Open access • Journal Article • DOI:10.4304/JCP.2.5.54-66

Aristotle's Square Revisited to Frame Discovery Science — Source link

Mohammad Afshar, Christopher Dartnell, Dominique Luzeaux, Jean Sallantin ...+1 more authors

Published on: 07 Jan 2007 - Journal of Computers

Topics: Discovery science, Business process discovery, Paraconsistent logic and Scientific method

Aristotle's Square Revisited to Frame Discovery Science

Jean Sallantin, Mohammad Afshar, Dominique Luzeaux, Christopher Dartnell, Yannick Tognetti

Aristotle's Square Revisited to Frame Discovery Science

Mohammad Afshar, Christopher Dartnell, Dominique Luzeaux, Jean Sallantin, Yannick Tognetti
Ariana Pharmaceuticals, Pasteur Biotop, 28 rue Dr Roux Paris 75724 - France
Email: {m.afshar, y.tognetti}@arianapharma.com
Euriware, 44 Rue des Vindits, 50130 Cherbourg-Octeville - France
Email: christopher.dartnell@gmail.com
Centre Technique des Systèmes d'Information, 24 av. Prieur de la Côte d'Or 94117 Arcueil Cedex France
Email: dominique.luzeaux@polytechnique.org
LIRMM, CNRS UM2, 161 rue Ada, 34392 Montpellier Cedex 5 - France
Email: js@lirmm.fr

***Abstract*— The paper attempts to give a formal framework to capture the entire process of scientific discovery including hypothesis formation, reasoning, identifying contradictions, peer reviewing, reformulating and so on. Data mining can be seen as one step in this complex process of interactive learning of an empirical theory This paper uses the terminology from paraconsistent logic and paracomplete logic that extends Aristotle square in a hypercube of oppositions which defines or substantiates any step of the discovery process.**

The central formal notions are validated on a mathematical scientific discovery game, and an industrial application in the field of Drug Discovery illustrates how the presented framework combines different learning processes to predict pharmaco-kinetic properties (ADME-T) and adverse side effects of therapeutic drug molecules.

***Index Terms*— Machine Learning, Scientific Method, Logical Reasoning Framework, Aristotle's Square of Oppositions**

I. Introduction

Our objective is to propose a logical framework to assist scientists in supervising the entire process of theories formation while studying a phenomenon, as it is done by chemists studying properties of a new molecule, such as absorption or toxicity.

The first piece of the framework's logical structure is known since medieval logic as Aristotle's square [1], and has appeared in logical texts ever since: Aristotle defines syllogistic reasoning by differentiating universal and particular statements, and linking them by deduction and negation. The detection of contradictions then occupies the diagonal of such a square and can warn the scientist that a revision of knowledge is necessary (Figure 1).

The framework is formulated in a logical form such that contradictions can occur: in logic, a contradiction is produced by the conjunction of a formula and its negation. Each logic defines its own negation by an axiomatic schema expressing relations between universal and existential modalities, as "'Necessary"' (positive universal) and "'Contingent"' (negative existential). In the case of scientific discovery, events can be declared contingent when they sometimes occur by chance, or can be declared necessary if they have to exist. A theory then presents a contradiction when it allows to deduce, for instance, that a contingent event occurs by necessity.

Since Platon and Aristotle, logicians created different logics by selecting axioms that impose for instance that *a necessary action is a possible action*. But all these logics exhibit paradoxes provoked by their axioms, which often lead to their triviality. However, we admit that contradictions occur during a causal reasoning, and they are used in this framework to alert about problems concerning the consistency or the completeness of the ongoing theory.

Let us sketch such a formation process (a definition of unintuitive terms is provided in sections V-A and VI-A). Since Popper, it is admitted that a scientific theory must be refutable by experimentation and empirical data. A scientific experimentation, designed to enable empirical proof or refutation, requires the use of an accurate and efficient instrumentation to determine the existence of positive observations used to formulate postulates and conjectures. This experimentation design is a tool to ensure the progression of the ongoing theory by revealing facts. A fact is a piece of information (data) having an undeniable empirical evidence for scientists, and techniques such as data mining consist in the induction of a model from these facts. Such a model is built to match with positive information describing facts and to prevent the prediction of non observed facts. Predictions are made using a theory completed by a model, and the consideration of the overfitting and underfitting of these predictions informs the scientists on the bias conditions making it possible to decide correctly with the ongoing theory.

The logical framework presented in this paper is designed to take into account the two dynamics of scientific discovery [2] [3]. The first dynamic, which we refer to as the *personal dynamic*, embraces the supervision of a computer assistant by a scientist. This dynamic is centered on individual behaviors, and depends strongly on the research strategy of each individual and on its use of computer assistants. In fact, more and more learning techniques and data mining tools are used to find correlations in data and propose models to explain a studied phenomenon

and make predictions. We will see how the framework presented in this paper takes into account this personal dynamic by giving logical modalities to the statements occurring during the construction of a theory, and by placing the interpretation of contradictions in the heart of the interactive process leading to the construction of a model which can be discussed, justified, and proven in pure logic. The second dynamic we are concerned with, which we refer to as the *social dynamic*, comes from the collective behavior emerging from the social game during which scientists publish their theories and confront them to understand how models in a given domain match or influence models concerning other domains. In fact, scientists are experts in the different scientific fields involved in the understanding of a phenomenon, and we suppose that it is not always possible to merge all the scientific domains in a only one. In other words, scientists can all use their own models and data to compute and predict (personal dynamic). But by confronting their works, they may realize that a good solution for a given model can be a bad one for another model, and therefore point out the need to improve their understanding of the phenomenon. For instance, when dealing with drug discovery, a model of Absorption may not take into account the toxicity of the molecule. To enable this dynamic, the framework proposed in this paper takes into account the process of crisis, conflict and transactions which continuously rhythms Science and contributes to the formation of theories admitted by the society of scientists as being scientifically valid.

II. Outline of the paper

We illustrate in section III the recurrent problems posed by the personal and social dynamics, by describing the domain of Drug discovery and the prediction of ADME-T properties. The pharmaceutical industry is indeed confronted to a pressing need to analyze ever growing quantities of collected data and convert them into relevant decisions, using cheminformatics methods. Most often, reliable predictions are only possible on molecules very similar to the learning set, and these predictions use descriptions which are not easy to be translated in better molecular structures. This domain concentrates many challenges for inductive learning : the bias problem, the underfitting/overfitting problem, the constraint satisfaction problem, the multiparametric decision problem, the empirical testing and the interactive problem solving by scientists assisted by machine learning. As there are hundreds of descriptors used to describe the surface of the molecules and thousands related to the computation of properties for a single molecule, the risk of overfitting is permanent. When we try to limit the number of descriptors to reduce this overfitting, we create a risk of falling in the adverse case of underfitting. So the balance between completeness and consistency corresponds to the balance between underfitting and overfitting. As learned rules in drug design can conjugate more than three terms, the constraint satisfaction problem is divided in two domains: the domain where deciding with learned rules is easy but produce errors, and the domain where finding rules to decide is so difficult that the system learns by heart and can only decide for molecules which are very similar to the examples used during learning. Finally the multiparametric decision, required to deal with a distributed set of constraints that is not convex, often leads to an antagonism between optimized decisions. Furthermore, examples from disparate domains such as medical, chemical, legal, ... have different theoretical basis. Therefore, the regularities learned from these examples cannot be justified only by causal arguments.

