A Formal Ontology Framework for Societies of Communicating Agents through Arithmetical Theories endowed with Self-reference and their Models

Paola Forcheri*, Paolo Gentilini**

*Istituto di Matematica Applicata e Tecnologie Informatiche,
Consiglio Nazionale delle Ricerche
Genova, Italy
forcheri@ima.ge.cnr.it

** Istituto Regionale di Ricerca Educativa - IRRE Liguria
Genova,Italy
gentilini@irre.liguria.it
and
Istituto di Matematica Applicata e Tecnologie Informatiche,
Consiglio Nazionale delle Ricerche
Genova, Italy
gentilini@ima.ge.cnr.it

Formal Ontology Framework for Societies of (cognitive) Communicating Agents

Aims and approach

* The focus is on an ontological taxonomy allowing an explicit and detailed representation of *knowing, thinking, communicating* of a complex cognitive agent

* The emphasis is more on the cognitive aspects than on the world representation aspects (object existence in time and space, time and space, and so on)

*An intrinsic representation of *knowledge*, *belief* and *reasoning* of an agent ins given, through Proof-theory, Provability Logic, Model theory of over-systems Ti of first order Arithmetic: agents are identified with such over-systems

* A formal expression of *intentionality* is proposed, that works in the introduced notions of *ontological commitment* of an agent and *communication willingness* of an agent

* A formal notion of *communication act* between complex cognitive agent is proposed, allowing the definition of *relevant communication act*

Introduced ontological categories

*Coherently with the stated aims and approach four ontological category are (formally) defined:

basic facts
agents
ontological commitments of agents
communication willingness of agents

Ontology for Societies of Complex Agents

The Category of Basic Facts

Operational approach to the ontology of the physical world:
the universe of the *basic facts* includes all situations and objects that an ideal operator can build, measure or define through effective processes

Assume the following form of the Church thesis: *each effective process can be represented by a recursive predicate to which a predicate letter in the language of the Primitive Recursive Arithmetic* **PRA** *corresponds*
Then
Basic fact ≡ entity that can be described by a recursive predicate.

A suitable encoding (godelization) of the objects of the world,
a temporal coordinate *t* for the physical time, ranging over natural numbers,
and
a Cartesian spatial reference system *(O,XYZ),*

are *a priori* established

The Category of Basic Facts

Therefore, each event *V* can be unambiguously temporally and spatially located by recursive terms *t*, *s*, that includes the codes of the temporal and spatial coordinates of *V*.

Each atomic formula representing a basic fact in the formal language is of the form *B(....,r,z)* where *r* has to be interpreted as a space term and *z* as a time term.

Agents and Societies

Definition (Agent)

A *formal agent* is a first order formal system constituting a theory T_i such that:

i) the T_i-language L includes the language of *Primitive Recursive Arithmetic* **PRA** plus a three-place function letter *f*, called *choice function letter*, and a four-place function letter *g*, called *communication function letter*, which do not belong to the **PRA**-language;

ii) the T_i-deductive apparatus includes the sequent version of the first order *Predicate Calculus* **LK**, a non empty set $\mathcal{A}(T_i)$ of proper T_i-axioms, and possibly a set $\mathcal{R}(T_i)$ of T_i -inference rules;

iii) T_i is consistent and recursively axiomatized. ■

Remarks

* Axioms $\mathcal{A}(T_i)$ and rules $\mathcal{R}(T_i)$ characterize T_i with respect to any other agent T_j $(j \neq i)$

*agent is able to perform and recognize **effective procedures involving numbers,** and objects coded by numbers

Agent: remarks

*an agent is able to distinguish between what he postulates (or takes as given from his environment) and what he draws from that through inference.

***proofs** of $\mathbf{T}_i$ (i.e. possibly sequent proof-trees) **are the reasoning** of the agent.

*Through the **provability predicates** $Pr_{\mathbf{Ti}}(.)$ any agent $\mathbf{T}_i$ can reason about his proofs, that is, about his reasoning and conclusions: therefore, **he is naturally endowed by a self-reference capability.**

Definition (Society)

A *communication society* is a set $E=\{T_1,...,T_m\}$ of formal agents ■

In principle, a dialectic communication between agents is allowed, since we do not impose that the T_i's must be pair-wise consistent.

