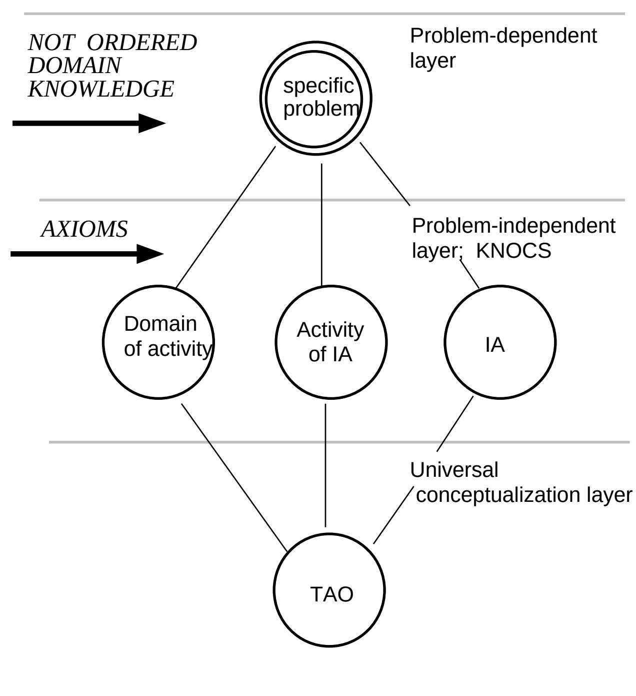

Fig. 3 Three layers of the KNOCS conceptualization.

- KNOCS conceptualization layer. It is problem independent, and includes three frameworks of interactions between IA and its domain of activity: domain-of-activity (d-o-a), IA model, and model of goal-oriented IA activity.
- Specific problems layer, i.e. problem-dependent conceptualization.

For example:
- specific concept *painting* is *an intervention* in the KNOCS layer , and a *relation* in the universal conceptualization layer,
- a *factory* is a *system* and an *object* subsequently.

2.3 Methodological Rules System

In the perspective of AIA, the Methodological Rules System (MRUS) is a methodological approach to a "top-down" and "goal-oriented" knowledge ordering for the specification of complex problems.
MRUS is based on previously introduced conceptualization.
It assumes that, at the beginning of a problem specification, the knowledge of the problem solver agent is incomplete and not goal-ordered. The problem specification activity is based on two fundamental mechanisms:
- the former is called the *top-down* mechanism, and indicates the specification direction: from very general statements on the top levels to the details which can be the elements of the problem solution;
- the latter is called the *goal-driven* mechanism; it always controls the links between the specified/identified object and the problem *goal* object; this mechanism creates bottom-up rules (synthesis rules).
MRUS can also be a tool for checking the correctness of the goal-oriented activity of designers, and for the validation/verification of their products.

3 Conceptualization of the *domain of activity*

3.1. SPG

One of the most important problems related to ICA design, as well as to human rational reasoning, is the goal oriented qualitative and complete modeling of its domain of activity which is recognized as a physical artificial dynamic system [Engelmore,90], for example, an industrial plant is a d-o-a of its human or artificial operators. In TOGA, such conceptualization framework is called SPG (System-Process-Goal approach) and is represented by objects' networks in the three-dimensional discrete space [Gadomski,88a,89].
The first dimension called Goal-System Interrelation (GSI), is divided into four layers: *design-goals, functions, processes*, and *systems*..

The design-goals are defined in the context of the the system user needs and society expectation. According to Lind's conceptualization [Lind, 82], the top design-goals of a goal-oriented system are: *production, economy*, and *safety*.
The system functions are direct consequence of the specification of its design-goals. They are goal-oriented properties of the projected system.
The key element of GSI is the formal distinguishing between functions and processes.
The status of system *functions* depends on *processes* and *system* current structure, i.e. configuration, integrity.
The processes are the *carriers* of functions.

The *carrier relation* is a relation between two abstraction levels, it is transitive and asymmetric.
The *property relation* is inverse to the carrier relation.
For ex., not all attributes of the heat transfer process are attributes of a heat transfer function necessary for a certain goal.
The systems are the carriers of processes.