After this introductory example, we present step by step the four levels of our formalism. Section IV describes Aristotle's square of oppositions, which defines the oppositions between universal and particular modalities, therefore enabling mathematical and logical reasoning on simple problems. In section V, this square is used to formalize the notions of postulate and conjecture, which are necessary to define a theory. The cubic structure presented in this section is obtained by linking the various squares together. section V-B offers an intuitive illustration of these notions. Section VI then presents a hypercubic structure obtained by linking the cube of oppositions to two new Aristotle's squares introduced to define the modalities of model, experimentation, prediction and facts. We emphasize the fact that the cube of oppositions lacked of temporality, and that this hypercubic structure links reasoning on facts and postulates to decision making and action, which are sufficient to define an agent. Finally, section VII establishes a link to deontic, autoepistemic, and defeasible modalities which are produced when agents using this hypercube to reason interact by implementing a conversational process of discovery and learning from each other. In conclusion we discuss how the objective to capture the entire process of scientific discovery is achieved starting from Aristotle's square to frame computer agents assisting humans during a problem resolution.

III. The problem of Drug discovery, ADME-T

Schematically, the pharmaceutical activity can be divided into three sectors: drug discovery (i.e. going from a target to a molecule that is ready to be tested in man), drug development (i.e. the proof of concept in man and the clinical trials) and finally the marketing and monitoring of the product.

It is widely accepted that out of one hundred drug discovery projects that are started, less than one would eventually reach the market ten to fifteen years later.

Despite over a decade of massive investment by the pharmaceutical industry into high throughput methods (Genomics, High Throughput Screening and combinatorial chemistry), efficient identification and optimization of potent and quality lead molecules is still the highest and riskiest hurdle in current drug discovery and development. The only clear outcome of high throughput methods has

been an unparalleled production of large quantities of data that need to be analyzed.

In order to reduce risks in the clinical stages of development, in a typical lead optimization process, 40 to 60 assays are run in parallel or in a cascade to evaluate the potential of each candidate molecule, its specificity, its good Absorption and Distribution, good Metabolism and Excretion profiles and limited Toxicity (ADME-T). In this multi-parametric space, identifying "quality" molecules which display desirable properties is a true challenge.

The use of computational tools (data mining, predictive modeling etc) has been seen as the potential solution to this dramatic inefficiency.

QSAR (Quantitative Structure Activity Relation) equations are standard examples in predictive modeling for drug discovery where an overall fitness score is developed as a weighted sum of numerous descriptors. In Docking, the score includes ligand internal energy, interaction energy and entropic considerations in the form of a weighted sum of terms [4]. Typically the score is developed empirically by analyzing a set of examples and deriving a weighted sum. The weights are fitted to the learning set and may not necessarily be relevant or precise for other complexes.

QSAR attempts to relate a numerical description of a molecular structure to a known biological activity. Large numbers of readily computable descriptors are available, in combination to sophisticated techniques that improve the initial linear regression analysis methods used in deriving QSAR equations (PCA, PLS, NN, GA, SVM etc). In general, QSAR equations relate one objective (such as activity for example) with a number of descriptors. QSAR equations are constructed by the combination of a number of weighted terms (descriptors).

These methods rely on the choices of (1) the descriptors for "generalization" and (2) the examples in the learning set to avoid overfitting.

An inadequate choice of either parameters will generally lead to useless models that do not generalize or are not interpretable. This is also true for initial ILP approaches [5]. In addition, search strategies can be compromised when confronted to non-convex solution fronts, i.e. when a solution "between" two valid solutions might be invalid. Furthermore scale invariance is not always true, i.e. even for a continuous property such as molecular weight, its use and therefore significance is distinct for different ranges (for example 200-600 range correspond to small molecules, a molecular weight greater than 2000 does not). This is to say that some relations are sensitive to scale. More generally, qualities can be converted into quantities (binning) but the reverse is not always true. This requires defining domains of validity for all parameters, in both the search and the objective spaces. In turn the notion of domain is linked to boundaries and hence allows characterization of paradoxical combinations or conflicts. Here, conflicts are real mutual exclusions rather than a competition between several continuous parameters.

All in all, it is fair to say that the current state of the art in cheminformatics is insufficient: "In general, reliable predictions are only possible for molecules similar to those in the training set" [6] hence undermining their predictive use, and "most models [...] use descriptors that are not easily understood by the chemist and not easy to translate into better molecular structures", and hence have little impact in drug discovery.

This discussion about the limits of QSAR shows how contradictions occur each time a numerical description of a molecular structure is related to biological activities. These contradictions are logical events and have to be framed in such a context by a logical framework placing them at the core of the user/assistant interaction in order to enable their understanding and to control the proof process of conjectures generated by learning from examples. Section IV recalls some definitions about Aristotle's square, the building brick of this logical framework. This article presents how this classical logical structure is extended to frame all the facets of scientific discovery illustrated by the ADME-T problem.

IV. Aristotle's square of oppositions

In this section, we recall what the classical square of oppositions is, and we provide a common sense interpretation to introduce the modalities which take sense in the context of theory formation and scientific discovery.

A. Classical square of oppositions

The doctrine of the square of opposition originated with Aristotle in the fourth century B.C. and has occurred in logic texts ever since. It connects various quantified propositions and their negations by introducing various notions of oppositions: contradiction, contrariety and subcontrariety (Figure 1).

Definition 1 (Contradiction): Contradiction for two terms is defined as the impossibility for them to be both true or both false at the same time.

Definition 2 (Contrariety): Contrariety is defined as the impossibility for two terms to be both true, but the possibility to be both false.

Definition 3 (Sub-contrariety): Sub-contrariety is defined as the impossibility for two terms to be both false, but the possibility for them to be both true.

According to these definitions, opposition is based on various degrees of truth difference. A last useful notion is sub-alternation between two terms, also better known as implication:

Definition 4 (Sub-alternation / Implication): Sub-alternation is defined as the impossibility of having the first term true without having also the second true.

The square of oppositions is represented by the following geometrical relations (Figure 1).

The column with A and I corresponds to affirmative propositions, while the column with E and O corresponds to negative propositions. The line with A and E corresponds to universal propositions, while the line with I and O corresponds to existential (also called particular) propositions. Several extensions have been proposed in order to

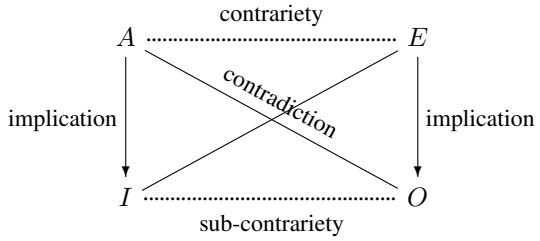

Figure 1. Aristotle's square of oppositions

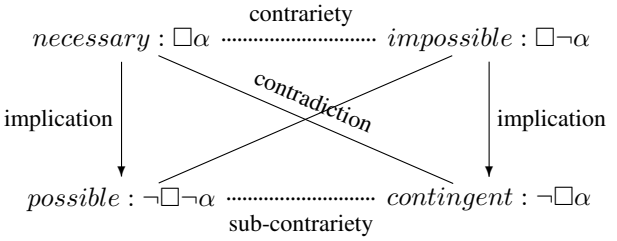

Figure 2. A modal interpretation of Aristotle's square of oppositions

palliate the logical drawbacks and develop the inference capabilities of the traditional Aristotelian square. Various modal decorations on the vertices can be found in [7].

To express our formalism, we need to define, starting from this square, a closed set of modalities to distinguish the logical status of the statements occurring during the process of theory formation.

B. An interpretation of Aristotle's square

It is very important to understand that the modalities are the key of the communication between the scientist and his computer assistant: they are used by the assistant to reason logically, and therefore must be interpretable easily by the human to enable a control and a supervision of the theory formation. Several interpretations may be possible, and the ones presented here were chosen because their definitions as found in the dictionary are related in the same way as the modalities they represent.