Moreover, all agents share the same language L, that thus constitutes the language of the society

For the sake of brevity we assume here that each T_i of E is an over-system of **PRA**

Example: Minimal *classical* PRA-based agent

The Structure of Agent Reasoning: sequent proof trees
[a sequent is $X \vdash Y$ where X, Y are sets of formulas; if $X = \{A1,..,An\}$, $Y= \{B1,...,Bm\}$ then $X \vdash Y$ has the meaning of $A1 \wedge .. \wedge An \rightarrow B1 \vee .. \vee Bm$]

<u>Basic logical reasoning</u> is the sequent version of first order Predicate Calculus **LK**:

The sequent calculus LK:

Axioms: $A \vdash A$

Positive propositional logical rules:

$$\frac{B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \wedge\text{−L} \qquad \frac{B, \Gamma \vdash \Delta}{B \wedge A, \Gamma \vdash \Delta} \wedge\text{−L}$$

$$\frac{\Gamma \vdash \Delta, A \qquad \Lambda \vdash \mathrm{X}, B}{\Gamma, \Lambda \vdash \Delta, \mathrm{X}, A \wedge B} \wedge\text{−R}$$

$$\frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, A \vee B} \vee\text{−R} \qquad \frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, B \vee A} \vee\text{−R}$$

$$\frac{A, \Gamma \vdash \Delta \qquad B, \Lambda \vdash \mathrm{X}}{A \vee B, \Gamma, \Lambda \vdash \Delta, \mathrm{X}} \vee\text{−L}$$

The sequent calculus LK:

$$\frac{A, \Gamma \mid\!\!— \Delta, B}{\Gamma \mid\!\!— \Delta, A \to B} \quad \to\text{- R}$$

$$\frac{\Gamma \mid\!\!— \Delta, A \qquad B, \Lambda \mid\!\!— \mathrm{X}}{A \to B, \Gamma, \Lambda \mid\!\!— \Delta, \mathrm{X}} \quad \to\text{-L}$$

$$\frac{\Gamma \mid\!\!— \Delta, A}{\neg A, \Gamma \mid\!\!— \Delta} \quad \neg -\text{ L} \qquad \frac{A, \Gamma \mid\!\!— \Delta}{\Gamma \mid\!\!— \Delta, \neg A} \quad \neg\text{ - R}$$

Quantifier rules:

$$\frac{[t / x]A,\ \Gamma \mid\!\!— \Delta}{\forall x A, \Gamma \mid\!\!— \Delta} \quad \forall - \text{L} \qquad \frac{\Gamma \mid\!\!— \Delta, [b / x]A}{\Gamma \mid\!\!— \Delta, \forall x A} \quad \forall - \text{R}$$

$$\frac{[b / x]A,\ \Gamma \mid\!\!— \Delta}{\exists x A, \Gamma \mid\!\!— \Delta} \quad \exists - \text{L} \qquad \frac{\Gamma \mid\!\!— \Delta, [t / x]A}{\Gamma \mid\!\!— \Delta, \exists x A} \quad \exists - \text{R}$$

t arbitrary term and b free variable which does not occur in Γ, Δ. Moreover, t may be not fully quantified while b must be uniformly replaced by x

Structural rules:

Weakening rules: $$\frac{\Gamma \mid\!\!— \Delta}{\Gamma \mid\!\!— \Delta, A} \quad W - R \qquad \frac{\Gamma \mid\!\!— \Delta}{A, \Gamma \mid\!\!— \Delta} \quad W\text{-}L$$

Cut rule: $$\frac{\Gamma \mid\!\!— \Delta, A \qquad A, \Lambda \mid\!\!— \mathrm{X}}{\Gamma, \Lambda \mid\!\!— \Delta, \mathrm{X}} \quad \text{Cut}$$

Primitive Recursive Arithmetic PRA= LK + AxPRA

PRA language: a letter for each primitive recursive function *plus* constant **0** *plus* equality predicate **=(.,.)**

Proper PRA-axioms AxPRA:

1) Definition of primitive recursive function (variables x_i, y_j, are free):

1j) basic recursive functions

(*zero function, successor function, projection function*):

$$\vdash Z_k(x_1,\ldots,x_k)=0 \quad \text{for each} \quad k \geq 1$$

$$S(x)=0 \vdash$$

$$S(x)=S(y) \vdash x=y$$

$$\vdash P^i{}_k(x_1,\ldots,x_k)=x_i \quad \text{for each} \quad k \geq 1 \quad i \leq k$$

1jj) *Composition* schema:

$$\vdash f(x_1,\ldots,x_n)=h(g_1(x_1,\ldots,x_n),\ldots,g_m(x_1,\ldots,x_n))$$

where $g_1,\ldots,g_m\ldots$ are n-ary function letters and h is a m-ary function letter.