The Fig. 4 illustrates the GSI concept. The relations among: goals, and among goals and functions are cause/consequence relations. The relations between systems are synthesis/decomposition relations.

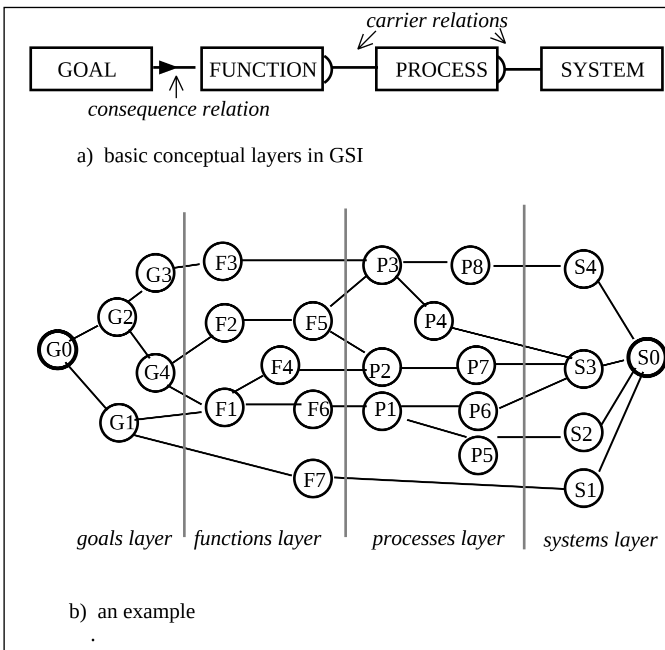

Fig. 4 Rerepresentation of the knowledge about a dynamic system according to the GSI pattern [Gadomski,88a].

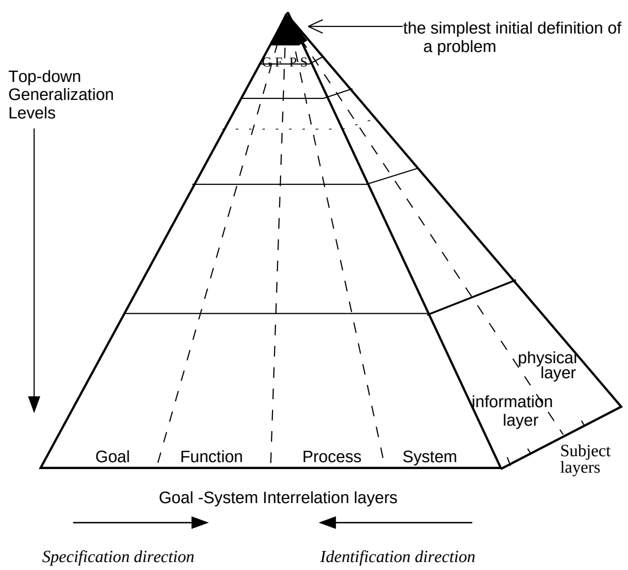

Fig. 5 The TOGA framework (SPG) for the representation of problems related to goal-oriented systems.

Remarks:
- One selected process can be the carrier of many functions.
- One function can be a property of different processes/systems.

The second dimension gives the possibility of setting up the description of systems, processes and functions on different *generalization levels* **(GL)**. GLs are organized arbitrarily, from assumed initially generalization up to the details level, where, for example, an implementation of system model is possible.
If the current d-o-a D is an abstract activity domain and it is represented as a w-o-o (for ex. an image of predefined end-activity-domain, then D can be a base for other descendent ws-o-o obtained by *abstraction operations*.
Therefore, knowledge structured in GL levels is always referred to the same pre-chosen system.

The third dimension is used for a set up of the system description in the *hierarchy of subjects*, from directly measured (physical) attributes to highly abstract conceptualizations, for ex. it can be composed of material, structural, automatic and information layers, or hardware, software and methods layers.

The pyramidal, top-down problem structuring requires bottom-up goal-driven evaluation and synthesis of available to the agent, informations, knowledge and its preferences.
In general, information about the same physical object can be "filtered" and conceptualized from different points of view (p-o-v), for example, supervisory p-o-v, diagnostic p-o-v, advisory p-o-v, or safety p-o-v.

