Foundations of Autonomy and Its (Cyber) Threats: From Individuals to Interdependence: Papers from the 2015 AAAI Spring Symposium

Algebraic Models of the Self-Orientation Concept for Autonomous Systems

Olivier Bartheye
CREC St-Cyr
Écoles de Saint-Cyr Coëtquidan
F–56381 Guer Cedex, France
olivier.bartheye@st-cyr.terre-net.defense.gouv.fr

Laurent Chaudron
ONERA
BA 701
F–13661 Salon-de-Provence, France
laurent.chaudron@polytechnique.org

Abstract

The aim of this paper is to define a both pragmatic and formal method allowing a social or technical entity to define its own strategic goals and plans. Indeed, by definition, an autonomous entity ought to be governed only by its own principles and laws. Thus, the core concept of autonomy is the capability of defining this principles regarding its own objectives and plans. Thus, the robustness of any autonomous system relies on the pivotal concept of **Self-Orientation**. This paper focuses on the first formal steps of Self-Orientation theories for any group of agents.

As an introductory test, we propose a pragmatic formal approach to the modelling of the Self-Orientation of the generic case of a research lab: a descriptive model for the autonomously defined orientation of a scientific program in a research establishment is described thanks to a Galois lattice based model.

Then in the sequel, a theoretical model is developed so as to build relevant algebraic, topological and axiomatic properties of any self-orientation systems.

This formal apparatus aims also at allowing to provide measures of the complexity of any self-orientation computation process.

Introduction: Self-Orientation of a lab

The concept of Self-Orientation Process

In the research activity, the evaluation and orientation methodologies are included in the organization of the laboratories. Indeed as science research aims at producing new knowledge, the management of this production is naturally conducted not only thanks to the publication activity but also through a permanent process of scientific committees chaired by external experts. But in a perfect world, the scientists that are evaluated are supposed to be the best experts of their own domain; thus why couldn't they propose the orientations of their own discipline?.. Beyond the ethic issues, a methodological drawback appears: on what philological and theoretical grounding could such a self-orientation be based on? The autonomy of the laboratory relies on its capabilities to manage a dynamic of its self-orientation process. That is, on one hand, such a social organization cannot exist outside a modern society whose foundations were built by antic Greeks and on the other hand, such an entity cannot be subordinated to the modern society in which it lives. In other words, the transformation process involved by such an entity must agree with conservation laws whereas it generates entropy at the same time.

In this introductory section, a very simple algebraic model is proposed. Let us consider a research entity as a prototype of any social system claiming for autonomy (with the complexity and diversity of its agents) i.e. a living self-analyzing system, capable of defining in a thoughtful process its own objectives. If this conceptual challenge is fixed, the challenge becomes: *which mathematical structure can capture the* **self-orientation process***,* **SOP***, i.e. the aggregation of the individual orientation proposals in a consistent collective program responding to the ambient social needs?*

A Formal Model of a SOP

The concepts will be simplified so as to allow formal definitions: SOP is a set of scientists: $S = \{S_n\}, n \in [1, \mathbb{N}]$. In any research entity, thus in the SOP, a fundamental ordering is the research activities based on scientific *specialties* (automatic, quantum physics, aerodynamics..) addressing social challenges through *purposes* (noise reduction, air safety, ..). This universal binary relation between the set $Spec$ of specialties and the set $Purp$ of purposes: $Spec \times Purp$. If all elements of S (both the CEO and the scientists) of a SOP may get some synthetic view of "what is done and for what in our lab.", the $Spec \times Purp$ table, cannot provide any orientation cue. Those two sets have to be described through detailed subcategories so as to capture the natural expression of any orientation statement: *"in order to improve these parts of this purpose and for this performance level, we have to program the development of these techniques of this scientific specialty in that way"*. Thus each purpose P of the set $Purp$ is described through a spectrum of parameters $P = \{p_i\} i \in [1, I]$ (eg: 10 main issues of the air safety...).

The societal challenges are combinations of various purposes (eg: improve air safety *and* reduce noise *and* improve speed of aircraft..) and theses societal challenges constitute a database: $E = \{E_m\} m \in [1, M]$. Naturally each E_m is decomposed on the spectrum of parameters P : $E_m \rightsquigarrow (p_{m,1}, .., p_{m,i}, .., p_{m,I})$. As expert in her/his own domain, a scientist, element of S, will propose to valuate the development of some characteristics of the scientific specialties she/he considers as pivotal for the success of the challenges thanks to a set of evaluation characteristics: $V = \{V_j\}, j \in [1, J]$ given one subject S_n and one challenge E_m, the final result of her/his valuation is a vector: $(v_{(n,m,1)}, v_{(n,m,2)}, ., v_{(n,m,j)}, .., v_{(n,m,J)})$.

At this point, a standard requirement of the SOP will be successful if all the members of the SOP can submit their proposals, and if an aggregation means allows to summarize the individual proposals in a collective consistent program. It is easy to consider that the global set of the individual valuation is a relational database of $N \times M$ items: *(challenge, subject, list of parameters of challenge, list of valuations by the subject)*.

Formally this "Programmatic Database" is a set of identifiers *challenge*$\times$*subject*, say ES verifying a list of values, ie:
$(ES_{m,n}, \{p_{m,1}, .., p_{m,i}, .., p_{m,I},$
$v_{(n,m,1)}, v_{(n,m,2)}, ., v_{(n,m,j)}, .., v_{(n,m,J)})$.