1jjj) *Recursion* schema :

.... $$\vdash f(x_1,\ldots,x_n,0)=g(x_1,\ldots,x_n)$$

$$\vdash f(x_1,\ldots,x_n,S(y))=h(x_1,\ldots,x_n,y,f(x_1,\ldots,x_n,y))$$

where g is a n-ary function letter and h is a n+1-ary function letter.

2)Equality axioms: $\vdash y=y$

$$x_1=y_1,\ldots,x_n=y_n \vdash f(x_1,\ldots,x_n)=f(y_1,\ldots,y_n)$$

$$x_1=y_1, x_2=y_2, x_1=x_2 \vdash y_1=y_2$$

Induction rule

The induction rule *Ind* has the following form:

$$\frac{F(x), X \vdash Y, F(S(x))}{F(0), X \vdash Y, F(t)}\ Ind$$

the free variable *x*, the *eigenvariable* of the rule, does not occur in *X,Y,t* and *t* is an arbitrary term. In the **PRA** Ind rule F(x) is quantifier free

The induction axiom schema is: (S successor function)

$[F(0) \wedge \forall x(F(x) \rightarrow F(S(x)))] \rightarrow \forall x F(x)$

Background: provability predicates

Tj includes the Provability Logic of **each** recursively axiomatized system T by the predicate $\mathbf{Pr_T(.)} \leftrightarrow \exists y \mathbf{Prov_T(y, .)}$ such that $\mathbf{Pr_T(B)}$ means**"B is T-provable"**

For each rec.axiom. system T we have:

T /---- B **implies** Tj /---- $Pr_T(B)$

in general Tj /----- $Pr_T(B)$ **does not** imply Tj /----- B

in general Tj **does not** prove $Pr_{Tj}(B) \rightarrow B$

Intrinsic representation of knowledge and belief

Definition

Let $E=\{T_1,..., T_m\}$ be a communication society. Let B be a sentence of the society language L. Then:

i) We say that the formal agent T_i, *declaratively knows the state of the world described by the sentence B* if T_i proves B;

ii) Let $\alpha \equiv sup\{\beta:\beta\in\varepsilon_0$, such that $T_i + IND_\beta$ is consistent$\}$ and let $U_i \equiv T_i + \{IND_\beta\}_{\beta\in\alpha}$. We say that *agent* T_i *believes that B holds*, if T_i proves $Pr_{U_i}(B)$;

iii) We say that *agent* T_i *believes that agent* T_j , $i\neq j$, *knows B*, if T_i proves $Pr_{T_j}(B)$. ■

The induction rule IND_β belongs to the following denumerable hierarchy:

$I(0),I(1),I(2),...,I(k),...,I(\omega),I(\omega+1),...I(\gamma),...,I(\varepsilon_0)$

$\equiv\{IND_\delta :\delta\in\varepsilon_0+1\}$

on the induction-rule schemata in sequent form such that: $I(k)$ is the induction rule schema over the set of natural numbers N with at most k quantifiers in the prenex form of the induction formulas;

$I(\beta)$, $\omega\leqslant\beta\leqslant\varepsilon_0$ is the induction rule schema up to β with arbitrary induction formulas;

ε_0 is the proof-theoretic ordinal of Arithmetic **PA**

Remark on the definition of belief:

The extension U_i of T_i in point (ii) is introduced to allow T_i to have beliefs which, though true in the standard model, are not obvious, whenever T_i could be extended by inductive inferences stronger than its own.

E.g., if T_i is consistent with $I(\varepsilon_0)$, then it may have the belief *Con(***PA***)*, that is, it believes that **PA** is consistent, without knowing (i.e. proving) *Con(***PA***)*, and without loss of the standard model.

We recall that $Pr_{\mathrm{Ti}}(B) \rightarrow B$ is T_i - provable if and only if B is a T_i -theorem, so that our notion of belief never trivializes; that is, in general, T_i does not declaratively know B, even if T_i proves $Pr_{\mathrm{Ti}}(B)$.

The intrinsic approach: provability predicates and knowledge representation

*If we represent $\mathbf{T}_i$ as **PRA** over-system, and assume that the essential properties of the predicate $Pr_{\mathbf{Ti}}(.)$ hold, e.g.:

D1) *If* $\mathbf{T}_i$ *proves* B *then* $\mathbf{T}_i$ *proves* $Pr_{\mathbf{Ti}}(\#B)$.