Here, we should stress that important problems of the ICA designers related to the particular applications, are:
(a) specification of descriptional ps-o-v (depending on the possible for ICA intervention-goals),
(b) finding connections between such constructed knowledges, and
(c) integration of these knowledges into the frame of a pre-assumed knowledge representation formalism.

The methodology for knowledge conceptualization suggested in this paper starts from the specification of the goals of the designed system and the goals of the ICA activities as two fixed main points-of-view. Such approach also enables the consideration of different ICA actions as specific points-of-view for knowledge functional structuring.

In the case of identification of the natural systems, the role of these systems in social or biological contexts can be treated in the same way as a design goal, and can be considered as pseudo-goals. Then, such reconceptualization enables the application of SPG to the above systems.
More precise description of SPG is presented in [Gadomski,88a].

4 Conceptualization of the activity of IA

Let as to begin from the example related to a plant operator activity [Gadomski,92]. The operator is an element of the behavior of any industrial-plant aggregate.
During normal plant exploitation, the plant status and its observability and controllability are inside the thresholds established by the plant designers, and operators change the plant status according to externally pre-defined tasks. From operator operational perspective, these tasks are his *intervention-goals***.** Other intervention-goals are established by the operator himself as the consequence of the recognition of plant abnormal situations or alarm signalizations.

In order to achieve the intervention-goals, the following operator top functions can be distinguished:
1) monitoring of plant attributes,
2) evaluation of plant safety, production and economy,
3) foresight of plant status,
4) interventions planning and execution.
They are the consequence of this IA role.

The conceptualization of activity of IA related to a physical system is represented in KNOCS as a formal interrelation between an *intervention goal* and its *execution environment*.
Execution environment includes IA and its end-d-o-a,

i.e. the domain where the final *intervention goal* is localized.

From the our point of view, considered as an external singular and intelligent object, we assume that:

If the d-o-a of an AIA X includes another AIA Y then X's knowledge related to a selected goal-oriented activity of Y may be conceptualized in terms of formally defined :

intervention-goals, *tasks* , and *actions* referred to the Y's means and its d-o-a .

Def.

The *task* is the property of an *action* oriented on predefined *intervention-goal*, it is expressed in terms of d-o-a description, and describes which changes must be introduced in the AIA d-o-a for achieving an *intervention-goal*.

Def.

The a*ction* is a specification what AIA is to do in d-o-a for a realization of tasks, i.e. in order to achieve the predefined goal .

For such reason, an *action* must depend on executor possibilities, and in different circumstances, one *task* can be performed by the execution of different alternative *actions*.

From the identification point of view, one selected *action* can be recognized as a *carrier* of different *tasks.*

Tasks System (task scenario) is executor independent and it depends on goal constrains.

Interrelation between an intervention-goal and its execution environment is represented on different *generalization levels*, GL, and can be related to different types of means in the abstraction hierarchy, *means level*, ML.

MLs dimension can be divided for example on: physical means, communication means , and reasoning means levels.

Finally, the proposed conceptualization is similar to SPG illustrated on the Fig. 5.

We can notice that the frameworks presented above structuralize responses on the following questions:

Who makes the intervention?

Where is it?

Why is it done ?

How is it performed?, and

How deep our knowledge is about the analyzed event?

The answer to the question: When the intervention goal is created? is omitted yet in this chapter.

It is interesting to stress that in KNOCS, if a goal was established then an unknown ignorance cannot exist, ignorance must have attributes because they defines relations for a closure of every real w-o-o.

Human reasoning referred to a d-o-a is based on many conceptual systems and associative processes. A mixed, not verified in "real time" changes of conceptualization contexts, frequently lead to false conclusions and intuitive convictions, i.e. to the construction of a false or "fiction" domain-knowledge which no longer has reference to any physical or abstract d-o-a of human agents. We can notice that the human mind is full of such types of constructions. These situations however can be omitted in the design of an artificial AIA.

From the perspective of the AIA knowledge, there are three types of objects/systems/sub-domains in the d-o-a:

Dt1, which is driven only by physical principles; hypothetical goal is unknown for the agent (a singular object),

Dt2, which is fixed goal-driven, their design-goal is known for the agents,

Dt3, which is driven by temporal intervention-goals.