Such a data base is a natural candidate to be analyzed by Galois Lattice theory (say also: Formal Concept Analysis) (Ganter and Wille 1999) thanks to which a relevant aggregation process can be defined.

A large body of literature exists on the "association rules" based on Galois Lattices, and a generic approach has been defined in (Chaudron, Maille, and Boyer 2003). This theory allows to generate a Rule data base S^* containing all the rules of the template: *(rule identifier, confidence, portance, conclusion, list of premises)*. The confidence c is by definition the proportion of *challenge*$\times$*subject* that verify the rule among those who verify all the premises in the "Programmatic Database" (conversely: $1 - c$ is the proportion of counter-examples). The portance is the global proportion among all the database (thus $0 < p < c < 1$). Many algorithms exist so as to compute extensively these rules (and generally considering only those for which: $0.5 < c$).

The "association rules" approach is known to be the more generic non-supervised rule induction method.

Intuitively a Galois lattice expresses in an algebraic fashion the ability to compose compatible sub-goals in order to define a constructive orientation for the SOP. That is, the modelling task leads to provide a separated representation of the SOP consistent dynamic into relevant sites connected with association rules and afterwards to provide amalgamation operators in order to ensure the global consistency of the living entity to represent.

Autopoïesis and Self-Orientation

Back to a conceptual level, so-called autopoïetic systems distinguish themselves from their environment and produce themselves their own be and their own identity: autopoïetic comes from the Greek "auto" (self) and "poien" (to produce). Autopoïesis is a closed process in the sense where it specifies itself. By its organization an autopoïetic system is able to regenerate itself, to make its own survival, in fact to behave as being "alive". In (Varela 1995) Varela states simply: "Organization of alive systems *is* autopoïesis".

Autopoïesis implies that the evolution and all the transformations of the system are subordinated to the preservation of its organization and such a preservation of its unity and of its identity is based on the capability to self-orient the actions.

Thus S^* is by definition the SOP which represents the autopoïesis by aggregation (to the c and p thresholds) of the scientists $S = \{S_n\} n \in [1, N]$ responding to the societal challenges $E = \{E_m\}, m \in [1, M]$. This model of autopoïesis by aggregation was programmed in constrained logic programming language and applied to a research lab of Onera (70 persons), during up to ten years, thus proving the feasibility of such an approach. Nevertheless, no formal model of the semantic aspects of the Self-Orientation processes is yet defined in this first step. Thus in the sequel, a second model dedicated to a general algebraic description of any self-orientation process is detailed.

The problem is as follows; any SOP lives according to an inherited information system and this information system holds according to some compact support. Therefore one has to assume that some folding process is to be performed in order to provide that support.

One can define autopoïesis according to the following equivalence : production is ensured by an autonomous system and conversely, if the latter doesn't hold, no production is to be expected.

In terms of algebras, it consists to express autonomous consistency according to some involutive operator (Barr 1979; Blute 1993); that is autonomous systems are self-dual. However, in order to compute a consistent behaviour, one has to embed any autonomous entity inside a global structure called the environment Ω; if the environment is cultural, then it is a social organization.

If one expresses the behaviour of any SOP according to operators and if the set of operators is the smallest non-trivial possible, then either a SOP *sustains* the social organization in which it lives (according to the positive polarity $\uparrow$) or it *denies* it (according to the negative polarity $\downarrow$).

The formal theory of self-orientation consists of algebraic tools required in order to compute a consistent self-orientation merging both conservation laws $\uparrow$ and the chaos generated by the entropy of the system $\downarrow$ according to autopoïesis. It turns out that the expected algebra in order to express the 2-element set $\{\uparrow, \downarrow\}$ is based on a non-commutative composition of irreducible components, the rising edges $\nearrow$ and the falling edges $\searrow$.

Recall that falling edges are not productive and must be handled in the computation as alternative self-orientations *incompatible* with rising edges. Managing exclusion rules between rising edges and falling edges inside a sequence of edges $\nearrow\searrow \dots \nearrow\searrow$ leads in general to endless loops. The main assumption in this paper is to consider that autopoïesis performed by a SOP leads to compute a *continuous* single transition as a quantum step *inside* a double consistent box $\boxdot$. What we're actually looking for is a double box-algebra, i.e. a simplicial complex[1] able to host inside its convex en-

[1] A simplicial complex is a topological space, constructed by "gluing together" points, line segments, triangles, and their n-dimensional counterparts.

velope the set of merging edges $\rtimes$ in order to compute a single transition.

The formal Theory of Self-Orientation

The main issue of this second part can be depicted as follows: assume that candidate self-orientations form a set S or a vector space V or more generally an algebraic structure A (e.g. a partially ordered set P , a monoid M, a semigroup $S, \ldots$); if the computation of a self-orientation refers to a motivated choice among possible orientations inside an algebraic structure, one should be able in this structure to **separate** "*good*" and "*bad*" self-orientations. That is, consistency is required and as such needs to be defined. Take a *self-orientation process* S; a "*good*" valuation is called a S-**model** and is noted $\mathsf{S}^\top$ and a "*bad*" valuation is called S-**counter-model** and is noted $\mathsf{S}^\perp$.

Therefore, *consistent* self-orientations can be called *formal actions* provided that the algebraic structure must agree with the *separable property* which in terms of polynomial algebra corresponds to the *reducible property*. Intuitively, if a SOP is able to compute the algebraic roots of its self-orientation vector, then it gets consistency.