D2) $\mathbf{T}_i$ *proves* :

$$Pr_{\mathbf{Ti}}(\#A \rightarrow B) \rightarrow (Pr_{\mathbf{Ti}}(\#A) \rightarrow Pr_{\mathbf{Ti}}(\#B))$$

D3) $\mathbf{T}_i$ proves

$$Pr_{\mathbf{Ti}}(\#B) \rightarrow Pr_{\mathbf{Ti}}(\# Pr_{\mathbf{Ti}}(\#B))$$

<u>some of these properties are similar to the standard S5 properties,</u> **but we have deep differences**:

For example, the so called negative introspection $\neg KiA \rightarrow Ki\neg KiA$ which would correspond to $\neg Pr_{\mathbf{Ti}}(\#B) \rightarrow Pr_{\mathbf{Ti}}(\#\neg Pr_{\mathbf{Ti}}(\#B))$ **is rejected** from the $\mathbf{T}_i$ - declarative knowledge,

<u>whenever $\mathbf{T}_i$ preserves among its infinite models an expansion $\mathbf{N}_i$ of the standard model **N** of **PRA**</u>:

indeed, if we have that $\mathbf{T}_i$ proves $\neg Pr_{\mathbf{Ti}}(\#\bot) \rightarrow Pr_{\mathbf{Ti}}(\#\neg Pr_{\mathbf{Ti}}(\#\bot))$ then it proves $\neg Pr_{\mathbf{Ti}}(\#\bot) \rightarrow Pr_{\mathbf{Ti}}(\#\bot)$ which gives $\neg$Con($\mathbf{T}_i$), which is **N**-false, as $\mathbf{T}_i$ -provable.

Also the so called *arrogance principle* $KiA \rightarrow A$ is rejected by our framework since $\mathbf{T}_i$ cannot prove $Pr_{\mathbf{Ti}}(\#B) \rightarrow B$

Choices and Ontological Commitments of an Agent

Knowing is not choosing:
the awareness of the knowledge expressed by sentence *A,* is not a commitment of an agent T_i, $T_i \in E$, on the existence of the fact described by *A*.

Thus, we introduce a category of entities, that is the *choices* of an agent T_i, which are distinct from the basic facts.

Definition (Choice)

$E=\{ T_1,..., T_m\}$ a society of agents.

Let S be the set of the codes of the effectively definable spatial regions in the Cartesian reference system (O, *XYZ*) shared by the society *E*.

Let N be the set of natural numbers.

Then, a *choice function f*: $\{i\}\times N\times S\rightarrow N$ is defined by each proper axiom sets $A(T_i)$, $T_i \in E$.

We call a *choice-sentence* of T_i each formula of the language of the form $f(i,t,v)= \#B$, where *B* is a sentence of the language of *E*.

A *choice-theorem* of T_i (briefly a T_i -choice) is any T_i-theorem of the form $f(i,t,v)= \#B$.

We say that a T_i -choice is an *ontological commitment of* T_i *on the state of the world described by B*, if *B* is not **PRA**-equivalent to a Δ_0-formula. ■

Choices

The intended meaning of $f(i,t,v)= \#B$ is "agent i chooses the state of the world described by the sentence B as real at the time indicated by t in the place coded by v ".
B can be arbitrarily complex.
If B describes an action of T_i, then the T_i -theorem $f(i,t,v)= \#B$ says that the agent decides to produce such action.

An attempt to formalize intentionality

Choices are not basic facts, through the following assumption:

Assumption *Axioms* $A(T_i)$ *are such that the function* $f(i,\bullet,\bullet)$ *is not recursive.*

Thus, the set $\{F : F \text{ is a } T_i \text{-choice theorem}, i=1,\dots,m\}$ represents the category of choices or ontological commitments

Remark: choices of any agent cannot be foreseen by any mechanical tool

Discussion: knowledge and reasoning *versus* ontological commitment

We introduce a strong distinction between the T_i-*declarative knowledge* of (the fact represented by) *A* and the T_i-decision to assign an ontological status to (the fact represented by) *A,* i.e. the *ontological commitment of* T_i *on A*.