Definition 5 (Proof: $\Box\alpha$): A *proof* is a process which enables the verification of a computation's exactitude, or of the pertinence of problem's solution. To prove α is to establish with reasoning the truth of α, and $\Box\alpha$ denotes the fact that α has been demonstrated as being necessarily true (or necessary).

Definition 6 (Refutation: $\Box\neg\alpha$): A *refutation* is a process which enables to demonstrate the falsity of an affirmation by contrary proofs. $\Box\neg\alpha$ denotes the fact that α has been demonstrated as being impossible, or necessarily false.

Definition 7 (Contingent: $\neg\Box\alpha$): Contingent denotes the fact that α has not been proven. α may then occur, without any explanation or "by chance".

Definition 8 (Possible: $\neg\Box\neg\alpha$): Possible denotes the fact that the formula has not been refuted.

If we gave a probabilist interpretation of these four modalities, we would say that α has a high risk of occurring, whereas $\neg\Box\alpha$ will most likely not occur. Of course, $\Box\alpha$ has 100% chances of happening, and $\Box\neg\alpha$ has none.

Figure 2 represents Aristotle's square decorated with these modalities.

Aristotle's square is the building block of our logical structure. This formalism closes the set of modalities used to type statements treated during logical reasoning, instead of closing the world on which one has to reason. In the following sections, we explain why at least three of these squares are needed to formalize the formation of theories in the context of incomplete and inconsistent knowledge (section V), and five of them to take into account the scientific aspect of the discovery process (section VI).

V. The Cube of Oppositions

The process which supplies evidence for the validity, or for the invalidity, of certain inferences and conversions (of a proposition into its negative) is based on the square presented on Figure 2, which opposes Proof ($\Box$) and its negation ($\neg\Box$). Therefore the square of opposition appears as a geometrization of the inference process. But as we mentioned in the introduction, each logic defines its own negation, and this square actually corresponds to the definition of proof in classical logic, which could be sufficient to help solving problems in a closed and perfectly described world. But to deal with scientific discovery, learning, and theory formation, we need to take into account the inconsistency and the incompleteness of the knowledge used to reason, which justifies the use of paraconsistent logic (in a paraconsistent logic there can exist a proposition which is true and the negation of which is true, without implying the triviality of the theory, i.e. the truth of any proposition) and paracomplete logic (in a paracomplete logic there can exist a proposition which is false and the negation of which is false, without implying triviality of the theory, i.e. the truth of any proposition).

Extending the classical square by introducing conjunctions and disjunctions of the terms, and working on the geometrical aspects of the so formed hexagons as well as on its various modal decorations, [7] introduces new modalities with two other hexagons corresponding to a paraconsistent and to a paracomplete definition of negation (Figure 3).

Each of these hexagons shows how the bottom modality is opposed to the top modality. Looking at the three hexagons, we have thus the most general sub-alternation (or implication) relation between these various negative terms: $\Box\neg\alpha \longrightarrow \neg\alpha \longrightarrow \neg\Box\alpha$ [8]. This is no surprise, since these terms are known as expressing various kinds of negation in classical and modal proposition logics with the corresponding weakening relations: [9], [10] show that $\Box\neg$ is an intuitionistic paracomplete negation, and [11] shows that $\neg\Box$ is a paraconsistent negation.

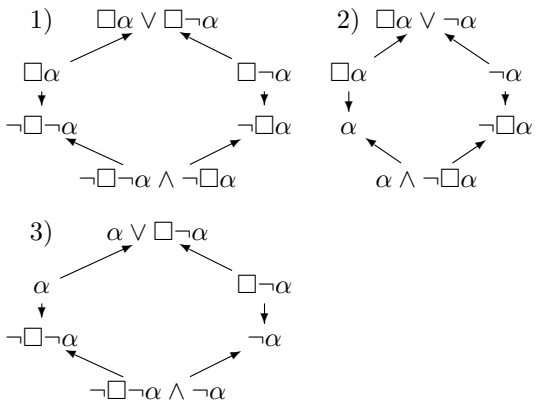

Figure 3. 1) "classical" negation, 2) "paraconsistent" negation , and "paracomplete" negation

A. Modalities attached to a theory formation

We derive from these negations the formal definitions of a Postulate ($\alpha \wedge \neg\Box\alpha$) and Conjecture ($\neg\alpha \wedge \neg\Box\neg\alpha$) (Figure 4), introduced to deal respectively with the inconsistency and the incompleteness of knowledge during the formation of a theory. An ongoing theory is then determined by a set of postulates and conjectures that should finally be instituted as being its principles and its theorems to be demonstrated.

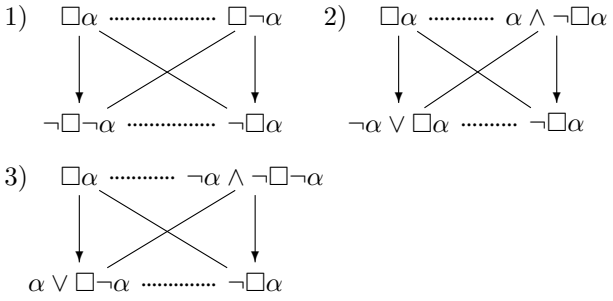

Figure 4. 1) classical definition of a refutation (as on Figure 2), 2) paraconsistent definition of a conjecture, and 3) paracomplete definition of a postulate

Definition 9 (Observed: α): α is often interpreted as "true" (and $\neg\alpha$ as "false"), but here, the interpretation of α as "is observed" (and $\neg\alpha$ as "not observed") is preferred. This notion of observation is refined in section VI in which observations are related to experimentation in a context of modeling: there is a glimpse of temporality under this notion, which should be also refined by the process of prediction and factualisation. In the cubic structure presented in the current section, the notion of observation "forgets", in a certain sense, all the meanings required to design the modeling actions of an agent.

We introduce the use of postulates to fix some limits to what is arguable or not, and conjectures to restrain the objectives, to fix some limits to what is provable or not (for example, one could state that the conjecture $P = NP$ is not to be proven)

Definition 10 (Postulate: $\alpha \wedge \neg\Box\alpha$): A *postulate* is a primary principle, indemonstrable or undemonstrated.

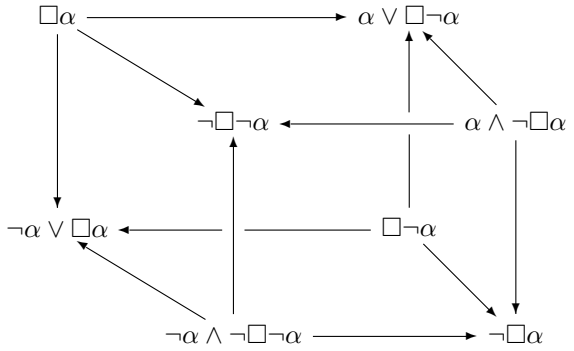

Figure 5. The cube of oppositions

Formally, a postulate is a statement supposed on the basis of incomplete observation;
($\neg\alpha \vee \Box\alpha$: the formula is not a postulate).

Definition 11 (Conjecture: $\neg\alpha \wedge \neg\Box\neg\alpha$): A *conjecture* is a simple supposition founded on apparent possibility or probabilities, a hypothesis which has not received any confirmation. Formally, a conjecture is a statement which existence is assumed without being observed, since it is not refuted. For instance, *tomorrow is a new day* is a conjecture.
($\alpha \vee \Box\neg\alpha$: the formula is not a conjecture)

Definition 12 (Theory): A theory includes a set of postulates and a set of conjectures: an ongoing theory is not supposed to be complete and consistent. The use of paraconsistent and paracomplete logics to formulate theories simply comes from the fact that postulates and conjectures are expressing contradictions when using respectively a paraconsistent and a paracomplete negation.