In order to understand the dynamic of autonomous self-orientation models, one can be interested in behavioral models for *cognitive entities*. Cognition is a collection of processes including attention, memory, producing and understanding language, solving problems, and making decisions and is studied in various disciplines such as psychology, philosophy, linguistics, science and computer science. Cognition is also a faculty for the processing of information, applying knowledge, and changing preferences.

The two associated terms *knowledge* and *change* mean to use knowledge as some stable information devoted to allow a self-orientation process as a cognitive entity to evolve safely inside dynamic environments. That means that *action* is necessary associated to knowledge and knowledge is required in order to ensure some *space/time stability*.

Assume that the central issue of a SOP is actually : *how to organize the orientation of actions, goals and plans* ? The algebraic apparatus should help us in order to tackle the separability issue as : *what is consistency in that framework ?*

In order to define what a self-orientation and what an action can be, one can select common widely used mathematical tools; in particular, one can start from 17th century's analytic geometry, matrices, systems of linear equations and Euclidean vectors and one can retain the 19th century's definition of a vector space able to extend classical geometric ideas like lines, planes and their higher-dimensional analogs. That is, at first sight, an *orientation* should be efficiently represented as an angle valuation θ characterizing a *vector* ξ, an element of a vector space V. That is, the set of possible self-orientations form an *operator algebra* A acting on a k-vector space V where k is a field[2]; an action a acts on vectors as a k-linear operator on V and the set of actions is a k-matrix algebra A.

[2] a field k is a triple $(k, +, .)$ where k is an abelian group under $+$ and $k - \{0\}$ acts multiplicatively on k

If the computation of a self-orientation refers to a motivated choice among possible self-orientations inside the vector space V, then a S-**model** is a subset $A^\top \subset A$ and a S-**counter-model** is a subset $A^\perp \subset A$. For instance, the "*good*" attribute means that elements of $A^\top$ are invertible matrices whereas the "*bad*" attribute means that elements of $A^\perp$ are singular matrices[3]. But in that case, inconsistency is everywhere and must be managed according to a combinatoric decomposition of A using a finite set of multiplicative subsets of A which correspond to A-ideals. Thanks to the Stone representation theorem (Johnstone 1982), A becomes *a boolean lattice* if A-ideals are expressed according to irreducible maximal-ideals which are in one-to-one correspondence thanks to De Morgan laws with ultrafilters.

If a SOP S would have a *complete* representation of the environment Ω in which it evolves, for instance as a k-vector space V, one can use standard convergent algorithms; but the knowledge of S on its environment Ω is *partial* and in that case, there is a *random issue* preventing to ensure unconditional stability; therefore consistency is required and is that way, a discrete combinatoric process as for instance *deduction*. This random issue is assumed to be unavoidable and can be summarized by the following statement :

Statement 0.1 *Any SOP* S *must encode a inner* compact *representation of the environment* Ω *(a compact* Ω*-encoding) in which it evolves (therefore the* S*-representation of* Ω *is* partial *and a random issue holds).*

Consequently, autopoïesis is ensured by a compact Ω-encoding and the following statement holds :

Statement 0.2 *Self-orientation for a SOP* S *inside* Ω *is necessary consistent and the self-orientation structure must agree with the* separable property *in order to separate* S-**models** $\mathsf{S}^\top$ *and* S-**counter-models** $\mathsf{S}^\perp$.

One has to refine the compact representation as follows

Statement 0.3 *A SOP* S *has a definite consistent* ***inner*** *subspace* $\top$ *and a definite* ***outer*** *subspace* $\perp, \top \neq \perp$ *with the incompleteness axiom* $\top \cap \perp \neq \emptyset$ *due to the partial representation of* Ω.

Therefore a SOP must be taken as a dynamic system represented by an algebra A inside which information was compacted by some densification operator $\uparrow$ and with which some decision process can be performed according to some unfolding step $\downarrow$. The unfolding process generates entropy, i.e. some chaos in order to control and to transform in a constructive fashion $\nearrow$ the environment Ω according to autopoïesis. The adequate structure for this dual folding/unfolding process is a Hopf algebra.

Definition 0.1 *A vector space* A *is a* Hopf algebra *(Blute*

[3] A non invertible square matrix C is is called singular or degenerate; C is singular if and only if $det(C) = 0$.

1993) if it is equipped with morphisms of the following form:

$$\textit{bialgebra}\begin{cases}\textit{coalgebra} & \begin{cases}\Delta : A \to A \otimes H\\ \epsilon : A \to k\end{cases}\\ \\ \textit{antipode} & S : A \to A\\ \\ \textit{algebra} & \begin{cases}\mu : A \otimes A \to A\\ \eta : k \to H\end{cases}\end{cases}$$

The unfolding process is represented by the coalgebraic part (Δ, ϵ); resp. the coproduct and the counity are able to decompose some element $a \in A$ according to two operands $a_1, a_2 \in A^2$ such that $a = a_1.a_2$. This process is non deterministic and must be confirmed by the algebraic part (μ, η) resp. the product and the unity recomputing the product $a = a_1.a_2$ according to the antipode bridge S which is able to connect the coalgebraic and the algebraic components. If the computation converges, then a self-orientation is expected as a result.

Therefore the main issue can be stated as follows:

1. How to get a *faithful consistent* computation of a self-orientation for a given SOP S due to the incompleteness assumption ?

