

## IL COGNITIVISMO GIURIDICO DI GIOVANNI SARTOR Francesco Romeo\*

Thus, we may conclude by affirming our faith  
in the *objectivity* of rationality.  
Giovanni Sartor

### 1. Il ragionamento giuridico: un approccio cognitivo al diritto<sup>1</sup>

Questo è il titolo del quinto volume dell'imponente "*Treatise of Legal Philosophy and General Jurisprudence*", che ha visto la luce nel 2005; opera d'inizio secolo e destinata sicuramente a fungere da pietra miliare per la scienza del diritto. In questo scritto mi propongo di analizzare alcune delle tesi avanzate da Sartor; la scelta è motivata dalla provenienza di Sartor dalle ricerche proprie dell'intelligenza artificiale e dell'informatica giuridica in senso stretto, che dimostrano in questo volume - che ben potrebbe essere chiamato *Summa Intellegentiae Facticiae* - la loro forza dirompente nell'ambito della scienza giuridica, alla quale appartengono.

Il rimprovero più ricorrente, che il giurista muoveva - negli anni addietro - alla ricerca sull'intelligenza artificiale applicata al diritto, era di non saper cogliere la specificità del diritto, la particolarità della scienza giuridica rispetto alle altre scienze, qualificando così queste ricerche come estranee al settore giuridico, più propriamente appartenenti all'ambito d'altre scienze, più o meno applicate. Corre quest'anno il cinquantenario dell'intelligenza artificiale<sup>2</sup> ed in questa occasione viene pubblicato il volume di Sartor che dimostra vero il contrario, quanto di nuovo portino questi studi alla teoria generale del diritto e come essi

---

\* Università G.d'Annunzio, Chieti-Pescara.

<sup>1</sup> G.Sartor, *Legal Reasoning, A Cognitive Approach to the Law*, in: *A Treatise of Legal Philosophy and General Jurisprudence*, a cura di E. Pattaro, Springer, Heidelberg/Dordrecht 2005.

<sup>2</sup> La fondazione di quest'ampio campo di ricerche si fa risalire, convenzionalmente, al 1956, data del Dartmouth Summer Research Project on Artificial Intelligence.

siano una strada obbligata nella comprensione degli ordinamenti giuridici, non solo contemporanei.

Uno dei più rilevanti punti innovativi è l'unione, in modo non contraddittorio e nuovo rispetto alle passate teorizzazioni, dei due modelli di fondazione del diritto: il modello giusnaturalista e quello positivista. È questo punto che trovo di particolare interesse, sia per la netta posizione antinaturalista di Kelsen<sup>3</sup> ed in genere della teoria generale del diritto, difficilmente risolvibile all'interno della scienza del diritto kelseniana, sia per l'unione, operata da Sartor, della scienza del diritto con altre scienze, particolarmente con le scienze cognitive e gli studi sulla mente, ma anche con altri rami della scienza, assai vivaci nell'ultimo scorso del ventesimo secolo, quali, ad esempio, la teoria dei giochi, l'analisi economica e le teorie evoluzioniste. Si sta realizzando così proprio quello sviluppo della scienza del diritto che Kelsen stesso poneva come scopo alla teorizzazione di una 'dottrina pura'<sup>4</sup>: la creazione di una scienza del diritto, capace di svilupparsi autonomamente, senza i vincoli derivanti dalle contingenti volontà politiche, al governo o all'opposizione, rivoluzionarie o conservatrici, ognuna delle quali non vede nel diritto altro che lo strumento per la realizzazione dei propri disegni. Un'espansione della scienza giuridica quindi, non un'invasione di campo da parte di altre scienze<sup>5</sup>, né una negazione della sua specificità ed autonomia.

Nei cinquanta anni trascorsi dal ritrovo di scienziati che diedero origine alle ricerche sull'intelligenza artificiale si è verificato un mutamento particolarmente rilevante per la scienza giuridica e non solo per essa: questi studi hanno fornito la possibilità di testare sull'elaboratore le diverse ipotesi che venivano proposte sul ragionamento umano, riproducendolo, ivi compreso il ragionamento giuridico. Alcune ipotesi si sono rivelate vere, altre false; ciò che alcune teorie spiegavano sul diritto non era tutto, non era sufficiente per riprodurre sulla macchina il ragionamento giuridico. Ad esempio, se si suppone che il ragionamento giuridico sia esclusivamente un'operazione

---

<sup>3</sup> La Dottrina pura kelseniana dichiara fin dall'origine il suo scopo di purificare la scienza giuridica dalla 'contaminazione' delle altre scienze che studiano la natura, ivi compresa la sociologia che può avere come oggetto di studio il diritto come fatto sociale, non la norma come significato, v. H. Kelsen, *Lineamenti di dottrina pura del diritto*, Einaudi, Torino 1952, pp.47-55.

<sup>4</sup> H. Kelsen, nella prefazione dei *Lineamenti cit.*, pp.44-45.