As [12] and [13], we combine the three squares presented on Figure 4 to form a higher order cubic structure, the *Cube of Oppositions* (Figure 5), relating all the modalities presented so far.

The squares are not visible on the faces, but on the diagonals of the cube. Indeed, since the relations of contradiction are visible on the diagonals of Aristotle's square, it is natural that the different squares form the diagonal planes of the cube of oppositions.

The latter is built from two distinct tetrahedrons. The one of contrariety, which vertices are those from which the implication arrows start, opposes the proof $\Box$ to the modalities that can imply its contrary $\neg\Box$, and the one of sub-contrariety which vertices are those to which the implication arrows lead. The latter opposes the contradiction of a proof $\neg\Box$ to the modalities that can be derived from a proof. Any vertex of the cube is then contradictory to the furthest lying opposite vertex (easily obtained by central symmetry).

A proof is the result of a dynamic process of constant revision: a new proof is interesting when it proves some conjectures or eliminate surnumerous postulates, and is reciprocally suspected when it proves some conjecture that are reputed unsolvable or false. Logicians appreciate that solvers reason with consistent and complete theories, this is why theories which take into account inconsistency

and incompleteness model them by believes, intentions, and defaults. However, during the interactive phases of learning, it is illusory to try being consistent and complete by considering that errors are the defaults of some known consistent and complete theories, which would lead to making a theory of the whole. A discovery is triggered by the resolution of a contradiction. For instance, it is contradictory to consider a statement as a proof and as a postulate, since proof is related to postulate by a contrariety relation. So when a postulate is proven then either it has to be removed, either the proof is false. This shows how this set of modalities is closed by negation, and why paracomplete and paraconsistant contradictions are needed to reason with incomplete and inconsistent knowledge.

By chance, this cube of oppositions is not an exotic logical structure, but simply a fragment of a classical logic which is used to raise an alert when a theory needs to be revised, i.e. when a postulate or a conjecture is proven. How can a postulate or a conjecture be empirically proven? This question shows the limits of this formalism, as illustrated in section V-B, and leads to the hypercubic structure presented in section VI.

B. E+N, a game of scientific discovery

The application chosen to illustrate this first part of the formalism is *E+N*, a toy game inspired from Abott's mathematical game "Eleusis" [14], and an interaction protocol "Nobel" designed by cognitive scientists to study collective behaviors in a controlled and parametrized environment [15]. As described in [2], E+N game was designed to simulate the discovery phenomenon and the formation of theories by implementing an Angluin like machine learning interaction protocol [16]. It is currently experimented in with PhD students, to gather enough data to enable the comparison with their results with those of computer assisted players, but also with children to validate the didactic impact of the game on their learning dialectics.

E+N is a card game in which players have to discover hidden rules determining the valid card sequences that can be formed during the game. Players have access to personal experimentation spaces in which they test sequences to observe their validity for the selected hidden rule. They also have access to a public environment in which they can publish their own theory explaining a hidden rule, read the ones submitted by other players, and possibly refute them when they find a sequence which is irrelevant with what was published. The game ends when the flow between published and refuted theories stops, and the winner is the player with the higher score. A player scores by publishing and refuting theories, and loses points when his own theories are refuted.

The choice to use cards was made to make this game accessible to the largest backgrounds as possible (including children), but this simple game is actually a constructive and collective process of theory construction during which players have to observe the sequences at their disposal, and formulate postulates and conjectures to build a publishable theory. The social game confronting players who do not have tested the same sequences create cases of contingency: when publishing a theory without testing sequences to prove or refute possible and contingent statements, there is a risk that another player has tested them. We insist on the fact that this game has been simplified for experimentation reasons: the length of the sequences is fixed to two cards, and the description of the cards is also fixed and known by every player, we will discuss in section V-B.2 the consequences of these simplifications, after illustrating with the game's concepts the definitions given so far.

1) Illustration of Definitions:

- Observed (α): By creating a new card sequence, one can observe if it is accepted or not. On the contrary, $\neg\alpha$ denotes that a sequence has not been played. For instance, the sequence S=[King of hearts][Ace of spades] is either played, or not.
- Proof ($\Box\alpha$): When a sequence is valid, then the statement describing it is proven. Considering that the way to describe cards is fixed and known, and supposing that the card sequence S given above is valid (accepted by the hidden rule), then the following statement St, is proven: *A card which form is "hearts", which color is "red", which strength is "King", and which is a court card, can be followed by a card which form is "spades", which color is "black", which strength is "Ace", and which is not a court card.*
- Refutation ($\Box\neg\alpha$): Supposing that the card sequence S given above is not valid (rejected by the hidden rule), then the St is refuted. If the statement in question, or part of it, has been published, then the player who observes the counter example can refute it and score points. For instance, a theory as "A red card is followed by a black card" can be refuted if S is not valid.
- Possible ($\neg\Box\neg\alpha$): Possible denotes the fact that a sequence has not been played, and therefore a statement describing it is not refuted or contradicted.
- Contingent ($\neg\Box\alpha$): Contingent denotes the fact that a sequence has not been played, so a statement describing it is not proven. Surely, there is a temporal notion underlying these two last modalities, since what is possible or contingent has not been played *yet*. Contingency is strongly linked to the notion of action, and in this game, every card sequence is contingent, as the result of the player's choice. It is not as if cards were played continuously, following a specific rule, as Earth turning around the sun in a bit less than 365 days, which is a phenomenon strictly observed, and which cannot be influenced by the observer. Moreover a sequence, until it is played, can be at the same time possible and contingent since these modalities are linked by a contrariety relation.
- Postulate ($\alpha \wedge \neg\Box\alpha$) A postulate can be seen as a statement describing only partially an observed and