In order to answer that question, one has to look carefully at topological and algebraic structures tailored for deductive systems theory and proof-theory.

Algebraic characterization of operator algebras

Topological characterization of deductive systems

A first characterization of deductive systems can be described in terms of topology : for sake of *completeness* and *correctness*, the topology of deductive systems is Hausdorff[4], i.e. *discrete* and *separable*.

In the frame of *operator systems* or of *action systems*, consequences are as follows:

Statement 0.4 *Consider an action system $\mathcal{A}$, in which one can select instances of generic action patterns and one can order them along a time line indexed by a denumerable subset of relative integers $\mathbb{Z}$; if a deductive system holds in order to compute consistent action sequences,* infinite *action loops are* perfectly valid.

This corresponds to the following problem for action computation: there is not enough *reaction* generated by the affine $\mathbb{Z}$-time line of ordered actions sequences. Therefore, one can exploit the following topological assumption:

Statement 0.5 *A suitable Ω-encoding requires a compact representation which is not Hausdorff, i.e. with a certain loss of consistent information.*

Consequently:

[4]A topological space X is *Hausdorff* if for every $x, y \in X, x \neq y$, there are open neighbourhoods $\mathcal{O}_x \ni x, \mathcal{O}_y \ni y$ so that $\mathcal{O}_x \cap \mathcal{O}_y = \emptyset$

Statement 0.6 *A compact non Hausdorff Ω-encoding means that connectedness "$=$" and disconnectedness "$\neq$" are both valid, i.e. components are* inseparable.

Therefore the *separability problem* is the central issue for the computation of operator sequences. Assume that such a computation is equivalent to compute an action plan Π; the inner consistency expressed by the *separable* condition cannot be ensured at start-up but is actually *the aim of the computation* for the whole plan Π.

In order to reinforce the relevance of algebraic characterization, one can propose to replace separability by *reducibility*. Consequently, the separability issue becomes the *reducibility issue*. Irreducible components can be depicted using the *simple property* for algebras. According to abstract algebra definitions, this corresponds to identify properties of two-sided ideals[5] (i.e. special multiplicative subsets of the operator algebra A).

In a way, this corresponds intuitively to force any action sequence $\alpha_1, \alpha_2, \ldots,$ to be of length 1 (that is, $|\alpha_1, \alpha_2, \ldots, | = 1$) and to force possible reactions between elements of the operator algebra by maximizing operator interactions. The simple assumption is very restrictive and prevents classical consistency representation to be used. Recall that the required structure in order to encode both knowledge and entropy as *continuous non-commutative simple morphisms* is the double box $\boxdot$.

Simple algebras

In mathematics, specifically in ring theory, an algebra A is *simple* if it contains **no** intermediate non-trivial two-sided ideals and the set $\{ab | a \in A, b \in A\} \neq \{0\}$ (this condition ensures that the algebra has a minimal nonzero left ideal). Note that this corresponds to the assumption "*no action sequence is of length less than 1 and greater than 1*".

An immediate example of simple algebras are *division algebras*, where every element has a multiplicative inverse, for instance, the real algebra of quaternions $\mathbb{H}$[6]. The physical interpretation of a division ring D in the frame of action sequences is as follows : every action α has an inverse α^{-1}; this would be the more powerful representation of an action system but this representation is to be altered into something weaker : the algebra of $n \times n$ matrices with entries in a division ring D which is a simple algebra. In fact, this charac-

[5]For an arbitrary ring $(R, +, .)$, let $(R, +)$ be the underlying additive group. A subset I is called a two-sided ideal (or simply an ideal) of R if it is an additive subgroup of R that "absorbs multiplication by elements of R". Formally we mean that I is an ideal if it satisfies the following conditions: (i) $(I, +)$ is a subgroup of $(R, +)$, (ii) $\forall x \in I, \forall r \in R : x.r \in I$, (iii) $\forall x \in I, \forall r \in R : r.x \in I$

[6]The quaternions $\mathbb{H}$ are a four-dimensional $\mathbb{R}$-algebra generated by the identity element 1 and the symbols i, j and k, so $\mathbb{H} = \{r_0 + r_1 i + r_2 j + r_3 k : r_0, \ldots, r_3 \in \mathbb{R}\}$. Quaternions are added together component by component, and quaternion multiplication is given by the quaternion relations $ij = -ji = k, jk = -kj = i, ki = -ik = j, i^2 = j^2 = k^2 = 1$ and the distributive law. The quaternion algebra is not commutative, though it does obey the associative law. The quaternions are a division algebra (an algebra with the property that $ab = 0$ implies that $a = 0$ or $b = 0$).

terizes all finite dimensional simple algebras up to isomorphism, i.e. any finite dimensional simple algebra is isomorphic to a matrix algebra over some division ring.

The simple qualifier in the frame of the planning calculus

In order to qualify computation issues for self-orientation as formal actions, one can examine carefully the definition of the so-called *AI Classical Planning Calculus*, a quite old fashion concept (Newell and Simon 1959), (Fikes and Nilsson 1971) imagined in the early days of Artificial Intelligence (1959) and check it out in which case the *simple* qualifier holds in that frame.