<sup>5</sup> C'è chi ha definito queste invasioni di campo, nei confronti della scienza giuridica da parte di altre discipline, 'giochi a somma zero', vedi: J.M. Balkin, *Interdisciplinarity as Colonization*, Wash. & Lee L. Rev., 1966, p.949.

mentale basata su simboli, o significati oggettivabili, e su operazioni logiche su di essi, allora un elaboratore programmato adeguatamente dovrebbe riuscire a condurre a termine un ragionamento così come un qualsiasi giurista: ma così non è stato. Ugualmente, se si ritiene che il giurista operi, mentalmente, in modo esclusivo su informazioni provenienti da atti appartenenti all'ordinamento giuridico ed individuabili secondo un criterio formale di validità, allora una banca dati contenente tutti questi atti e regole di inferenza dovrebbe essere in grado di trovare la soluzione adatta: anche così non è stato. E' quindi essenziale conoscere quali parti del ragionamento giuridico manchino all'elaboratore, perché, essendo esse parte del ragionamento giuridico, non solo sono necessarie per riprodurlo su elaboratore, ma devono essere prese in considerazione da qualsiasi teoria sul diritto. A mio avviso è sulla scorta di queste ricerche, dei successi e dei fallimenti di questi tentativi ed ipotesi, che Sartor giunge ad elaborare una nuova concezione della scienza giuridica e del suo oggetto di studio, il diritto, dimostrando, là dove necessario, la falsità o parzialità di molte ipotesi correnti. Un'ulteriore particolarità, sempre derivante da quest'approccio scientifico, risiede nel modo in cui Sartor arriva alle conclusioni: non criticando o confutando questa o quella teoria, bensì costruendo - partendo dai 'dati di fatto' della mente - la teoria che ha una maggior portata esplicativa su questi e solo secondariamente, ed ove strettamente necessario, confutando le altre teorie sul diritto. Tante confutazioni si trovano solo 'tra le righe', o meglio nei fatti mentali esposti da Sartor, non esplicitamente enunciate.

## 2. Le premesse

Sartor dichiara *in limine* il fine di voler giungere ad una riunificazione delle diverse dottrine che si sono divise, nei secoli passati, sulla specifica concezione del ragionamento giuridico, giustificando, sulla diversa rappresentazione di questo, le differenti concezioni del diritto stesso, nel senso delle varie opposizioni giusnaturalismo - giuspositivismo, giurisprudenza dei concetti - giurisprudenza degli interessi, formalismo - realismo, ecc.. Il punto unificante, che intende dimostrare valido, è costituito dall'asserzione che "il ragionamento giuridico può essere concepito come un'applicazione di una più vasta competenza umana, cioè: la cognizione pratica o razionalità pratica", vale a dire l'abilità di

processare informazione al fine di giungere ad appropriate determinazioni<sup>6</sup>.

Non è possibile qui dare conto esaustivamente della teoria di Sartor, ma ne riteniamo utile una breve sintesi, al fine di verificarne la portata teorica nell'ambito giuridico. Per favorire oggettività e chiarezza, separo la parte espositiva della teoria di Sartor dal commento e dalle eventuali notazioni su di essa, che pongo nell'ultimo paragrafo; quando invece sono presenti nell'esposizione della teoria di Sartor esse sono sempre notate come tali. È impossibile in poche pagine dare un resoconto esatto del pensiero dell'autore, cerco comunque di estrapolare dal testo in modo il più possibile attinente alla lettera.

### 3. Mente e cognizione

Sartor trae dalle scienze cognitive, ma a mio avviso soprattutto dall'intelligenza artificiale, la concezione della mente umana come un insieme di competenze diverse che, unitamente, conducono l'agente al volere ed all'atto di realizzazione di ciò che l'agente vuole. In questa rappresentazione un agente cognitivo, oltre a possedere cognizione implicita, possiede anche una cognizione esplicita, cioè "*conoscenza rappresentata nell'agente in modo tale da essere accessibile allo stesso agente*". Sartor presuppone uno "stato interno dell'agente preposto specificamente alla rappresentazione della conoscenza, in modo tale che l'agente possa utilizzarla come base per i suoi ulteriori processi interni o comportamenti esterni". In questo modo, sempre seguendo quest'impostazione, si sposterebbe l'oggetto di studio dalla relazione inputs esterni-comportamento osservabile, propria del behaviorismo, allo studio dei modi in cui l'informazione viene processata internamente, oggetto di studio proprio della psicologia cognitiva. Ciò conduce Sartor a differenziare diverse *funzioni cognitive svolte da diversi organi cognitivi*, sia nella cognizione epistemica che pratica: la mente può essere concepita come un set di organi mentali connessi, il cui lavoro è

---

<sup>6</sup> "Our contribution will consist in an attempt to provide cognitive foundations not only for conclusions pertaining to natural law, but also, as we shall show in the following pages, for conclusions concerning positive law. For this idea too, one may find illustrious predecessors, like Aquinas, who famously said that human reason *from the precepts of the natural law [ . . . ] needs to proceed to the more particular determination of certain matters. These particular determinations, devised by human reason, are called human laws.* (Aquinas, *Summa Theologiae*, I-II, q. 91 a. 3, c)." G.Sartor, *op. cit.*, § 3.3.5..

predisposto geneticamente. Sartor fa sua l'asserzione di Pinker<sup>7</sup> in base alla quale la mente è ciò che fa il cervello e ciò che fa il cervello è organizzato in moduli o organi mentali, ognuno specializzato in un determinato campo d'azione; la logica base d'ogni modulo è determinata geneticamente. Tra l'altro Sartor osserva che, se si vuole comprendere o prevedere il comportamento delle persone, è necessario presupporre che esse posseggano stati cognitivi, cioè che passino da uno stato cognitivo ad un altro in accordo con la razionalità. Nello specifico, ciò è necessario per comprendere molti fenomeni sociali, particolarmente il diritto<sup>8</sup>.

La cognizione viene distinta in due tipi: epistemica e pratica. La cognizione epistemica produce nuovi stati epistemici partendo dai precedenti stati, ed è la capacità di formare stati mentali consistenti nella rappresentazione del mondo dell'agente. Esistono due tipi di stati epistemici: i percetti<sup>9</sup>, consistenti negli stati mentali causati dalla percezione, e le credenze (*beliefs*)<sup>10</sup>, consistenti nell'accreditare (*endorse*) una proposizione come premessa dei propri ragionamenti e delle proprie azioni<sup>11</sup>.