<u>Consider the following example:</u>
Agent Y is an electronic engineer and works in a factory projecting and producing components for computer hardware (transistors, microchips, ecc)
By quantum physics, Y knows that elementary particles have a *spin* (*angular momentum* of the particle, if it is considered as a material body, which is a mathematically defined entity) and in her job she applies formulas and calculations involving the *spin* of electrons.
<u>Y knows that electrons have a *spin*.</u>
<u>However, Y do not assume any ontological commitments on the *spin* of electrons,</u>
<u>i.e. Y does not include the *spin* of elementary particles in her own ontology</u>

(As wellknown, nobody can *properly see* or *directly perceive* the *spin* of an elementary particle)

Communication Acts between Agents

Definition (communication willingness)

$E=\{T_1,\dots, T_m\}$ society of agents

S set of the codes of the effectively definable space regions in the Cartesian reference system (O, *XYZ*) shared by the society E

N set of natural numbers

Then, a *communication function* $g:\{i\}\times N\times S\times\{1,\dots m\}\rightarrow N$ is defined by each proper axiom sets $A(T_i)$.

We call a *communication sentence of* T_i *to* T_j, T_i , T_j $\in E$, $i\neq j$, each formula of the form $g(i,t,v,j)= \#B$ where B is a sentence of the language of E.

A *communication willingness of* T_i is any T_i -theorem of the form $g(i,t,v,j)= \#B$. ■

The intended meaning of the T_i-communication willingness $g(i,t,v,j)= \#B$ is "The agent T_i wishes to communicate to agent T_j the information expressed by B at the time indicated by t in the place coded by v "

The set $\{G : G \text{ is a } T_i \text{ - theorem of the form } g(i,t,v,j)= \#B,\ i=1,\dots,m\}$ represents the category of communication willingness

Communication willingness too require intentionality:

Assumption. *Axioms* $A(T_i)$ *are such that the function* $g(i, \bullet,\bullet,\bullet)$ *is not recursive*.

Communication Acts between Agents

Definition (communication act)

$E=\{ T_1,..., T_m \}$ be a society of agents.

Let the sentences B and C represent the knowledge of T_i and T_j respectively, $i\neq j$.

Let $g(i,t,v,j)= \#B$ and $g(j,t,v,i)= \#C$ be the communication willingness of T_i and T_j respectively, having the same time-term t and the same space term v. Then, we say that the pair

$$\langle g(i,t,v,j)= \#B,\ g(j,t,v,i)= \#C\rangle$$

is a *communication act between* T_i *and* T_j *in the society* E.

We say that the systems $T_i +C$ and $T_j +B$ are the *extensions by communication* of agents T_i and T_j respectively. ■

The intended meaning of the communication act $\langle g(i,t,v,j)= \#B,\ g(j,t,v,i)= \#C\rangle$ is "Agents T_i and T_j communicate their knowledge to each other, at the time indicated by t and in the in the place coded by v".

From the standpoint of a meta-observer of society E there are four ontological categories: basic facts, agents, choices/ontological commitments, communication willingness.

Formal Semantic Universes of the Agents

Given the society $E=\{T_1,\ldots, T_m\}$, we wish to characterize, at the level of E, the ontological status of the knowledge of an agent.

By model-theoretic properties of *Peano Arithmetic* **PA** and its subsystems (like **PRA**) each consistent agent $T_i \in E$, has an uncountable set of non isomorphic denumerable models:

then, we treat every denumerable model of T_i as a semantic universe for it.

As the T_i *'s* form a communication society, these models, in general, are such that there may exist items of knowledge and beliefs of T_i which are false in T_j 's semantic universe, $\forall j, j \neq i$.

It is well known that the standard model **N** of **PRA** is defined so that the domain is the set N of natural numbers and the function and predicate letters of the **PRA**-language have the usual intuitive meaning.

We consider the natural interpretation of the function letters f and g of the society proper language and we get an expansion **W** of the standard model **N**. We call **W** the *standard semantic universe* with respect to the society $E=\{T_1,\ldots, T_m\}$.

In general, due to the peculiarity of its knowing and thinking, an agent T_i looses the standard model.

Absolute ontology and objective knowledge of society E

Definition (ontological compatibility inside E)

Let $E=\{T_1,..., T_m\}$ be a communication society.
Let A be a sentence of the language of E.
We say that the fact described by A is *ontologically compatible with agent* T_i, if the system T_i +A is consistent. Otherwise, it is *ontologically incompatible*. ■

Definition (absolute ontology of society E)

Let $E=\{T_1,..., T_m\}$ be a communication society. Let A be a sentence of the language of E.

We call *absolute ontology of E* the class of entities described by the set of sentences

$$AbsO(E) \equiv \{A\colon A \text{ is ontologically compatible with each } T_i \text{ of } E \text{ and } A \text{ is } \mathbf{W}\text{-true}\}.$$ ■

AbsO(E) always includes **PRA**.
The basic facts are always included in the absolute ontology of *E*.