Definition 0.2 (AI Classical Planning Calculus) *The* AI Classical Planning Calculus *is a process used in Artificial Intelligence which is intended given :*

- *a world description,*
- *a goal to achieve,*
- *a set of actions which can be brought into play in this world,*

to find a **correct** *sequence of actions to apply in this world in order to* **transform** *the initial state Σ_i into a state Σ_f compatible with the goal to achieve.*

Definition 0.3 (State) *A* planning state Σ *is taken to be a* non void *conjunction of planning atoms,* $\mathsf{at(ball1,roomb)} \wedge \mathsf{at(ball2,roomb)} \wedge \mathsf{free(left)} \wedge \ldots$*; the set of planning states $\mathcal{S}_\wedge$ is the free subalgebra of propositional conjunctive formulæ $\mathcal{F}_\wedge$ defined wrt to the domain $\mathbf{D}$ minus the empty formula $\emptyset$. A* planning transition *is an element of $\mathcal{S}_\wedge \times \mathcal{S}_\wedge$.*

It is crucial to note that one cannot identify a planning state using the empty formula $\emptyset$!!! Therefore:

Statement 0.7 *Unlike deduction, the planning calculus agrees with a inner differential feature; it is equivalent to assume that the deduction theorem*

$$\overbrace{A}^{left} \vdash \overbrace{B}^{right} \cong \overbrace{\quad}^{left} \vdash \overbrace{A \supset B}^{right}$$

no longer holds. In effect, deduction requires to void the left hand side of $\vdash$ (or to set the assumptions equal to $\emptyset$) in order to prove consistency.

Definition 0.4 (Heyting's paradigm) *The* HEYTING*'s paradigm* — a proof of $f : A \supset B$ (A "implies" B) is a function which associates a proof of B to every proof of A *— allows deductions to be representable arrows in a* discrete *category. The consequence relation for the intuitionnist deduction is* functional.

From the previous definition, one can infer

Statement 0.8 *The planning arrow $f : \Sigma_i \to \Sigma_f$ doesn't verify a consistent partial order relation $\subseteq$ and cannot be expressed as an Σ_f-algebraic extension of the consistency of the initial state Σ_i as in the Heyting's paradigm.*

Finally

Statement 0.9 *The planning arrow $f : \Sigma_i \to \Sigma_f$ is not a causal law whereas a causal representation holds in the neighborhood of Σ_i due to the consistency of the initial state.*

If the last remark is ignored, then one can propose endless diverging algorithms whose main useless activity is to manage mutex (mutual exclusions), i.e. to remove unwanted pairs ($\nearrow, \searrow$); we claim that a SOP doesn't produce any interesting result that way. Mutex are computed because of the stability of the initial state Σ_i which turns out to be the support of the consistent environment Ω (recall that Ω is a social organization); therefore computing the final state Σ_f means to define autopoïesis as the *annihilator* of the social organization. Therefore the SOP consistency can be defined as follows :

How to represent the annihilator applying on Ω
as a constructive transformation ?

This constructive representation requires is to condense pure action (simple groups) according two kinds of *orthogonal arrows* : the "*state creator*" $\diagdown\!\!\nearrow$ and the "*state annihilator*" $\diagup\!\!\searrow$ and to apply them both on the initial state Σ_i and on the final state Σ_f.

The simple qualifier and the quantum step

Since the planning arrow f is not a causal law, one should define what sort of algebraic property holds in the neighborhood of the final state Σ_f. From the conjunctive expression of the final state,

$$\Sigma_f = \mathsf{at(ball1,roomb)} \wedge \mathsf{at(ball2,roomb)} \wedge \mathsf{free(left)} \wedge \ \ldots$$

one can compute the self of local orientations

$$\mathcal{O} = \{\nearrow \mathsf{at(ball1,roomb)}, \nearrow \mathsf{at(ball2,roomb)}, \nearrow \mathsf{free(left)}, \ldots\}$$

as the *bundle* of proofs allowing the set of goals to be activated locally (i.e. each activation function for given positive literal ends with a rising edge in the neighborhood of the final state Σ_f).

The *simple* assumption means that these activation functions between two boundaries,

1. the initial state Σ_i as a left boundary $\triangleright$,
2. the final state Σ_f as a right boundary $\triangleleft$.

are *irreducible as polynomials*.

In other words, it exists a *single quantum step* $\to$ from $\triangleright$ to $\triangleleft$ which does not agree with causal laws but which is nevertheless consistent. This single quantum step (Majid 2000) requires a polynomial functional space of activation functions, those with a rising edge in the neighborhood of the final state Σ_f.

Reactive representation of consistency and bi-orthogonality

One can force the completion of the triple ($\triangleright, \to, \triangleleft$) by several ways; one can assume that it may exist *something before* $\triangleright$ and *something after* $\triangleleft$ (in other words, one has to build a 5-sequence whose first component is a "*state creator*" $\diagdown\!\!\nearrow$ and whose last component is a "*state annihilator*" $\diagup\!\!\searrow$).

One obtains a 5-tuple $(\to_{-1}, \triangleright, \to_0, \triangleleft, \to_{+1})$ on which one can restore *involutive* symmetry by equalizing the first component $\to_{-1}$ and the fifth one $\to_{+1}$ which corresponds to the torsion condition

$$-1 \cong +1 \tag{1}$$

One obtains a quaternary structure $(\rightarrow_{-1}, \rhd, \rightarrow_0, \lhd)$ where the two boundaries are assumed to be dense states whereas the two arrows aren't and one can define reactive consistency in some looping algebraic structure **modulo** 4 which cannot be deductive. Assume that $\rightarrow_{-1}$ is a *fork*; then it has two *codomains* and therefore one can propose another diagram for the simple quantum step assumption

$$\begin{array}{ccc} & & \lhd \\ {}_{-1} & \nearrow & \\ & & \uparrow{\scriptstyle 0} \\ {}_{-1} & \searrow & \\ & & \rhd \end{array} \tag{2}$$

which looks like a Larmor precession[7] (the fork -1 forms a cone whose base is a circle containing $\lhd$ and $\rhd$ and for which 0 is a diameter.