---

<sup>7</sup> S. Pinker, *How the Mind Works*, Penguin, Londra 1999, p.21, citato nel testo.

<sup>8</sup> Qui il richiamo di Sartor è a S. Haack, *Evidence and Inquiry*, Oxford, Blackwell 1993, pp. 170 ss., allo storicismo tedesco, W. Dilthey, *Selected Works Volume I: Introduction to the Human Sciences*, Princeton N. J., Princeton University Press 1991, nonché alla sociologia di Max Weber, *The Theory of Social and Economic Organization*. Ed. T. Parsons. Trans. A. M. Henderson and T. Parsons, Glencoe Ill., Free 1947, pp. 87 ss., citati nel testo.

<sup>9</sup> "Percepts are mental states that are caused by the various mechanisms of perception, which are activated when some input is provided to the agent's sensors [...]. Perception may concern the external environment (heteroception: sight, hearing, touch, smell, and taste) but also the state of the agent's body (proprioception: the perception of the position and movements of one's limbs)". G. Sartor, *op. cit.*, § 1.2.3.

<sup>10</sup> G. Sartor, *op. cit.*, § 1.2.3. in riferimento a K. Lehrer, *Knowledge*, Routledge, Londra 1990, p.36: "*to believe a proposition means: to adopt it as a basis for further attitudes, choices and behaviour, i.e., to consider the information expressed by it as correct and proceed accordingly in thoughts, and actions.*"

<sup>11</sup> "So we shall say that one believes a proposition, not only when one is absolutely certain about that proposition, but also when one is aware that one's evidence for that proposition may be overridden by evidence to the contrary [...]. We also speak of a belief in a proposition when one only adopts the

#### 4. Il ragionamento pratico

Sartor delimita in modo sinonimico ragione<sup>12</sup> e razionalità<sup>13</sup> in diversi luoghi del testo, è forse possibile sintetizzare una definizione, sensu Sartor, nel modo seguente: 'ragione o razionalità sono un organo mentale la cui funzione consiste nel modo appropriato di processare l'informazione, tramite il ragionamento, il suo contenuto sono le ragioni cioè gli stati mentali'.

La cognizione pratica è la capacità di formare stati mentali, così come l'epistemica, ma questi stati mentali, invece di semplicemente rappresentare il mondo dell'agente, fungono da guida al processo deliberativo dello stesso, che porta l'agente a determinare il suo comportamento. Gli stati mentali della cognizione pratica vengono chiamati stati conativi. Così come il ragionamento epistemico consiste nel muovere dalle attuali proprie credenze e stati percettivi verso nuove credenze, il ragionamento pratico viene definito come la procedura per acquisire (o abbandonare) stati conativi, consistente nel muovere dai propri stati epistemici e conativi verso nuovi stati conativi<sup>14</sup>.

Il volere è l'ultimo di una serie di quattro differenti tipi di stati conativi dell'agente: preferenza o simpatia, desiderio o fine, intenzione, volere.

La preferenza o simpatia (*liking, preference*) è l'individuale posizione dell'agente a favore o contro qualcosa.

---

propositions as a *hypothesis*, i.e., as a premise for one's further epistemic inquiries, but not (yet) for one's action". G. Sartor, *op. cit.*, § 1.2.3..

<sup>12</sup> "We shall just focus on *reason* or *reasoning*, which we tend to view as a *mental organ*, like Marr's perception modules or Chomsky's language faculty. [...] Our particular perspective—approaching reason as a mental organ, coupled with viewing reasons as mental states—allows us to set aside the distinction between *guiding reasons* and *explanatory reasons*." G. Sartor, *op. cit.*, § 1.2.2..

<sup>13</sup> "Practical rationality is a fundamental aspect of *rationality*, by which we mean the appropriate way of processing information through reasoning. As we shall see, practical rationality complements *theoretical rationality*, the other fundamental aspect of rationality." G. Sartor, *op. cit.*, § 1.

"In the following, we shall tend to present rationality[...] as a 'mental organ' that belongs to the individual mind, to the individual psychology. According to this psychological approach the inputs of rationality—its reasons—will be constituted by certain mental or *cognitive* states of the individual reasoner." G. Sartor, *op. cit.*, § 1.2.2..

<sup>14</sup> G. Sartor, *op. cit.*, § 1.2.4..

Il desiderio o fine è una preferenza specificata in un determinato scopo.

L'intenzione è uno stato della mente dell'agente che è giunto alla determinazione che una certa azione debba venire posta in essere <sup>15</sup>.

Il volere è un impulso a porre in essere un'azione<sup>16</sup>.

Lo schema completo di razionalità pratica vede quindi un ragionamento che dalle preferenze individuali porta ai desideri, da questi alle intenzioni e quindi al volere. La razionalità pratica è rappresentata come una tecnica che permette all'agente di raggiungere stati conativi più elevati realizzando quelli di più basso livello: realizzare il volere significa implementare le proprie intenzioni e quindi raggiungere i propri fini, con ciò soddisfare i propri desideri e con ciò adattare il mondo alle proprie preferenze. Questo spiega perché la funzione di ogni stato conativo sia quella di produrne uno di più basso livello<sup>17</sup>.