Objective knowledge of society E

AbsO(E) allows us to define the objective knowledge:

Definition

E={T_1,..., T_m} communication society, *A* sentence of the language of *E*.

Let the sentence *A* belong to the knowledge or beliefs description of an agent T_i, $T_i \in E$.

Then we say that *A represents an objective knowledge (belief) of* T_i *in the society E* if *A* belongs to *AbsO(E)*.
■

Constraint on choices and communication willingness

A "small" absolute ontology would mean that the objective knowledge of each single agent is a "small" subset of the whole agent'sknowledge
But by definition of **W** the following holds:

Proposition Let *E={T_1,..., T_m} be a communication society. Then, the set of choices/ontological commitments and the set of communication willingness of the agents of E belong to the absolute ontology AbsO(E).* ■

In general, the whole knowledge of any agent is not included in *AbsO(E)*.

Ontology of subjective knowledge and thinking in society E

Definition (relative ontology of an agent)

$E=\{T_1,\dots, T_m\}$ communication society,

A sentence of the language of E.

Then we call *relative ontology of the agent* T_i *in the society E*, the class of entities described by the set $O_i \equiv\{A$: A is at least believed by T_i and is **M**-true in each model **M** of T_i }. ■

Definition (shared ontology)

The *shared ontology of the society* $E=\{T_1,\dots, T_m\}$ is the class of entities described by the set $Co(E)\equiv \cap_i\{O_i\}$ where O_i describes the relative ontology of T_i ■

Co(E) always includes **PRA** and the basic facts

The absolute ontology of *E* is independent of the knowledge and belief of the agents, while the shared ontology is linked at least with the beliefs of a society

Reality, Acceptability and Relevance of Communication Acts

Definition

Let **C**≡⟨ *g(i,t,v,j)=B*, *g(j,t,v,i)=D*⟩ be a communication act between T_i and T_j in the society $E=\{T_1,..., T_m\}$. Then:

i) **C** is *real* if the ontological commitments of T_i on B and of T_j on D exist, i.e. T_i proves *f(i,t,v)=B* and T_j proves *f(j,t,v)=D*;

ii) **C** is *acceptable* if both the extensions by communication T_i+D and T_j+B are consistent;

iii) **C** is *relevant* if it is real, acceptable, both the extensions by communication T_i+D and T_j+B are proper extensions (i.e. T_i does not prove *D* and T_j does not prove *B*) and both *B* and *D* belong to the absolute ontology *AbsO(E)*. ■

Point i) above states that in a relevant communication the information conveyed must be assumed into the ontology of the communicating agent.

Existence Theorem of Societies allowing Relevant Communication Acts

It is not trivial to establish that our formalized ontology for communication societies is possible:

Theorem *For each m≥2 it is possible to define a denumerable infinity of communication societies* $\{T_1,\dots, T_m\}$, *each admitting an absolute ontology AbsO(E) such that each pair of agents* $< T_i, T_j>$, $i \neq j$, *can produce relevant communication acts.* ■

The proof is neither easy nor brief, due to the following points:

- the property that each agent is endowed by an induction rule, i.e. the fact that it is at the level of *inductive rationality*, is an important feature of the general case

- the case of the pair-wise inconsistency of the agents must be considered, . i.e. the case such that for each pair $< T_i, T_j>$, $i \neq j$, at least a sentence A exists such that T_i proves A and T_j proves $\neg A$ (at least one fact exists on which T_i and T_j have opposite opinions)

- the axiomatization of the non recursive choice function f and of the communication function g must be explicitly given for each T_i

Existence Theorem of Societies allowing Relevant Communication Acts

The proof of a corresponding theorem in a different context, for societies $\{T_1,...,T_m\}$ of interacting game-players can be found in:

Benassi-Gentilini, *Paraconsistent Provability Logic and Rational Epistemic Agents*, ***Paraconsistent Logic with No Frontiers***, J.Y. Beziau, W.A. Carnielli eds., Elsevier, 2006, pp 189-226