Self-orientation and algebraic characterization of the initial state Σ_i

One can use the statement 0.3 page 3 in order to define a representation of self-orientation according to the initial state Σ_i and this kind of quaternary box. One should add another crucial assumption:

Statement 0.10 *The Ω-encoding is ensured by a compactification operator condensing knowledge on a finite set of vertexes on which one can define a maximal clique structure (a hyper connected graph); this kind of information is a phase structure (i.e. a representation of periodic functions wrt to a orthogonal base of characters, the harmonics of the function).*

From the previous statement and statement 0.9 page 5 , one can infer the following assumptions:

1. Σ_i is causal
2. Σ_i is a phase structure
3. Σ_i is a stabilizer ([8])
4. Σ_i is an idempotent operator **1**
5. Σ_i is a knowledge representation

Therefore one can assimilate the initial state Σ_i to inner consistency $\top$ and one can generate *reflections* in the quaternary box by facing against the initial state and the final state as follows:

1. Σ_f is *anti*-causal
2. Σ_f is a *anti*-phase
3. Σ_f is a *anti*-stabilizer (i.e. an orbit [9])
4. Σ_f is an *anti*-idempotent (i.e. a nilpotent operator) **0**
5. Σ_f is a *anti*-knowledge representation

However, one cannot assimilate the final state Σ_f to outer consistency $\perp$ because Σ_f is not outside due to the torsion condition 1 page 5 . The best constructive representation of the *anti*-functor would be an *inverse* operator able to encode Σ_f in the same Σ_i-algebra; in that case the operator algebra is a unified field F; but this is not possible because actions are not invertible;

The consistent interpretation is as follows:

Statement 0.11 *Due to the partial representation assumption, any non trivial consistent theory Θ has two roots (i.e. is ambivalent): a low-level root $\rhd$ and a high-level root $\lhd$; the consistency of Θ always stabilizes in the neighborhood of the* idempotent *low-level root $\rhd$ in which $\top \rightarrow \top$ holds and which is an altered form of consistency. A quantum step is required in order to reach the local form of plain consistency in the neighborhood of $\lhd$ and to stabilize around this high-level root.*

One can interpret that way the meaning of the simple assumption: thanks to the simple assumption one can separate in an orthogonal fashion stabilizer subgroups and orbit subgroups; that is the planning algorithm looks like a three step consistency process:

1. perform a stabilizer in the neighborhood of the low-level root $\rhd$
2. perform an quantum orbit $\rightarrow$ from $\lhd$
3. perform a stabilizer in the neighborhood of the high-level root $\lhd$

The splitting operator

The purpose is to *split* consistency in two parts $\top \rightarrow \top \times \top$ and to work in that weak consistent space on which we hope to define a tensor product in order to encode the direct product in a functional algebra and to refine the previous arrow as follows $\top \rightarrow \top \otimes \top$ in the frame of strong *involutive* structures : *Hopf Algebras* defined in definition 0.1 page 3 . The quantum step corresponds to the following twist :

$$(\top_0, \top_1, \perp_0, \perp_1) \rightarrow (\top_0, \perp_0, \top_1, \perp_1)$$

Intuitively the initial state is the idempotent part to remove $\top_1$ and to be replaced by the final state $\perp_0$ by a switch between the second and the third component; $\top_0$ remains consistent while $\perp_1$ remains inconsistent. It is interesting to note that one can represent duality $(\top, \perp)$ by something like $\mathbb{C}$-algebras; the dense part $\top$ is the $\mathbb{R}$-part whereas the nowhere dense part $\perp$ is the *imaginary* part. Therefore the quaternion algebra $\mathbb{H}$, as a simple four dimensional $\mathbb{R}$-algebra can be a starting point in order to encode quaternary boxes.

$$((\top_0, \top_1), \overline{(\top_0, \top_1)}) \rightarrow (\top_0, \overline{\top_0}, \overline{\top_1, \overline{\top_1}})$$

In the frame of self-orientation, the problem is to motivate an orientation for a SOP S and consequently a *change* whereas the Ω-encoding is currently stable (i.e. where the idempotent acts as a stabilizer). In other words, the SOP

[7] In physics, Larmor precession (named after Joseph Larmor) is the precession of the magnetic moments of electrons, muons, all leptons with magnetic moments, which are quantum effects of particle spin, atomic nuclei, and atoms about an external magnetic field

[8] For every $x \in X$, we define the stabilizer subgroup of x (also called the isotropy group or little group) as the set of all elements in G that fix x : $\{g \in G | g \circ x = x\}$

[9] Consider a group G acting on a set X. The orbit of a point $x \in X$ is the set of elements of X to which x can be moved by the elements of G. The orbit of x is denoted by $Gx = \{g \circ x | g \in G\}$

S performs the orthogonal of a stabilizer : *the orbit operator* in order to locally curve (and therefore to locally control) the environment Ω in which it evolves. Assume that biorthogonality means both *faithful* (not altered) and *stable*; in the neighborhood of the initial state, a SOP S has not the biorthogonal condition, and computes the self-orientation in order to gain biorthogonality and to perform a consistent quantum step extension (the consistency extension by biorthogonality is not causal). Biorthogonality corresponds to some consistent result : the action plan to compute; that is biorthogonality characterizes some truth valuation; in particular, in deductive system theory, a type is bi-orthogonal (one can define for each subset a its orthogonal $a^{\perp} = a \multimap \perp$ and a fact a is any subset equal to its biorthogonal $a \cong a^{\perp\perp}$).