---

<sup>15</sup> "The third conative attitude consists in having *intentions*. By an intention we mean in general the state of mind of an agent who has determined that a certain action is to be executed, possibly under certain conditions. As we shall see in the following, by an *instruction* we mean a specification of an action to be performed, under given conditions. Therefore we may also say that an intention is the adoption or the endorsement of an instruction. Adopting intentions is the way in which we shall store the plans we have adopted, and keep ready for executing them: Instructions contained in chosen plans provide the content to our intentions. Thus, an agent having formed an intention is *committed* to implement the corresponding instruction, since this is required by the way in which practical rationality works: One may rationally withdraw one's intention, but it would be irrational for one not to implement one's intention, while having it. For example, when I select the plan of going to the ice cream shop round the corner (rather than to a better shop, further away), I adopt the intentions to go there and buy an ice cream. I may withdraw this intention (since there is something more important I must do), but it would be irrational for me both to keep the intention and not to implement it (at least when there is no conflict with stronger intentions)". G. Sartor, *op. cit.*, § 1.3.1..

<sup>16</sup> "A want is an impulse towards performing an action, which the agent feels when the action needs to be performed. For instance, when I look at my watch and I see that the time has come for getting my ice cream (according to my intention), I feel the impulse (want) to get up and go out [...]. When we get to the level of wants, rationality must abandon the scene. There is no time for reasoning now, and one should just let one's wants be transformed into actions". G. Sartor, *op. cit.*, § 1.3.1..

<sup>17</sup> "By realising one's wants, one implements one's intentions; by implementing one's intentions one achieves one's goals; by achieving one's

Sartor schematizza l'intero procedimento come segue<sup>18</sup>:

**"Reasoning schema: *Desire adoption***

- (1) liking A; AND
- (2) believing that A can be achieved
- IS A REASON FOR
- (3) desiring A

[...]

**Reasoning schema: Teleological inference**

- (1) desiring A; AND
- (2) believing that plan B is a sufficiently good way of achieving A
- IS A REASON FOR
- (3) intending to achieve A; AND
- (4) intending to perform all instructions in B in the appropriate combination

[...]

**Reasoning schema: *Want adoption***

- (1) having the intention that [now, I shall do A]
- IS A REASON FOR
- (2) wanting to do A"

E' da notare che l'inferenza teleologica, che si riferisce all'intenzione, non conclude solo con l'avere l'intenzione di ottenere A, bensì contemporaneamente con l'intenzione di performare convenientemente tutte le istruzioni contenute nel piano B: accettare l'ottenimento di A significa dunque accettare contemporaneamente le istruzioni che ne permettono l'ottenimento: le istruzioni sono il contenuto dell'intenzione. Il modello di Sartor è caratterizzato anche da questa scelta, che deriva da alcune impostazioni delle ricerche di Intelligenza Artificiale<sup>19</sup>: "We

---

goals one satisfies one's desires, and by satisfying one's desires one adapts the world to one's likings. This explains why the function of each conative state consists in producing, under appropriate conditions, corresponding lower level conative states." G. Sartor, *op. cit.*, § 1.3.2..

<sup>18</sup> G. Sartor, *op. cit.*, § 1.3.2..

<sup>19</sup> P. R. Cohen and H. J. Levesque, *Intention is Choice with Commitment*, in: *Artificial Intelligence*, 42, 1990, pp. 213-61; C. Castelfranchi, *Commitments:*



use the term *instruction* to denote the content of an intention. By an *instruction*, we mean a cognitive structure, which is characterised by the function of specifying the content of an intention: The instruction indicates what needs to be done, by whom and in what circumstances, to fulfil the corresponding intention.

We might also say that instructions are *imperatives*, to emphasise that they are meant to direct action, but it must be clear that here as an "imperative" we do not mean a speech act (the act of commanding), but rather the content of a mental state (the intention)<sup>20</sup>. Notiamo che la specifica caratteristica dell'intenzione è quella di essere condizionata, o costituita, non solo da informazione relativa all'agente cognitivo ma anche da informazione esterna ad esso, che l'agente accetta nei suoi nessi di causa ed effetto, nessi che quindi condizioneranno l'azione dell'agente. Un agente razionale non è un idealista sognatore, e quindi l'accettazione dei nessi di causa ed effetto, che regolano la realtà nella quale l'agente deve operare, diventa condizione dell'agire razionale. Ci soffermiamo solo un paio di righe ad osservare che in questo modo l'oggetto di studio, la mente, non è più qualcosa di isolato dal mondo esterno, avente informazioni sue proprie e funzionante in un determinato modo a lei caratteristico; in questo momento la mente dell'agente (meglio sarebbe il suo cervello) ed il mondo esterno si uniscono, le informazioni del mondo esterno ed i suoi nessi di causalità diventano informazioni relative alla mente dell'agente. È una conseguenza della scelta di Sartor di eleggere a punto di partenza la mente e non il cervello, e, con operazione non priva di interpolazioni, di unificare funzionalmente mente e cervello (v. *supra* nota 7).

Sartor distingue quindi tra istruzioni comportamentali ed istruzioni cognitive, queste ultime portano non direttamente ad un comportamento bensì all'adozione di un determinato stato mentale<sup>21</sup>.

---

*From Individual Intentions to Groups and Organizations*, in: *Proceedings First International Conference on Multi-Agent Systems*, V. Lesser ed., AAAI Press and MIT Press, San Francisco Cal., 1995, pp. 41–8; citati nel testo.

<sup>20</sup> G. Sartor, *op. cit.*, § 1.4.1..

<sup>21</sup> "We may say that the first type of instruction expresses a *behavioural policy*, while the second type express a *mental policy*. Thus, in the second type of instruction, the required action consists in the adoption, or in the acceptance, of a mental state." G. Sartor, *op. cit.*, § 1.4.3.

## 5. Ragionamento conclusivo e defeasible

Vengono poi distinti due tipi di schemi di ragionamento: quello conclusivo e quello defeasible<sup>22</sup>.