In the Dt1 domain, the agent can navigate and manipulate.

The Dt2 domain is supervised and controlled by the agent according to its intervention-goals.

The Dt3 domain can also be the domain of cooperations and other social interactions.

We can assume that AIA is, at first, event driven, for the recognition of environment status; after this, it is consequences-driven (on the base of his temporal knowledge) what includes a risk assessment; finally, it is time-driven, i.e. it searches better solutions according to time constraints.

Therefore, before any intervention, AIA should know (in order of priority):

- how to avoid extreme negative consequences,
- which are acceptable consequences,
- how to obtain required environment status,
- how to optimize some attributes of the action.

From the ICA designer perspective, every rational human mental function can be allocated to ICA system if:

- it is formalized and represented in a unique manner by means of input/output processes,
- it is formally disjointed from other human reasoning behaviors,
- a clear conceptual interface with humans or other modules of ICA can be designed.

From the perspective of human agent identification, strong relations exist between AIA and its physical carrier agent. This problem is discussed in the next chapter.

5 Basic patterns of IA and AIA models

5.1 General assumptions

If every abstract system must be a property of physical one then AIA can be separated from other properties of physical intelligent agents.

"Not intelligent " and physical part of IA is called CIA (Carrier of an abstract Intelligent Agent).

AIA, as a structural property of CIA has no possibility to change directly "its" physical environment but can activate these changes in CIA according to its own conceptualization of them. Therefore CIA is also a direct executor of the all interactions between AI and its environment.

In the case of human IA, his carrier, CIA, can be considered a partially autonomous agent which, on the subsymbolic level, not only realizes perception and execution functions, but it communicates with AIA and changes the performance of AIA's basic internal functions.

5.2 Model of AIA

The AIA model represents a semantic context of the term *knowledge*. Thus, *knowledge* is in *reference* relation with a predefined domain of activity of the agent, in *attribute* relation with this agent, and in *activation* relation with the *preference system.*

In TOGA, the basic element of AIA is called *Abstract Agent* (AA). It is composed of three fundamental functional elements :

- *Domain-of-Activity (d-o-a),*
- *Knowledge System,*
- *Preference System.*

Domain-of-activity is always considered as the *information source*.

Information is referred either to the state of the d-o-a itself or to the state of other objects which are symbolically represented in this d-o-a.

The *d-o-a* of an *AA* is the *reference domain* of its knowledge system, and, from the point of view of its observer it is called Knowledge Reference Domain (KRD).

Knowledge system is the *carrier* of different *reasoning processes,* for ex. the information processing or the information choice.

The *goal-oriented* reasoning processes start from the *activation* of the knowledge system by some states of the preference system.

The results of these processes can change the state of the d-o-a of the AA.

Preference system is activated by information from the agent d-o-a and it activates the knowledge system by generation of a so called *intervention-goal*.

The *intervention-goal* is the conceptualization of the state of the d-o-a to be achieved.

The functional scheme of AA is presented on the Fig. 6

Now let us introduce a definition of an *abstract intelligent agent:*

Def.

An Abstract Agent which is able to reason about and modify its own knowledge and preferences is called an Abstract Intelligent Agent.

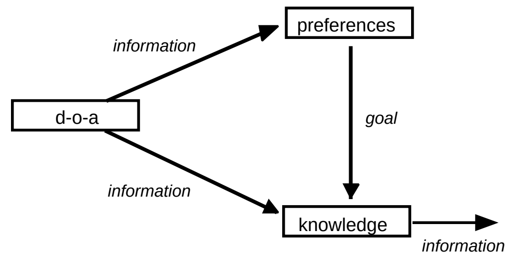

Fig. 6 Basic functional module of a reasoning process of Abstract Agent .

Functional structure of an AIA

The functional structure of an AIA is a tree network of AAs (see Fig. 7). The AA at the root of the tree network is called the *basic-AA.* Starting from this basic-AA, the *knowledge system* and the *preference system* of each AA, on any level of the tree, are the domains-of-activity of the AAs on the subsequent level. The d-o-a of the basic-AA is called the *basic-domain-of-activity (b-d-a)* of the AIA.