Bi-representations and planning consistency

In order to express the action planning consistency in the frame of representation theory, one can use Kelvin-Planck's second principle of thermodynamic : "*let a monotherm cycle; it cannot provide any motion* " which becomes that way "*let a unique representation system; it cannot provide any interesting computation*". That is, a double representation arrow (i.e. a biarrow) $(\rightarrow, \leftarrow)$ holds and one can propose a reflector to term/type duality; the expression $t \in T$ which means that the term t is of term T can be reversed as $t \ni T$. One can define *idempotent* consistency $\top_{\mathsf{e}} \in \Omega$ as a maximal clique on which one can define the spectrum of periodic functions and *nilpotent* semantic consistency $\perp_{\mathsf{e}} \ni \Omega$ as the anti-support of this maximal clique.

When one examines carefully deductive systems and in particular the role of inference rules as geometric operators on a functional set of proofs, one can note that consistency requires strong duality results to hold in order to encode negation. An interesting framework is to connect in a a *bialgebra* (a same full dual structure), an algebraic part and a coalgebraic wrt involutive rules. This involutive implementation must be valid in the frame of planning calculus as well except that the consistency support is not binary but quaternary in order to be implemented using quaternary boxes.

Reflections in algebraic characterization of deductive systems

According to Gerhard GENTZEN's formal model (Szabo 1969), the support of a deductive system is set of **involutive** inference rules; the carrier of consistency is a multi-functional sequent f

$$f : \Gamma \vdash \Theta \qquad \cong \qquad f : \overbrace{A_1, \ldots, A_n}^{\wedge} \vdash \overbrace{B_1, \ldots, B_m}^{\vee}$$

where formulæ occurring in Γ and Θ travel across the turnstyle symbol $\vdash$ as for instance the major *involutive* inference rule for the negation symbol $\neg$

$$\frac{\ldots_1 \vdash A, \ldots_2}{\ldots_1, \neg A \vdash \ldots_2}\ \neg_L \qquad\qquad \frac{\ldots_1, A \vdash \ldots_2}{\ldots_1 \vdash \neg A, \ldots_2}\ \neg_R \qquad (3)$$

which can be geometrically expressed as a *reflector* by the provable statement $(\neg\, \neg\, A) \supset A$ or equivalently "(*not not* A) *implies* A" which leads to the "excluded middle law" $(\neg\, \neg\, A) \cong A$ provided that one adds the statement $A \supset (\neg\, \neg\, A)$ which is true only in classical logic.

Proofs as sequents $f, g, \ldots$ are representable "*arrows*" in a **discrete category** (Lambek 1972) containing a "hole" $\perp$ (a hyperbolic subspace or a zero divisor) guarded by the consistency axiom $A \wedge (\neg\, A) \supset \perp$ (which means A *and* (*not* A) *implies absurd*) or equivalently $\top \cap \perp = \emptyset$.

Any deductive sequent system is defined according to the *identity group* : the **initial** rule and the **cut-rule**;

$$\frac{}{1_A : \ldots_1, A \vdash A, \ldots_2}\ \textbf{initial}$$

$$\frac{f : \overbrace{\ldots_1 \vdash A, \ldots_3}^{(a)} \qquad g : \overbrace{\ldots_2, A \vdash \ldots_4}^{(b)}}{gf : \underbrace{\ldots_1, \ldots_2 \vdash \ldots_3, \ldots_4}_{(c)}}\ \textbf{cut}$$

The **initial** rule stands for the identity axiom; if all *leaves* of the proof tree are initial rules, then a *correctness* result holds.

The **cut-rule** stands for free insertion of intermediate formulæ and free splitting of the main proof (c) occurring in the denominator in two *separate* subproofs, the numerator subproof on the *left* (a) whose conclusion is the principal formula A and the numerator subproof (b) on the *right* whose assumption is the same the principal formula; consequently, according to the cut-rule a *completeness* result holds.

Nevertheless, this rule is paradoxical because the **cut-rule** generates short-cuts in a proof and consistency of the deductive systems requires to replace the cut-rule by logical rules. For every connective of the logical group $\supset, \neg, \wedge, \vee$, inference rules are pairwise adjoint endo-functors elimination/introduction as for instance, one of them, the sequent representation of the deduction theorem (i.e. $\supset_L$ and $\supset_R$ rules)

$$\frac{\ldots_1 \vdash \ldots_3, A \qquad B, \ldots_2 \vdash \ldots_4}{\ldots_1, \ldots_2, A \supset B \vdash \ldots_3, \ldots_4}\ \supset_L$$

$$\frac{\ldots_1, A \vdash B, \ldots_2}{\ldots_1, \vdash A \supset B, \ldots_2}\ \supset_R$$

For instance a proof of **modus ponens** (from A and $A \supset B$, one can prove B) can be represented as follows:

$$\frac{\dfrac{}{\ldots_1, A \vdash A, \ldots_3}\ \textbf{initial} \qquad \dfrac{}{\ldots_2, B \vdash B, \ldots_4}\ \textbf{initial}}{\ldots_1, \ldots_2, A, A \supset B \vdash B, \ldots_3, \ldots_4}\ \supset_L$$