Uno schema di ragionamento conclusivo indica la transizione che un agente può operare tra premesse e conclusioni senza riguardo ad ulteriori informazioni<sup>23</sup>. Il ragionamento conclusivo è truth-preserving, se le premesse sono vere allora anche le conclusioni lo sono.

L'osservazione, che non è però questa peculiarità a caratterizzare ogni forma di ragionamento razionale o razicinativo<sup>24</sup>, permette a Sartor un ulteriore importante passo nell'allargamento del concetto di razionalità: il ragionamento conclusivo non delimita il campo della logica, che comprende quindi anche il ragionamento non truth-

---

<sup>22</sup> Non traduciamo il lemma inglese, di origine giuridica, che significa: "suscettibile di estinzione, sottoposto a condizione, risolubile o revocabile" F. de Franchis, *Dizionario Giuridico - Law Dictionary*, Giuffrè, Milano 1984; nella prassi dei *Torts* del *Common Law* il lemma viene usato per indicare un atto sottoposto a condizione o a termine. La ricerca in Internet dell'associazione "defeasible + logica + \*" ha dato i seguenti risultati: plausibile 20, annullabile 2, rivedibile 2, ricusabile 2, invalidabile 1, ritrattabile 1, revocabile 1, confutabile 0. Occorre però osservare che nella letteratura anglosassone si distingue tra *plausible* e *defeasible inference*, a seconda che le premesse o il metodo siano certi o incerti, nel testo però il lemma *plausible* viene utilizzato solo nei significati di senso comune e non viene differenziato da *defeasible*; preferiamo quindi non tradurre il lemma.

<sup>23</sup> "A *conclusive* reasoning schema indicates a transition that can operate regardless of any further information the agent possesses, as long as the reasoner instantiates the preconditions in the schema: Whenever one endorses the reason of the schema, one may safely endorse its conclusion. [...] **Definition 2.2.1** *Conclusive reasoning schema*. A reasoning schema R is conclusive if one can always adopt R's conclusions while endorsing R's premises (and one should never reject R's conclusions while endorsing R's premises)." G. Sartor, *op. cit.*, § 2.2..

<sup>24</sup> "Truth-preservation is a very useful and important property, but it does not characterise all rational reasoning patterns. Therefore, it does not delimit the scope of logical reasoning, if by *logic* we mean rational reasoning, or rational ratiocination. It is true that many authors tend to limit logic to truth-preserving reasoning: They view logic as having the specific function of providing truth-preserving ways of reasoning. Moreover, many logicians refer to truth-preservation by using the word *valid*. [...] Correspondingly, it is also said that logic is the study of valid reasoning." G. Sartor, *op. cit.*, § 2.2.1..

preserving, cioè il ragionamento defeasibile "This allows us to avoid qualifying all defeasible inferences as being 'invalid': Though they are not truth-preserving (it may happen that their preconditions hold, but their conclusions fail to be true), defeasible reasoning schemata are indeed "valid" or "sound" forms of reasoning, in the sense of being appropriate ways of approaching certain cognitive tasks"<sup>25</sup>. Lo schema di ragionamento *defeasible*, contrariamente a quello conclusivo, può essere *defeated* dal sopraggiungere di ulteriori premesse<sup>26</sup> e non porta necessariamente all'adozione delle conclusioni pur accettando le premesse<sup>27</sup>.

**"Reasoning schema: Conclusive syllogism**

- (1) all Y 's are Z's; and
  - (2) x is Y
- is a conclusive reason for
- (3) x is Z

**Reasoning schema: Defeasible syllogism**

- (1) all Y 's are normally Z's; and
  - (2) x is Y
- is a defeasible reason for
- (3) x is Z"<sup>28</sup>

---

<sup>25</sup> G. Sartor, *op. cit.*, § 2.2.1..

<sup>26</sup> "The schema indicates a transition that only operates when one has no prevailing beliefs (or, more generally, mental states) against applying the schema or against adopting its conclusions. When one endorses the premises of a defeasible schema but has such prevailing beliefs, one should:

- refrain from adopting the conclusion of the schema, and even
- withdraw the adoption of that conclusion.

**Definition 2.2.2** Defeasible reasoning schema. A reasoning schema is defeasible if one should, under certain conditions, refrain from adopting its conclusions though endorsing its premises. " G. Sartor, *op. cit.*, § 2.2..

<sup>27</sup> Defeasible reasoning schemata license *non-monotonic* reasoning: Their conclusions may need to be abandoned when new information is available. G. Sartor, *op. cit.*, § 2.2.2..

<sup>28</sup> "When believing the premises of a defeasible schema, we are lead to endorse the conclusions of the schema, though these conclusions are not truth-preservingly implied by our premises (and they may indeed turn out to be

Sartor<sup>29</sup> evidenzia che anche il metodo scientifico ammette diversi tipi di inferenza *defeasible*: inferenza percettiva<sup>30</sup>, inferenza mnemonica<sup>31</sup>, induzione per enumerazione<sup>32</sup>, sillogismo statistico<sup>33</sup>, persistenza temporale<sup>34</sup>, e che questi tipi di inferenza rappresentano il modo più comune e naturale di adattamento ad un ambiente mutevole, in cui l'agente cognitivo, attraverso il ragionamento, acquisisce alcuni stati cognitivi provvisori per rivederli successivamente, grazie al risultato di ulteriori ragionamenti. L'importante funzione del ragionamento *defeasible* è quella di fornire la possibilità di agire anche in presenza di informazioni solo parziali e provvisorie, fino a che non vengano ottenute nuove informazioni. Un'altra importante funzione consiste nell'attivare

---

false, even when the premises hold true). However, failure to satisfy truth preservation does not entail logical faultiness." G. Sartor, *op. cit.*, § 2.2.2..