In the structure of an AIA, the sequence of knowledge systems constitutes *metaknowledge levels.* The metaknowledge levels obey the following rule:

> the (n)th - metaknowledge is the reference-domain for the (n+1)th - metaknowledge,

where n=0,1,2,..., and n=0 denotes the knowledge referred to the basic d-o-a of AIA.

Every i-metaknowledge (i=0,1,2,..) is aggregated in the form of:

descriptive knowledge which is the set of frames of a relations (on different generalization levels),

operational knowledge which is a set of rules and procedures.

Procedures are sequences of ordered rules and frame relations which link two hypothetical states of d-o-a..

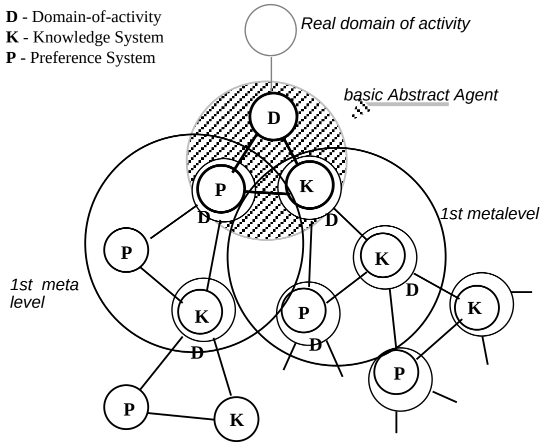

Fig. 7 The TOGA architecture framework of Abstract Intelligent Agent .

The reasoning of AIA is based on two following processes:

1) (information) ----> PREFERENCE SYSTEM -----> (intervention goal)

2) (information, intervention goal) ---> KNOWLEDGE SYSTEM ----> (information related to intervention goal, activation of a CIA execution function).

3) If ,after the 2), the goal is not achieved then the 1) and 2) are repeated on the next metalevel, where the domain of intervention is the previous knowledge or preference system.

In general, every preference and knowledge system can be the domain of activity for the preferences and knowledge from the higher metalevel.
The above reasoning capability should enable realization of all behavioral symptoms of intelligence as well as consciousness of the agent (if it is necessary for achieving an assumed particular goal).

From the cognitive perspective we can say that an "ideal" human goal-oriented activity, in a chosen domain-of-activity, depends only on :
- *information,* i.e. data related to current status of the d-o-a,
- *knowledge*, which processes or selects information and other knowledge,
- *interests*, which depend on interrelations outside the d-o-a, and
- *axiology* (hierarchy of preferences)[Gadomski,90, 91].

In the case of human agent, these elements are modified in different way, by CIA.
We can assume that CIA is a carrier of another *abstract agent* which works according to fixed preferences, buildin knowledge, and which directly communicates with AIA.
The basic relations between AIA and CIA illustrated in the Fig.8, could be very interesting and fruitful domain for cognitivistic and psychological research, for ex. for the problems presented in [Kan.89].

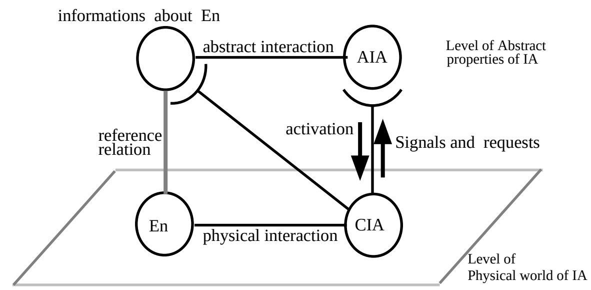

Fig. 8 Relation between physical and abstract elements of an IA.

Conclusions

The frameworks presented above, enable to define intelligence as an *operational capability of the agent to the construction of and reasoning on different abstraction levels in order to intervene in its domain of activity according to its own preferences.*
In the TOGA theory, intelligent agents operate only on information, knowledge, and preferences previously defined, all of them being conceptualized according to KNOCS and TAO frameworks.
One of the possible indicators of intelligence can be the number of meta-levels of reasoning available to the agent's problem solving.
We can notice that the above defined intelligence does not directly depend on the quality of available information, agent current knowledge about its intervention domain, and used preferences.
Therefore, a good adaptation of a physical agent to the complex real-world is a sufficient but not a necessary condition for to recognize it as intelligent.