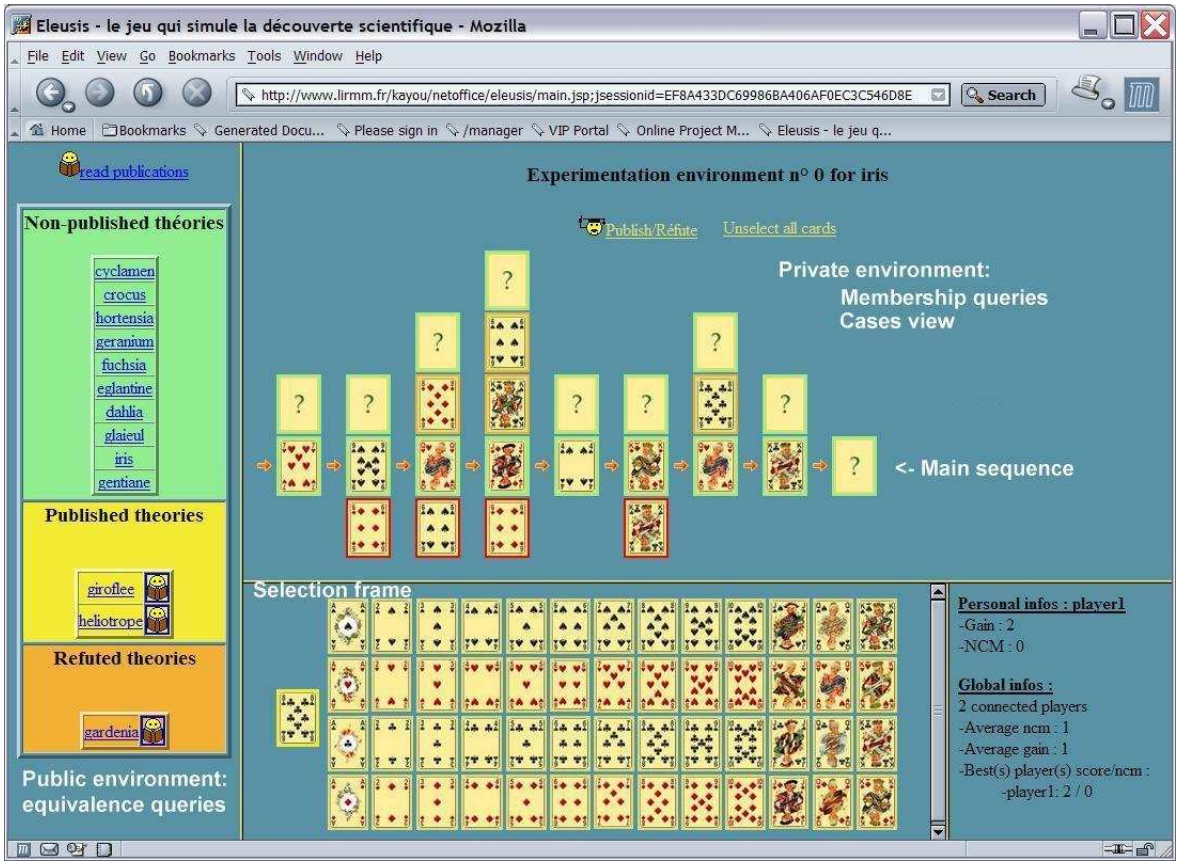

Figure 6. Eleusis + Nobel Game display

contingent card sequence. Indeed, the description may not be accurate or really relevant. Since the language to describe cards is given in this simplified version of the game, there can be no postulates, these statements are already proven.

- Conjecture ($\neg\alpha \wedge \neg\Box\neg\alpha$): A conjecture is a set of logical rules that produce predictions (statements) which are not observed and possible. For instance, if there is at least one sequence of a red card followed by a black card which has not been observed as valid, then the statement "a black card will be accepted after a red one" is a conjecture, but can be used to predict, associated to a risk of producing an error.

The dynamic of the interactive process is given by the "Nobel" part of the game during which each player, motivated by the acquisition of credits, publishes his own theories: each player has his own bias, and individual theories are improved by being confronted to others.

2) Discussion: As it was emphasized in the presentation of this game, some simplifications were made to make this game accessible to the children who experimented it at first: hidden rules are consistent and complete on the set of sequences of two cards, and the cards description is also fixed and known by each players. As we mentioned in the illustration of postulates (section V-B.1), the first consequence of these simplifications is to eliminate the use of postulates. A harder game could make players doubt about the most pertinent vocabulary to describe the properties of cards involved in the hidden rule, as well as about the relations between cards (hidden rule concerning n consecutive cards, or a card could be related to the one placed in the n^{th} position after it). In this harder game using postulates would make sense, since the observation of a sequence would only proof its existence. Moreover, in the context of scientific discovery, a publication not only contains postulates and conjectures, but also a model together with a device for experimental validation and facts confirming or invalidating the model's predictions, which are not taken into account in the Cube. The state of the art and the credibility of an author (credit) are not dealt with neither. However, some of these are already present in E+N and visible on Figure 6: the state of the art, for instance, is represented by a journal of publications, listing every theory published, as well as the eventual refutations (left hand side). On the right hand side is shown information as personal credits (score), the credits of the best player, the average number of cards played before publishing ...The main frame shows the private environment in which each player can freely play cards to form sequences, and this whole experimentation environment could be published as a device for empirical validation or refutation.

Section VI fills the gaps mentioned earlier, and presents how the cube is extended into a hypercubic structure by absorbing two other squares of oppositions defining the relations between modeling and proof or refutation on the

one hand, facts and predictions on the other.

VI. The Hypercube of oppositions

In this section, we suppose that each scientist publishes his own model and the experimental protocol to validate it, together with the postulates and conjectures defining the theory as mentioned in section V. A scientist also produces experimentable facts, as well as possible and contingent statements making factual some predictions coming from others scientists' model. Section V presents a set of modalities closed by negation and used to describe the formation of a theory, in a mathematical way, in a purely theoretic manner. But as Popper [17], we believe that science is pragmatic and must empirically rely on experimentation and be validated by the confrontation of a theory to reality. Every formal element has to be experimentally refutable by a scientific society, and proof and refutation must form the social accreditation process of a formalism.

In other words, the role of the computer assistant is to build, starting from facts, a model which is used to simulate a phenomenon and predict the behavior of a system, in an almost scientific manner. The modalities representing these notions are introduced in section VI-A and are based on a higher-order modal extension of the previous cube of oppositions [12], [18]. An illustration in drug design is provided in section VI-B. Section VII then present informally how agents using such a logical framework can interact to agree on a consensual vocabulary to describe a phenomenon.

A. Modalities in empirical Theory formation

Definition 13 (Model: $\sim$): Modeling is a method of study and measurement consisting in replacing a studied system by a model supposed to have an analogous behavior. $\sim$ is the universal modality associated to a formal statement representing a model.

Definition 14 (Experimentation: $\neg\sim$): An experimentation is an effective test realized to study a phenomenon. $\neg\sim$ is the existential modality associated to a statement expressing an experimentation.

Definition 15 (Prediction: $\Box \vee \sim \vee \Box\neg$): A *prediction* is the result of a computation, or a principle. $\Box \vee \sim \vee \Box\neg$ is the existential modality associated to a statement implied by a proof or a refutation or a model.

Definition 16 (Fact: $\neg\Box \wedge \neg\sim \wedge\neg\Box\neg$): A *Fact* is the concrete result of an action, and has an unquestionable empirical evidence for scientists. $\neg\Box \wedge \neg\sim \wedge\neg\Box\neg$ is the universal modality attached to a statement from which may be implied an experimental, possible and contingent result.

The two opposition squares relating these new modalities are visible on Figure 7 (only modalities are shown to simplify the diagrams). They can be linked to the previous cube to build a hypercubic construction of higher-order geometrical figures of oppositions as suggested by [12] (Figure 8 shows only a part of this hypercube for clarity reasons).

Figure 7. the squares defining simulation and experimention towards proof and refutation

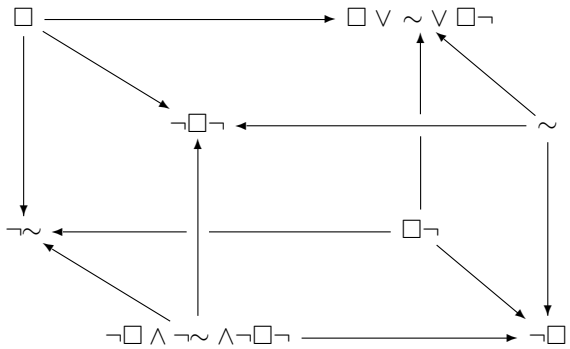

Figure 8. One facet of the hypercube of oppositions

Both experimentation and modeling produce results, and it is the confrontation between them that puts forward an eventual contradiction between a phenomenon and the model used to simulate it, which leads to put into question proofs, conjectures, and postulates to localize the theoretical error.

More generally, each face opposes two universal modalities by confronting the two particular modalities which can be derived from both of them. For instance, different models create different points of view, and modify postultes and conjectures. A first result of this methodology is to reveal the facets of different learning techniques. On the front face, the adequacy of "fact" and "model" is related to the production of possible and contingent statements which are linked by a subcontrariety relation. The corresponding learning method are version space or Galois lattice techniques. The left face as the bottom face links respectively the "fact" to the "proof" or to the "refutation" of the model. The three other faces are related to the analysis of the prediction. The back face is used to compare a prediction, i.e. a "theoretical result" given by a "simulation", to an "experimentation" which concretizes some "experimental results" (the fact that an object falls when it is released on earth is observable by experimenting it on a particular object in particular conditions).