Cut-elimination is expressed by the following result :

Theorem 0.1 *Gerhard* GENTZEN*'s Hauptsatz : if a sequent system has a cut-elimination result, each sequent has the* sub-formula *property (the proof data required in order to decide about the validity of the sequent is self-contained).*

Statement 0.12 *In a sequent system with the cut-elimination property, the knowledge cannot be partial.*

In the sequent calculus, one can relax idempotency $A \wedge A \neq A$ (idempotency is mandatory for classical and intuitionist logic and requires structural rules : contraction, weakening, exchange); then, one can get B, using modus ponens, from A and $A \supset B$, but in that case A *no longer holds* when modus ponens is performed. The conjunction $\otimes$ replaces $\wedge$, the "linear" implication $A \multimap B$ replaces $A \supset B$, negation $A^{\perp} \cong A \multimap \perp$ is orthogonal. That is, according to tensored involutive structures, one can select relevant structures for consistency.

A completeness result holds thanks to involutive rules: a valid coalgebraic proof tree is an element of a cut-free *cocommutative* comodule (i.e. a set of acyclic graphs whose leaves are separable occurrences of **initial** and nodes are inference rules); the cocommutativity condition for a Hopf algebra $\otimes H$ sets $S \circ S = \mathbf{id}$ (that is an involution between coalgebra and algebra able to encode negation). Proofs are representable arrows in linear cocommutative comodule category

In the frame of the planning calculus, the algebra (μ, η) is the compact *positive* component "$\oplus \top$" and the coalgebra (Δ, ϵ) is the unfolding *negative* component "$\ominus\perp$".

One can assume that the differential "$\ominus \perp \oplus \top$" is not commutative and corresponds to the simple quantum step with respect to the following law :
"In order to **gain** *something, I should* **remove** *something"*.

Therefore, one should encode the simple property in the dual frame of Hopf algebras. Planning operators are naturally definable as coalgebraic items whereas activation polynomial functions for positive literals are definable as algebraic items. In the bialgebra context, one should define an algorithm according to the quaternary box structure $H \otimes H \otimes H \otimes H$ mentioned in the fork diagram 2 page 6 .

Central simple algebras as a host algebra for the planning algorithm

A central simple algebra A (sometimes called a Brauer algebra) is a simple finite dimensional algebra over a field F whose *center* $Z(A)$ (the set of commuting elements) is F; intuitively, for an algebra A the "center" $Z(A)$ is real part of the faithful representation whereas the complement $A - Z(A)$ is the simple (maximally non-commutative) quantum step. They are finite dimensional associative algebras over a field F (or over a local ring R) with center F (or R) and without intermediate non-trivial two sided ideals (*twisted* "*non-separable*" forms of $n \times n$ square matrix algebras). One can define a sketch of the planning algorithm: compute planning operators in order to remove the twist and to finally obtain a $n \times n$ square matrix algebra which is *separable*.

Conclusion

If self-orientations are definable in the frame of consistent representations, one can propose relevant algorithms provided one is able to enrich the algebraic structure of states and actions in order to use enough reflection to encode and decode dynamic information.

Deductive system theory is undoubtedly a suitable framework although it should be used in that frame for totally different purposes.

References

Barr, M. 1979. $\star$-autonomous categories. volume **752** of *Lectures Notes in Mathematics*. Springer-Verlag.

Blute, R. 1993. Linear topology, hopf algebras and $*$-autonomous categories. *Research Report*.

Chaudron, L.; Erceau, J.; and Trousse, B. 1993. Cooperative decisions and actions in multi-agent worlds. In *International Conference on Systems, Man and Cybernetics*. Le Touquet, France: IEEE.

Chaudron, L.; Maille, N.; and Boyer, M. 2003. The CUBE lattice model and its applications. *Applied Artificial Intelligence* 3(17).

Erceau, J.; Benhamou, Ph.; and Zeghal, K. 1993. Collective intelligence patterns from antagonist dynamic. In *2nd European Systemic Congress*.

Fikes, R.; and Nilsson, N. J. 1971. STRIPS : A new approach to the application of theorem proving to problem solving. *Artificial Intelligence 2* 189–208.

Ganter, B.; and Wille, R. 1999. *Formal Concept Analysis, Mathematical foundations*. Springer–Verlag. ISBN 3-540-62771-5.

Johnstone, P. 1982. *Stone spaces*, volume **3**. Cambridge studies in advanced mathematics, Cambridge.

Lambek, J. 1972. *Deductive systems and categories III*. volume **274** of *Lectures Notes in Mathematics*. Springer-Verlag. 57–82.

Lupasco, S. 1947. *Logic and contradiction*. in french.

Majid, S. 2000. *Foundations of Quantum Group Theory*. Cambridge University Press.

Newell, A., and Simon H. A. 1959. Report on a general problem-solving program. 256–264.

Szabo, M. 1969. *The collected works of Gerhard Gentzen*. North-Holland, Amsterdam.

Varela, F. H.; and Uribe, R. 1974. Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems* 5:187–196.

Varela, F. 1995. *The Artificial Life route to Artificial Intelligence: Building Embodied, Situated Agents*. Lawrence Erlbaum Assoc. chapter The re-enchantment of the concrete, 11–20.