Example

$E \equiv \{ T_1, T_2\}$ communication society

Suppose that proper axiom set $A(T_1)$ is such that T_1 proves $\neg Con(T_2)$, i.e. $Pr_{T2}(0=1)$ i.e. that T_2 is inconsistent, that is T_1 knows (thinks) that T_2 is "crazy".
By definition of the society E each T_i is consistent. Then, $\neg Con(T_2)$ is *ontologically compatible* with both T_1 and T_2 (by Provability Logic properties) but *it is false in the standard semantic universe* **W** of *E, even it is true* in each model of the infinite set of the non isomorphic models of T_1 (which are the subjective semantic universes of T_1)
Thus, $\neg Con(T_2) \notin AbsO(E)$, the absolute ontology of *E*, and, moreover, $\neg Con(T_2)$ *is not objective knowledge* of T_1 in *E*.
Suppose that, however, T_1 wishes to ontologically persist in its conviction, i.e. that axioms in $A(T_1)$ are such that T_1 proves :

$$f(1, 3, 1237) = \#\neg Con(T_2)$$

where *f is the choice function*, i.e. that T_1 at time t=3, in the place coded by 1237, assumes the *ontological commitment* on the fact that agent T_2 is crazy.

Example

Note that $f(1, 3, 1237) = \#\neg \mathrm{Con}(T_2)$ belongs to *AbsO*(*E*): the idea $\neg \mathrm{Con}(T_2)$ of T_1 is (objectively) wrong, but it is (objectively) true that it assumes it as an ontological commitment.

Suppose that T_2 has analogous opinions and analogous ontological commitments about T_1, that is T_2 proves $\neg \mathrm{Con}(T_1)$ and so on, and that axioms $A(T_1)$ and $A(T_2)$ allow the following *communication willingness*:

T_1 proves $g(1, 5, 2617,2) = \#\neg \mathrm{Con}(T_2)$

T_2 proves $g(2, 5, 2617,1) = \#\neg \mathrm{Con}(T_1)$

and consider the *communication act*:

$\mathbf{C} \equiv < g(1, 5, 2617,2) = \#\neg \mathrm{Con}(T_2),\ \ g(2, 5, 2617,1) = \#\neg \mathrm{Con}(T_1) >$

Then we can say that:

C is *real* , due to the ontological commitments of the agents

C is *acceptable*, since by Provability Logic properties $T_k + \neg \mathrm{Con}(T_k)$ is consistent

C is not *relevant* , since even if $T_k + \neg \mathrm{Con}(T_k)$ is a proper extension, $\neg \mathrm{Con}(T_1)$ and $\neg \mathrm{Con}(T_2)$ do not belong to the absolute ontology AbsO(*E*) of the society

Work in progress

1) Exploring the paraconsistent case i.e. the information exchange between paraconsistent cognitive agent.
Paraconsistency allows to face the situations where relevant information is exchanged which contradicts the ideas of the receiver, i.e. such that the extensions by communication T_i+D and T_j+B *are not* classically consistent.
This requires suitable systems of paraconsistent Arithmetic, endowed with a paraconsistent Provability Logic. Such tools are provided, e.g., in the following

Gentilini: *Paraconsistent Arithmetic with a Local Consistency Operator and Global Selfreference*, ***Proceedings of Logic, Model, Computer Science 2006***, Camerino, April 2006, «Electronic Notes in Theoretical Computer Science»

2) Further investigation on the formalization of *intentionality*, both in selfreference process and in communication process of a cognitive agent

3) Formalization, through the presented approach, (inside a formal ontology framework) of some problems in philosophy of mind

APPENDIX

The expressiveness of the paraconsistent setting

Interactive societies of agents

In many relevant applications, in order to have significant societies of (communicating) agents we must impose that *agents may disagree on at least one statement*, and so we formally define an *interactive society*:

Definition The set $\{ \mathbf{T}_1,\ldots,\mathbf{T}_n \}$ of paraconsistent rational agents is an *interactive society* if for each pair $\mathbf{T}_i$, $\mathbf{T}_j$, $i \neq j$, at least one non Δ_0 -sentence A exists, such that $\mathbf{T}_i$ proves $\vdash A$ and $\mathbf{T}_j$ proves $\vdash \neg A$.

Paraconsistent Epistemic Interaction

Paraconsistency allows to express strong knowledge evolutions of epistemic interacting agents in a society. That is, we can define a kind of $\mathbf{T}_i$-declarative knowledge of $\mathbf{T}_i$ – contradicting choices or opinions that belong to a different agent $\mathbf{T}_j$, without trivializing $\mathbf{T}_i$:

Definition (Paraconsistent epistemic interaction) Let $\{\mathbf{T}_1,\ldots,\mathbf{T}_n\}$ be an interactive society of paraconsistent rational agents at the level of the inductive rationality. Let A_i be a $\mathbf{T}_i$ –theorem and let A_j be a $\mathbf{T}_j$ –theorem $i \neq j$. Let

$$\mathbf{T}_i * = \mathbf{T}_i + A_j \qquad \text{and} \qquad \mathbf{T}_j * = \mathbf{T}_j + A_i.$$

Then, the pair

$$< \mathbf{T}_i *, \mathbf{T}_j * >,$$

is a *paraconsitent epistemic interaction* if $\mathbf{T}_i *$, $\mathbf{T}_j *$ are not trivial and A_i, A_j are not Δ_0 -formulas.