In the following section, we illustrate the different facets of this supervision strategy of a learning process on an industrial application in Drug Discovery. The form of the "hidden rules" is not known anymore, but several models can exist.

B. Application to the prediction of Absorption

In this section, we present an application of the hypercube in scientific discovery, on the absorption problem. In the example E+N presented in section V-B, the objects

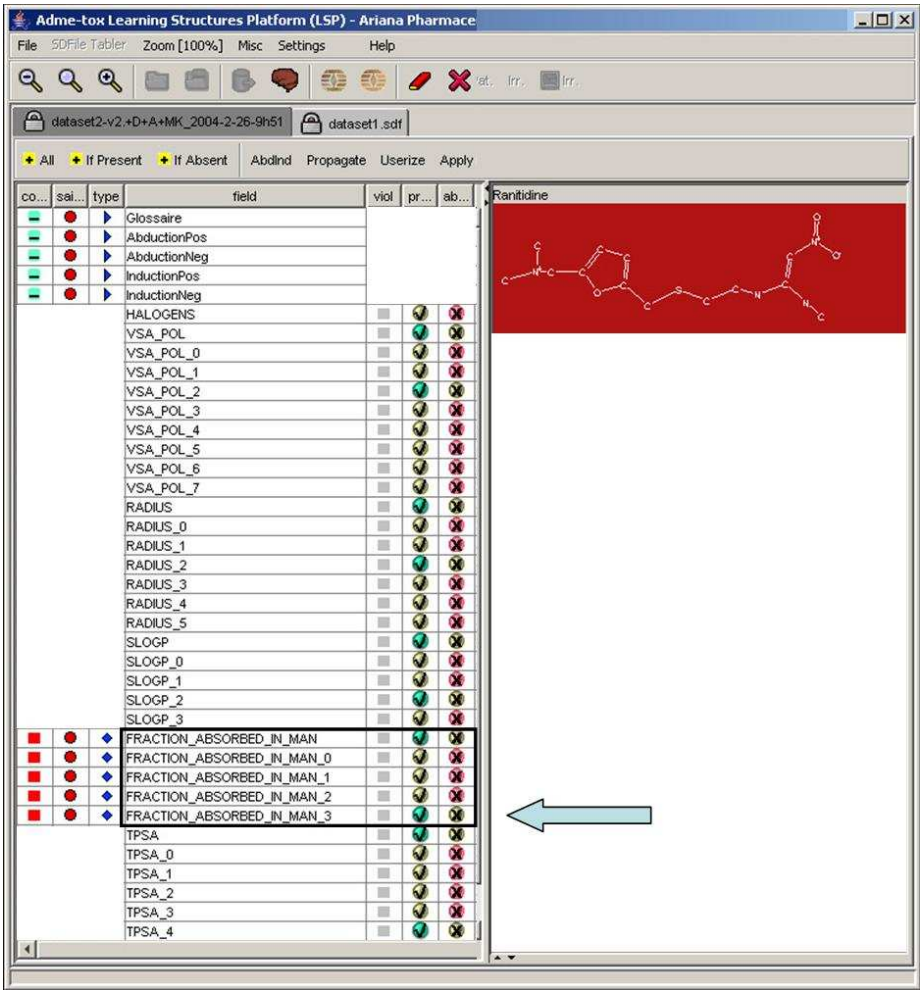

Figure 9. KEM^{TM}

were described using a fixed and known ontology and the hidden rules an error free answer to each query concerning a card sequence.

In this new problem a scientist is using a learning machine to learn a model from facts. This model is formulated as a constraint network, and each constraint concerns properties of the molecular graph's description. As the model is learned from examples, since predictions may be done in a context of overfitting and underfitting, contradictions are the events which allow the supervision, by the expert, of the learning assistant modeled as a hypercube.

This real application of learning in scientific discovery results from collaboration with Ariana Pharmaceuticals in Drug design [19]. KEM^{TM} can suggest specific molecular modifications to achieve multiple objectives, after analyzing a multi-parametric database.

In this example we focus on the prediction of absorption, a key issue in drug design since this is one of the important and early causes of failure in the drug discovery process. Indeed molecules need to be absorbed before they can perform any desired activity. Absorption is a complex process involving both passive (diffusion) and active (through transporter proteins) across cellular membranes. For passive transport, molecules need to be soluble (hydrophilic) in water and at the same time they need to be greasy (hydrophobic) to penetrate cellular membranes that are formed of lipids. This contradicting requirement is modulated by active transport, where molecules need to be recognized (i.e. complementarity of shape and charge) by another molecule (transporter) that helps them through membranes. Although no one can for sure predict the absorption of a new molecule, a number of empirical rules are known. This is an interesting context for applying our IA since our key requirement is to capture knowledge from the experimental data and then evolve and improve this model in a consistent manner.

To illustrate our approach we focus on a set of 169 molecules for which the absorption in human tissues has been experimentally evaluated (4 classes. 0 not absorbed, 3 highly absorbed) [20]. These molecules are described using a set of physico-chemical properties such as molecular radius, different calculated measures of their total polar surface accessible to water (*TPSA* and *VSA POL*), their hydrophobicity (*SLOGP*), presence of halogens etc.

To learn, KEM acts according the facet of the hypercube of oppositions visible on Figure 8: 1)(left face)A decision tree is used to find a good segmentation of

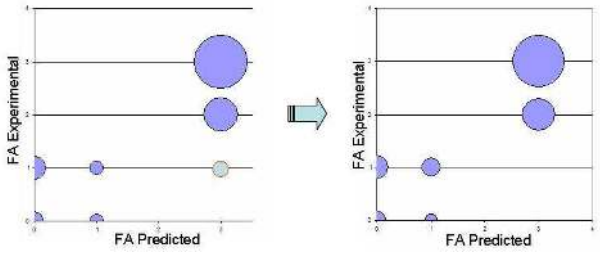

Figure 10. Predictions A and B

the numerical descriptors. 2)(front face) A Galois lattice method works on these binary descriptors in order to construct a lattice of regularities. 3)(upper face) sup-irreducible nodes are translated into logical constraints for prediction. 4)(back face) the prediction is confronted to the experimentation. 5)(right face) the study of the experimental error is done by a refutation of the simulation.

Initially, the system learns from the data set a set of rules linking the structure of the molecule to the absorption. The quality of the prediction is tested in a subsequent stage on a novel set of molecules. The results are shown on prediction A in Figure 10. Ideally the predictions should be on the diagonal. An error of one class is tolerated. However, it is clear that for one molecule, the error is larger (i.e. experimental: class 1 vs predicted: class 3). This confrontation between a theoretical result and an experimental result puts forward a contradiction in the model.

Let us sketch this dialog: the molecule (Ranitidine) has been predicted with *fraction absorbed in human 3* i.e. highly absorbed. However, if the user makes a postulate and forces *fraction absorbed in human 3* to be false, the system localizes the error that induced the contradiction by showing that the postulate contradicts the conjectural learned rule *VSA pol 2* $\rightarrow$ *fraction absorbed in human 3*. At this stage the user realizes that indeed this conjecture was true for the learning set, however this is not generally true and it can be eliminated. The user then goes back to simulating once more the test and results are shown in Figure 10, prediction B. As expected, the results have been improved. The important point is that the improvement has been done in a controlled way under the user's supervision, and this was only possible because the user and KEM^{TM} shared a common vocabulary to type statements.

In scientific discovery, there are in general no Oracles who can say a priori whether a prediction is correct or not. Experimentalists formulate a conjecture that is consistent with existing empirical data and then set about to test it. We believe that the key for a computational system is to adhere to the same process i.e. build up an explanation / reasons for predicting an outcome. If the system is able to provide enough arguments, the user will "trust" it and try the experience. This implies that the arguments are annotated with modalities which are meaningful both for the user and the machine.