References

T.Businaro, A.M.Gadomski, G.B.Meo, V.Nanni, G.Vialetto, "La Gestione delle Informazioni nei Sistemi di Supervisione",Tecniche dell'Automazione & Robotica, N.9, N.10, 1988 (in Italian).

C.Balducelli, A.M.Gadomski, G.Vicoli, P.Strula, "Knowledge Based Computer Aids for the Plant Operator: An Application in the Management of Chemical Pollution Crises", in Proceedings of the Congresso Annuale dell'Associazione Italiana per l'Informatica, Bari, Italy, September 19-21, 1990.

C.Balducelli, A.M.Gadomski, Intelligent agent in Computer Supported Training for Emergency Management, AIA93 Round-Table, Rome, 1993. In this Proceedings.

R.S.Engelmore, "Knowledge-Based modeling of Physical Devices", AIIA Notizie, N.1., March, 1990.

A.M.Gadomski, "An Application of System Process-Goal Approach to the TRIGA RC-1 Reactor System Description", in Proceedings of the 9th European TRIGA Users Conference, 1986, and ENEA report RT/TIB/88/2, 1988.

A.M.Gadomski, V. Nanni, "Design Criteria of TRIGA-RC 1 Supervisory System: System-Process-Goal (SPG) Approach", presented at the Workshop on Computer Applications for Nuclear Power and Industrial Process Plant Operation and Control - ENEA & OECD Halden Project, Rome, Italy, November 19-20, 1987; and ENEA Report RTI/TIB(88)12, 1988.

A.M.Gadomski, "From Know-how to How-to-Know: An Approach to Knowledge Ordering for Specification of Complex Problems (TOGA methodology)", presented at the International Symposium on Computational Intelligence, Milan, Italy, September, 1989.

A.M.Gadomski, M.Gadomska, "Environmental and Emergency Communication and Decision: Confrontation of Shallow Models from the Point of View of Computer Support Designing", presented at the Second SRA-Europe Conference, Laxenburg, Austria, April 2-4, 1990.

A.M.Gadomski, "A Model of Action-Oriented Decision-Making Process: Methodological Application", in Proceedings of the 9th European Annual Conference on Human Decison Making and Manual Control, 1990, pp.79-99.

A.M.Gadomski, V.Nanni, "Problems of Knowledge about Knowledge: an Approach to the Understanding of Knowledge Conceptualization", in Proceedings of The World Congress of Expert Systems, 1991, pp. 1519-1523.

F.Giunchiglia and T.Walsh, A Therory of Abstraction, Artificial Intelligence, V.57,N.2-3, 1992.

IAEA-TECDOC-542,ISSN 1011-4289, "Use of Expert Systems in Nuclear Safety", Report of a Technical Committee Meeting, Wien, Austria, October, 1988.

C.C.F.Kan, at al., "A Framework for modeling the Behaviour of a Process Control Plant Operator Under Stress", the Meeting on Artificial Intelligence in Nuclear Power Plants, Helsinki, Finland, October, 1989.

J.Kay, "Generalised User Modelling Shells - a Taxonomy", in Proceedings of IJCAI-91 Workshop W.4 'Agent Modelling for Intelligent Interaction', 1991, pp.169-185.

M.Lind, "Multilevel Flow modeling of Process Plant for Diagnostic and Control", RISO National Laboratories Report, RISO-M-2357, 1982.

G.Mancini, "Modelling Humans and Machines", in E.Hollnagel, G.Mancini and D.D.Woods (editors), Intelligent Decision Support in Process Environments, Berlin: Springer-Verlag, NATO ASI Series, 1986.

M.Minsky, "A Framework for Representing Knowledge", in P.W.Winson (editor), Psychology of Computer Vision, New York: MacGraw-Hill, 1975.

H.S.Nwana, User Modelling and User Adapted Interaction in an Intelligent Tutoring System, User Modeling and User-Adapted Interaction:An International Journal, V.1,N.1,1991

M.Stock, AI Theory and Applications in Process Control and Management, New York: Intertext Publications & McGraw-Hill Book Company, 1989.

T.Winograd and F.Flores, Understanding Computers and Cognition. A New Fundation for Design, Norwood, NJ: Albex Corporation, 1987.