The epistemic interaction is *relevant* if we have: $\mathbf{T}_i *$ proves a sequent S and $\mathbf{T}_j *$ proves a sequent L such that

$$\mathrm{S}, \mathrm{L} \in \{ \vdash A_i \wedge \neg A_i \quad \vdash A_j \wedge \neg A_j \}.$$

Relevant epistemic interaction

The epistemic interaction is *relevant* if we have :
$\mathbf{T}_i$ * proves a sequent S and $\mathbf{T}_j$ * proves a sequent L such that S,L $\in \{ \vdash A_i \wedge \neg A_i \quad \vdash A_j \wedge \neg A_j \}$

The relevant epistemic interaction expresses the actual updating in the agent's knowledge when it establishes, through interacting, new opinions and conjectures on the state of the world.

We remark that the classical setting does not allow such a direct and powerful notion of knowledge increasing through other's opinions, ideas or actions. In a society of *classical* agents, the systems $\mathbf{T}_i$ *, $\mathbf{T}_j$ * are generally trivial.

References

[1]Benassi C., Gentilini P. (2006), *Paraconsistent Provability Logic and Rational Epistemic Agents*, in *Paraconsistent Logic with No Frontiers*, J.Y. Beziau and W.A. Carnielli eds., Elsevier, in print.

[2] Benassi C., Gentilini P. (1999), *Common Knowledge and Common Rationality Through Provability Logic*, in H. De Swart ed., *Tutorials & Position Papers - LGS '99*, Faculty of Philosophy, University of Tilburg, Tilburg.

[3] Bottazzi E., Catenacci C., Gangemi A., Lehmann J. (2006), *From Collective Intentionality to Intentional Collectives: an Ontological Perspective*, Cognitive Systems Research, 7 (2, 3): 192-208.

[4] Buss S.R. (1998), *First Order Proof Theory of Arithmetic*, in S. R. Buss ed., *Handbook of Proof Theory*, Elsevier, Amsterdam, 79-147.

[5] Ferrario R., Oltramari A. (2004), *Towards a Computational Ontology of Mind*. in A.C.Varzi, L.Vieu eds, *Formal Ontology in Information Systems. Proceedings of FOIS 2004*, pp 287-297, IOS Press, Amsterdam.

[6] Fodor J.A. (1975), *The Language of Thought*, Thomas Crowell, New York.

References

[7] Gentilini P.(1999a), *Proof-Theoretic Modal PA-Completeness I: A System-Sequent Metric*, Studia Logica, 63, 27-48.

[8] Gentilini P. (1999b), *Proof-Theoretic Modal PA-Completeness II: The Syntactic Countermodel*, Studia Logica, 63, 245-268.

[9] Gentilini P. (1999c), *Proof-Theoretic Modal PA-Completeness III: The Syntactic Proof*, Studia Logica, 63, 301-310.

[10] Guarino, N. (1998), *Formal Ontology and Information Systems*, in N. Guarino ed., *Formal Ontology in Information Systems. Proceedings of FOIS 98*, IOS Press, Amsterdam, 3-15

[11] Japardize, G. and De Jongh D. (1998), *The Logic of Provability*, in S. R. Buss ed., *Handbook of Proof Theory*, Elsevier, 476-546.

[12] Kaye R. (1991), *Models of Peano Arithmetic*, Clarendon Press, Oxford.

References

[13] Masolo C., Gangemi A., Guarino N., Oltramari A., and Schneider L. (2003), *Ontology Library (final)-Library of Foundational Ontologies*. WonderWeb-Ontology infrastructure for the Semantic Web Deliverable D18.

[14] Quine, W.V.O.(1969), *Ontological Relativity and Other Essays*, Columbia University Press, New York.

[15] Shoenfield, J.R. (1967), *Mathematical Logic*, Addison Wesley, Reading (Ma).

[16] Smorynski C. (1985), *Self-Reference and Modal Logic*, Springer-Verlag, Berlin.

[17] Takeuti G. (1987), *Proof Theory*, North Holland, Amsterdam.