VII. The Institution Agent social game

This section is the final step of the formalism and links the classical formalisms used to express the individual and collective behaviors of agents. Our contribution is to describe the dynamic process of conventional formation of theories by agents (human as well as artificial agents) in respect with social norms regulating a community, in order to describe a community of agents interacting with each other to understand their environment and organize themselves to fix common objectives and chose the best actions to achieve them. The objective is then to give the intuitive idea about mathematical developments, which are our ongoing research, and to show that this is a constructive approach to build agents comparable to classical BDI agents, which use Beliefs, Desires, and Intentions to reason and decide.

The only constraint fixed on agents is to reason using a hypercube (section VI) and a normative system NS, which is a logical language using the cube modalities to communicate.

Definition 17 (Normative System): We call a Normative System (NS) the couple (L, $Cube$) formed by:

- L: a language formed by a hierarchy of concepts and the relations between them,
- $Cube$: a cube of oppositions.

Definition 18 (IA): An IA is an agent using an $Hypercube$ framework to reason and judge statements. As we emphasized, the Hypercube is a formalization of dialectic as it is practiced by humans, and an IA can then be either an artificial agent, or a human agent. The interaction protocol is compliant with both of them since the definitions of the modalities are shared by both of them.

Scientific discovery is a collective process, and needs interaction between researchers to exchange their points of view and judgments. This is how IAs interact (Figure 12): by exchanging judgments about statements, more exactly, by asking another IA if it agrees with a particular judgment: "this statement is a conjecture, is it not?", to which the answer is "yes" or "no, it is a result", or if a theory seems relevant or not. This type of interaction is well described in Machine Learning theories and these queries are known respectively as *Membership* and *Equivalence* queries. [16], [21] shows that the use of at least these two types of queries is required for an effective learning. Exchanging judgments creates the negation in the common frame of reference (language), and the revision of the normative system associated with one IA or the other. Two judgments are especially important: *judging one's conjecture as being a paradox*, and *judging one's postulate as already being a result*. KEM^{TM}, presented in section VI-B, illustrates this control by a scientist over the artificial IA assisting him.

We suppose that each IA can be represented by a particular normative system resulting from its own experience and adaptation during an interaction with other IAs, and we assume that the logic used during a decision process is the same for every IA. We then focus on

adaptation and interaction of IAs sharing a vocabulary and trying to build a common language or frame of reference with this vocabulary to describe a phenomenon and provide consensual definitions, expressed logically or by constraint networks.

Three logical properties, presented in the following, are needed to qualify this interaction protocol and to add a logical control to the adaptation process:

- deontic: an IA must be able to attribute credits to another IA, to interact, and to teach another IA,
- defeasible: Lower IAs must be able to adapt their behavior to the norms imposed by the higher ones,
- autoepistemic: an IA can be seen as composed by at least two interacting IAs and can therefore learn its own hierarchy of norms and auto-adapt.

section VII-A presents how a deontic and defeasible logic can be used to frame the collective dynamic during which IAs build a common NS, and section VII-B suggests how an IA can be itself composed by several IAs to form an autoepistemic agent performing an inner dialog while reasoning.

A. Collective dynamic

1) Deontic logic: Often used in multi-agent systems to constrain an agent's behavior, annotable deontic logic uses modalities expressing obligation, interdiction, advice, and warning. According to Frege's definition, these statements express a judgment, i.e. the recognition of the type of truth of the statement [22]. Imputations (gains or losses, risk estimation) are used to estimate the risk incurred in a given situation to decide what action to take or what behavior to adopt. A modality and an imputation have to be used to express statements of the following form: "The obligation to respect the speed limit is attached to a imputation of x". A credit value can also be associated to IAs, ordering them hierarchically, to define which one is the most qualified to rule in a given context, for example by defining a social organization as a government with a parliament, a senate...

2) Defeasible logic: It is possible to link two IAs by respecting a defeasible logic to take into account a hierarchy of Institution Agents. The resulting hierarchy of IAs has to be brought together with the transitivity axiom, that stands as follows: "What is necessary in an upper NS of is also necessary in a lower NS". In other words, no one should be unaware of the law, no one should go against a superior law. [23] gives a concrete usage of defeasible logic, that allows us to order rules and to supervise an IA, for example with another higher IA, as illustrated on Figure 13.

- Every Obligation of a lower IA belongs to the superior IA's advice.
- Every Interdiction of a lower IA belongs to the superior IA's warnings.

The middle line shows the conditions according to which an IA can be supervised by another one. The violation of this constraint ($O_2 = \Box_2 \subset I_1 = \Box_1\neg$ or $I_2 = \Box_2\neg \subset O_1 = \Box_1$) can put forward contradictions between the two IA's normative system. Finding a contradiction, and trying to eliminate it, leads to the initiation of a transaction between the two IAs, during which they adapt their theory. When no contradiction remains, a new IA can be created, formed by the association of the two precedent IAs, and this process ensures the traceability of all the events leading to an IA's creation.

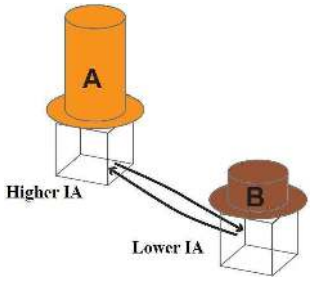

Figure 11. IA's credit

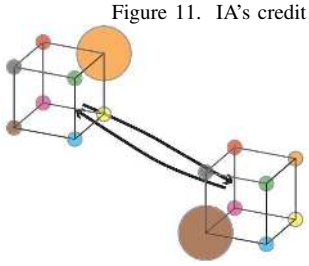

Figure 12. Exchanging judgments

B. Personal dynamic

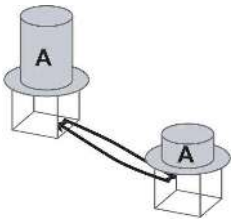

Figure 14. Autoepistemic dialog

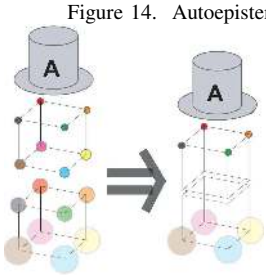

Figure 15. IA formation

1) Autoepistemic logic: Aristotle distinguishes endophasy as an inner dialog ("to think is to speak with oneself"). This is a constructive manner to build an intelligent agent as the result of an self learning. The inner IAs can be interpreted as managing believes, desires or intentions (BDI), for example. By applying the dialectic and deontic interaction we have presented in section VII-A.1, an IA is able to acquire its own NS, which prepares

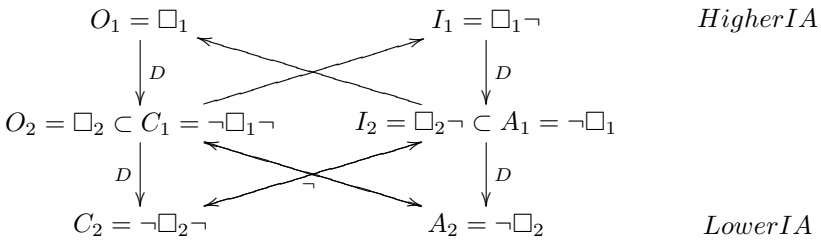

Figure 13. Normative system hierarchy

an efficient learning, and even enable self learning from examples.

In this section, we presented how an interaction process and a hierarchical control can be used to define an agent able to adapt its defeasible deontic and autoepistemic Normative System.