<sup>29</sup> In riferimento a J. L. Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, New York, N. Y. 1995 pp. 52ss..

<sup>30</sup> "*Perceptual inference*. Having certain perceptual contents is a defeasible reason to believe in the existence of corresponding external objects. More generally, having a percept with content  $\varphi$  is a defeasible reason to believe  $\varphi$ . For instance, having an image of a red book at the centre of my visual field is a defeasible reason to believe that there is a red book in front of me." G. Sartor, *op. cit.*, § 2.2.3..

<sup>31</sup> "*Memory inference*. Recalling  $\varphi$  is a defeasible reason to believe  $\varphi$ . For instance, my memory that yesterday I had a faculty meeting is a defeasible reason for me to believe that indeed there was such meeting." G. Sartor, *op. cit.*, § 2.2.3..

<sup>32</sup> "*Enumerative induction*. Observing a sample of  $F$ 's all of which are  $G$ 's is a defeasible reason to believe that all  $F$ 's are  $G$ 's. For instance, believing that all crows I have ever seen are black is a defeasible reason to believe that all crows are black." G. Sartor, *op. cit.*, § 2.2.3..

<sup>33</sup> "*Statistical syllogism*. Believing that [most  $F$ 's are  $G$ 's AND  $a$  is an  $F$ ] is a defeasible reason to believe that  $a$  is a  $G$ . For instance my beliefs that [most printed books have even numbered pages on their left side] and that [the volume on the top of my table is a printed book] are defeasible reason for me to believe that this volume has even-numbered pages on its left side." G. Sartor, *op. cit.*, § 2.2.3..

<sup>34</sup> "*Temporal persistence*. Believing that  $\varphi$  is the case at time  $t_1$  is a defeasible reason to believe that  $\varphi$  is still the case at a later time  $t_2$ . For instance, my belief that my computer was on the top of my table yesterday evening (when I last saw it) is a defeasible reason for me to believe that my computer is still there." G. Sartor, *op. cit.*, § 2.2.3..

un processo strutturato d'indagine, basato sull'acquisizione di conclusioni *pro-tanto*, cercare i defeaters, i defeaters dei defeaters ecc. fino ad ottenere conclusioni stabili.

È qui interessante porre in rilievo la scelta operata da Sartor di utilizzare la logica del *defeasible*, invece del calcolo delle probabilità nella rappresentazione del ragionamento su informazione incompleta. Sartor adduce tre diversi motivi per questa scelta. Il primo motivo riguarda il fatto che spesso non si possiede sufficiente informazione per attribuire una probabilità, si dovrebbe comunque farlo con una notevole dose di arbitrarietà<sup>35</sup>. Il secondo motivo riguarda il fatto che sembra avere poco senso ascrivere una probabilità all'informazione pratica<sup>36</sup>, ed il terzo riguarda la nota difficoltà degli esseri umani di ragionare seguendo corretti schemi di calcolo delle probabilità<sup>37</sup>.

Sartor compie qui una scelta ben precisa e non priva di conseguenze sull'intera costruzione della rappresentazione del ragionamento, molto chiaramente nega che l'inferenza statistica o il calcolo delle probabilità possano fornire la struttura del modello teorico, è una scelta che successivamente analizzeremo più approfonditamente.

La scelta di Sartor, che esclude il calcolo probabilistico, è comunque coerente con la scelta di considerare oggetto della razionalità solo la conoscenza liberamente accessibile all'agente, a nostro avviso, però,

---

<sup>35</sup> "The first issue concerns practicability: Often we do not have enough information for assigning numerical probabilities in a sensible way. For instance, how do I know that there is a 40% probability that John was driving and 60%probability the Mary was driving? In such a case, it seems that either we arbitrarily attribute probabilities, or, with equal arbitrariness, we assume that all alternative ways in which things may have gone have the same probability." G.Sartor, *op. cit.*, § 2.2.11..

<sup>36</sup> "The second issue is a conceptual one: Though it makes sense to ascribe probabilities to epistemic propositions, it makes little sense to assign it to practical information. What does it mean that a certain desire (goal) or intention (instruction) has a certain probability?" G.Sartor, *op. cit.*, § 2.2.11..

<sup>37</sup> "The third issue relates to psychology: Humans tend to face situations of uncertainty by choosing to endorse hypothetically one of the available epistemic or practical alternatives (while keeping open the chance that the other options may turn out to be preferable), and applying their reasoning to this hypothesis (while possibly, at the same time, exploring would be the case, if things turned out to be different). We do not usually assign probabilities, and then compute what further probabilities follow from such an assignment." G.Sartor, *op. cit.*, § 2.2.11..

nulla esclude che sia possibile 'ibridare' il modello di Sartor con un calcolo probabilistico.

## 6. La doxificazione

Adottando la teoria di Pollock<sup>38</sup> che unifica i moduli mentali dedicati ai ragionamenti defeasible epistemici e pratici, Sartor utilizza il procedimento di *doxificazione* in modo che sia possibile performare il ragionamento pratico usando le risorse del ragionamento epistemico, riformulando gli stati conativi *come se fossero stati epistemici*, e trasferendo nel ragionamento epistemico gli schemi ragionativi del ragionamento pratico<sup>39</sup>. La doxificazione consiste appunto nella riduzione del ragionamento pratico a quello epistemico. Sartor introduce, allo scopo, la nozione di *adoption-worthiness o bindingness*<sup>40</sup>:

**"Definition 3.1.1** *Adoption-worthiness (bindingness).*"

*A mental state (or a noema) is adoption-worthy, or more simply, binding, when it (or its content) deserves to be adopted."*

Ad esempio,

l'adozione diretta di un piano (vedi *supra* inferenza teleologica p. 238) comporta che:

- ◇ se l'agente ha un determinato fine (desiderio)
- ◇ e crede che un determinato piano sia una strada sufficientemente buona per raggiungerlo
- l'agente formerà l'intenzione di raggiungerlo.