VIII. Conclusion

We presented a logical framework in order to capture the entire process of scientific discovery as it is performed in scientific comunity. This framework is based on Aristotle's square, a conventional logical structure, and recent related works in logic. In this context, we focused on a geometrical representation of the notion of logical contradiction.

Our approach is progressive. We first presented with a cubic structure the formation of a theory as driven by the proof of its postulates and conjectures by scientists. Then we presented a hypercubic extension to capture the empirical modeling processes, to make the predictions of a model and the experimentations explicit. We showed how this theory construction can be done by human and by machines taught by humans in a formal and in a practical way. Finally we presented in an informal way how the institution agents (agents using this hypercubic structure to build theories), educated by human, can be constrained to respect normative systems and how this approach is only a constructive alternative to BDI agents.

We illustrated on an example coming from Drug Discovery how these annotations of common sense, which are now logically defined, are used by a learning machine and a scientist to interactively build a model which is coherent and complete with observations and experimental results. We prone that this hypercube describes in a universal way a rational agent and enables the supervision of its computing process.

In perspective, to complete our experimentation, we plan on experimenting this constructive approach to teach a humanoid robot and make him gain more autonomy. A mathematical formalism of agents based on this vision is currently developed in category theory.

References

[1] T. Parsons, "The traditional square of opposition," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Winter 2006. [Online]. Available: http://plato.stanford.edu/archives/win2006/entries/square/

[2] C. Dartnell and J. Sallantin, "Assisting scientific discovery with an adaptive problem solver." in *Discovery Science*, ser. Lecture Notes in Computer Science, A. G. Hoffmann, H. Motoda, and T. Scheffer, Eds., vol. 3735. Springer, 2005, pp. 99–112.

[3] J. Sallantin, C. Dartnell, and M. Afshar, "A pragmatic logic of scientific discovery." in *Discovery Science*, ser. Lecture Notes in Computer Science, L. Todorovski, N. Lavrac, and K. P. Jantke, Eds., vol. 4265. Springer, 2006, pp. 231–242.

[4] S. Morley and M. Afshar, "Validation of an empirical rna-ligand scoring function for fast flexible docking using ribodock." *Journal of Computer Aided Molecular Design*, pp. 189–208, 2004.

[5] M.-G. N., W. K. A. B.K., and K. R.D., "New approach to pharmacophore mapping and qsar analysis using inductive logic programming. application to thermolysin inhibitors and glycogen phosphorylase b inhibitors," *Journal of Medicinal Chemistry*, vol. 45(2), pp. 399–409, 2002.

[6] B. Faller and F. Wohnsland, "Physicochemical parameters as tools in drug discovery and lead optimisation," in *Pharmacokinetic optimization in drug research*. Testa, Waterbeemd, Folkers and Guy editors, Wiley-VCH, Zurich, 2004, pp. 189–208.

[7] J.-Y. Béziau, "New light on the square of oppositions and its nameless corner," *Logical Investigations*, vol. 10, pp. 218–233, 2003.

[8] R. Blanché, *Structures intellectuelles : essai sur l'organisation systématique des concepts*. Vrin, Paris, 1966.

[9] K. Došen, "Intuitionistic double negation as a necessity operator," *Publications de l'Institut Mathématique*, vol. 35(49), pp. 15–20, 1984.

[10] ——, "Negative modal operators in intuitionistic logic," *Publications de l'Institut Mathématique*, vol. 35(49), pp. 3–14, 1984.

[11] J.-Y. Béziau, "S5 is a paraconsistent logic and so is first-order logic," *Logical Investigations*, vol. 9, pp. 301–309, 2002.

[12] A. Moretti, "Geometry for modalities? Yes: through n-opposition theory," in *Aspects of Universal Logic*, J.-Y. Béziau, A. Costa-Leite, and A. Facchini, Eds. Travaux de logique 17, Neuchâtel, 2004, pp. 102–145.

[13] R. Pellissier, ""Setting" n-opposition," in *UNILOG05*, (to be published), 2006.

[14] M. Gardner, "Mathematical games," *Scientific American*, June 1959.

[15] D. Chavalarias, "La thèse de popper est-elle réfutable?" Memoire de DEA, CREA - CNRS/Ecole Polytechnique, 1997.

[16] D. Angluin and Krikis, "Learning from different teachers," *Machine Learning*, vol. 51, pp. 137–163, 2003.

[17] K. R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*. Harper and Row, 1963.

[18] D. Luzeaux, J. Sallantin, and C. Dartnell, "Logical extensions of aristotle's square," in *Logica Universalis*, (to be published), 2007.

[19] M. Afshar, A. Lanoue, and J. Sallantin, "New directions:

multidimensionnal optimization in drug discovery," *Comprehensive Medicinal Chemistry 2*, vol. 4, 2006.
[20] Y. Zhao, M. Abraham, A. Hersey, and C. Luscombe, "Quantitative relationship between rat intestinal absorption and abraham descriptors," *Eur J Med Chem.*, pp. 939–947, 2003.
[21] D. Angluin, "Queries revisited," *Theoretical Computer Science*, vol. 313, pp. 175–194, 2004.
[22] C. D. Pozza and C. Garola, "a pragmatic interpretation of intuitionistic propositional logic," *Erkenntnis*, vol. 43, pp. 81–109, 1995.
[23] K. Nakamatsu, T. Kato, and A. Suzuki, "Basic ideas of defeasible deontic traffic signal control based on a paraconsistent logic program evalpsn," *Advances in Intelligent Systems and Robotics*, 2003.

Mohammad Afshar is the CEO and co-founder of Ariana Pharmaceuticals. He was one of the founding members of the senior management group at RiboTargets, Cambridge, UK, where, as the Director of IT and Head of Drug Design, he set up and managed RiboTargets' structure based discovery platform RiboDock^{TM} - rDock. The patented technology developed within his team (RiboDock^{TM} - rDock) allowed the identification and validation of novel therapeutic molecules, becoming the central focus of RiboTargets. He was actively involved in the companys business development.

Christopher Dartnell is currently a PhD candidate at UM2, The Science University of Montpellier, France. He received his MS and degree in computer science in 2003 and is currently working in collaboration with the LIRMM (Laboratory of Computer Sciences, Robotics, and Mathematics of Montpellier) and Euriware, specialist in industrial IT and information systems for the energy sector (member of the Areva group). His research interests include machine learning, assistance for the supervision of complex systems, logical reasoning and robot control.

Dominique Luzeaux graduated from École Polytechnique (1987), École Nationale Supérieure des Techniques Avancées (1989), PhD from University Paris XI (1991). He held the position of an invited postdoctoral fellow at the University of Berkeley in 1991-1992. In 2001 he completed his professor thesis at the University Paris XI. He is currently director of the Technical Center for Information Systems within the Ministry of Defense. He also teaches robotics and theoretical computer science at the École Nationale Suprieure des Techniques Avancées and at graduate level.

Jean Sallantin is a Director of Research at the LIRMM Laboratoire dInformatique, de Robotique et de Micro-électronique de Montpellier of the CNRS in Montpellier, France. Phd from University Paris VI in Theoretical Physic (1972) and professor thesis in Computer Science (1979). He currently heads the Rationality and machine learning team that develops and studies the applications of machine learning techniques in a scientific discovery environment. One of the pioneers of bioinformatics in France, Jean Sallantin established the bioinformatics laboratory at the Curie Institute Paris (1983-85) and was the director of the GSDIARL research consortium (1985-89).

Yannick S. Tognetti is a research scientist at ariana pharmaceuticals, specialist in Artificial Intelligence. He obtained a PhD in Computer Science and Artificial Intelligence working on the modeling of chemical information systems using graph's theory and algorithmics, in 2002 at LIRMM Montpellier.