---

<sup>38</sup> J. L. Pollock, *Cognitive Carpentry, cit.*, pp.277 ss..

<sup>39</sup> "We shall now consider an important option that is available to an agent possessing both types of cognition. One can perform practical reasoning by using some of the resources one has for theoretical reasoning. This happens by reformulating conative states as if they were epistemic states, and in transferring into epistemic reasoning the reasoning schemata available for practical reasoning. We call this process the *doxification* of practical reasoning." G.Sartor, *op. cit.*, § 3.

<sup>40</sup> "[...] We say that a mental state (or its content) is *adoption-worthy*, to express *cognitive optimality*. In other words, by saying that a mental state is adoption-worthy, we mean that this mental state is worthy of, deserves or requires being adopted, i.e., that we should adopt it. Said otherwise, adopting this mental state is necessary to fulfil our cognitive duty, that is, to proceed correctly in our cognitive effort, according to the requirements of rationality. For the sake of conciseness, we shall also use the word *binding*, to convey this idea." G.Sartor, *op. cit.*, § 3.1.1..

L'adozione doxificata di un piano invece comporta che:

- ◇ se l'agente crede che un fine sia adoption-worthy
- ◇ e crede che un determinato piano sia una strada sufficientemente buona per raggiungerlo
- l'agente formerà la credenza che il piano sia adoption-worthy.<sup>41</sup>

Sartor quindi asserisce che qualora si giunga alla credenza che uno stato mentale sia adoption-worhty l'agente è giustificato nell'accreditare (endorse) lo stato conativo, procedimento che chiama de-doxificazione<sup>42</sup>.

Il procedimento proposto da Sartor aggiunge un grado intermedio tra il ragionamento e l'azione, permettendo di operare razionalmente sul ragionamento pratico, distinguendo i due aspetti: quello epistemico, espresso dalla domanda 'che desideri o intenzioni o voleri ho?', da quello pratico 'che desideri o intenzioni o voleri devo avere, cioè sono adoption-worthy?'<sup>43</sup>. Ciò permette quindi la distinzione tra stati conativi doxificati che, pur essendo accreditati dall'agente, non sono razionalmente adoption-worthy e quelli che lo sono<sup>44</sup>: "one should (rationally) believe that something is obligatory, exactly when one should (rationally) intend to do it. Rational (or true, if you prefer)

---

<sup>41</sup> "In general, doxification appears to be a reasoning technique that allows one to store one's conative states in the form of beliefs and manipulate them accordingly, in order to achieve rational determinations: Rather than memorising the intention to implement a plan, one memorises one's belief that the plan is adoption-worthy (binding); rather than memorising the desire for something, one memorises the belief that something is desirable (it is a valid object for desire)." G.Sartor, *op. cit.*, § 3.1.2..

<sup>42</sup> "**Reasoning schema: De-doxification**

believing that mental state (or noema) M is adoption-worthy

is a reason for

endorsing M" G.Sartor, *op. cit.*, § 3.1.2..

<sup>43</sup> G.Sartor, *op. cit.*, § 3.1.3..

<sup>44</sup> Nota che la nozione di bindingness, vincolatività, non equivale ad obbligatorietà, il significato è dato dalla regola condizionale 'se ci si vuole attenere alla razionalità, allora il contenuto x dello stato conativo doxificato è vincolante'. Sartor la definisce 'cognitive bindingness'. Inoltre: "though we assume a perfect logical equivalence between adoption-worthiness and bindingness, we shall speak of bindingness only when self-interest is not at issue." G.Sartor, *op. cit.*, § 3.1.5..

normative beliefs only accompany rational conative states, namely, the outcomes of correct practical cognition"<sup>45</sup>.

La doxificazione funziona con ogni stato conativo:

<b>Stato conativo</b>	<b>Doxificazione</b>	<b>Cognizione doxificata</b>
preferenza	≅	bene, valore
↓		↓
desiderio	≅	desiderio dover essere
↓		↓
intenzione	≅	inferenza teleologica

su queste premesse Sartor definisce:

*credenza normativa* = credenza che doxifica uno stato conativo

*proposizione normativa* = contenuto di una credenza normativa<sup>46</sup>.

La nozione di bindingness prima esposta permette di unire gli aspetti psicologici con quelli normativi: credere che qualche cosa sia un valore porta a credere che la preferenza per quel valore sia vincolante (deve essere preferito), ed ugualmente per lo scopo (deve essere perseguito) e per l'istruzione (si deve avere l'intenzione di eseguirla). La doxificazione proposizionale può essere operata anche sulle istruzioni cognitive e rappresenta il fondamento psicologico delle regole costitutive (intese come regole non-deontiche)<sup>47</sup>, Sartor le chiama scorciatoie del ragionamento, la loro applicazione porta ai medesimi risultati ai quali porterebbero le istruzioni cognitive; Sartor suggerisce di concentrare la ricerca non sul loro status ontologico, bensì sulla loro funzione, data l'equivalenza funzionale con le istruzioni cognitive<sup>48</sup>.

<sup>45</sup> G.Sartor, *op. cit.*, § 3.1.3..

<sup>46</sup> "**Definition 3.1.2** Normative beliefs and normative propositions. A normative belief is a belief that doxifies a conative state. A normative proposition is the content of a normative belief.

As we shall see in the following, our notion of a normative belief includes believing that something is a value to be appreciated, a goal to be pursued, a factor promoting an outcome, or a binding rule, and also believing that, under appropriate circumstances, certain actions are obligatory or permitted, certain persons have certain rights or other entitlements, various legal or moral qualifications hold. The diverse contents of all such beliefs constitute what we call normative propositions." G.Sartor, *op. cit.*, § 3.1.4..

<sup>47</sup> G.Sartor, *op. cit.*, § 3.1.5..

<sup>48</sup> Per meglio chiarire riportiamo l'esempio di Sartor: "For instance, assume that Mary, a cook in a kindergarten, has to face an epidemic that is related to food infection. She asks herself: What instructions (determinations) shall I



---

adopt, out of my concern for the health of the children, to prevent the spread of the epidemic? Assume that she concludes that she should adopt the behavioural instruction that she shall never give the children any uncooked food. This practical belief is supported by her belief that cooking the food will kill the germs causing the infection. Her belief in the adoption-worthiness (bindingness) of this policy leads her to adopt that policy, that is, to form the intention that she shall never give uncooked food to the children. Mary is happy with the choice of this plan of action, which allows her to pursue her mission of feeding the children, while being sure that she is not harming them. To be able to apply her policy, however, Mary has to solve an apparently theoretical question: "When is a food uncooked?" To answer this question she may refer to linguistic knowledge and consult a dictionary to find out the meaning of "uncooked." More ambitiously, she may investigate non-empirical realities and look for uncookedness in a dimension inhabited by conceptual realities: When does a food count as being uncooked? When is it endowed with the non-empiric property of uncookedness? When does it partake to the idea of uncookedness? However, being a practically minded woman, Mary does not care about such inquires, but only about the health of her children: Since she views cooking as a way of killing germs through heat, and since she believes that all germs will be killed at 150 degrees Celsius, she concludes that the following cognitive instruction is adoption-worthy to her: [I shall accept that food is uncooked, whenever it has not been heated above 150° Celsius].

Let us now see how Mary will apply this instruction. Faced with a food that has not been heated up to that temperature, she will conclude that she shall accept that it is uncooked. This belief, according to reasoning schema *executing cognitive intention* will lead her to believe that the food is uncooked, as you can see in Table 3.1 [...]. Note that the cook's reasoning in Table 3.1 [...] can be shortcut, if she translates the proposition

I shall accept that food is uncooked, whenever it has not been heated above 150° Celsius

into the following form:

food is uncooked whenever it has not been heated above 150° Celsius

**Table 3.1: Inference with cognitive intentions**

(1) I shall accept that food is uncooked whenever it has not been heated above 150° Celsius;

(2) this dish has not been heated above 150° Celsius

-----

(3) I shall accept that this dish is uncooked

-----

(4) this dish is uncooked

### 7.1 *normative states of affairs*

In effetti la tesi di Sartor porta a risultati assai innovativi per la scienza giuridica: se ad esempio ho la credenza che un'istruzione del tipo 'devo fare l'azione A' sia binding, ciò conduce l'agente a proiettare l'istruzione nel mondo esterno e ad affermare che l'istruzione ha la proprietà di essere obbligatoria. Similmente, afferma Sartor, la proiezione nel mondo esterno delle istruzioni condizionali che crediamo essere binding le trasforma in connessioni di tipo causale (causal-like) intercorrenti tra le precondizioni fattuali e gli effetti normativi. Così come già notato nell'ambito giuridico da Kelsen<sup>49</sup>. Quindi le credenze pratiche portano l'agente a ritenere esistente un '*practical state of affairs*'

---

**Table 3.2: *Doxification of the inference with cognitive intentions***

(1) food is uncooked whenever it has not been heated above 150° Celsius;

(2) this dish has not been heated above 150° Celsius

-----  
(3) this dish is uncooked

and stores the latter proposition in her mind. This allows her to reduce the inference in Table 3.1 above to the inference in Table 3.2, which looks like an epistemic inference. Observe that the epistemic-like inference of Table 3.2 leads the cook exactly to the same result (her belief that the food is uncooked) she could achieve by using the corresponding cognitive instruction (first inferring the intention to believe that the food is uncooked, and then forming that belief).

One may want to make a step forward and say that the fact that Mary has adopted (or should adopt) the belief [if a food has not been heated above 150° Celsius, it is uncooked] constitutes a new reality (a rule, or a rule-based connection), according to which the fact that a food has not been heated above 150° constitutes its uncookedness." G.Sartor, *op. cit.*, § 3.1.6..

<sup>49</sup> "A similar type of projection also holds for legal beliefs. So, my belief that I should adopt (in legal reasoning) the instruction: one shall compensate people whom one has damaged will be paralleled by my belief in the existence of a *legal connection*, namely, a causal-like connection existing as a legal ought-to-be, as Kelsen (1960, sec. 4.b) would say (see also Kelsen 1992, sec. 11).6 According to such a connection, producing damage causes an obligation to repair: if one damages another, then one ought to compensate the damage or even one's damaging another determines one's obligation to repair the damage." G.Sartor, *op. cit.*, § 3.2.1.

corrispondente alle sue credenze<sup>50</sup>. Ad esempio la credenza di dover adottare l'istruzione 'devo aiutare mio fratello quando ha bisogno' mi fa formare la credenza che esista una connessione tra l'aver bisogno di mio fratello ed il mio obbligo di aiutarlo e da ciò 'il fatto che mio fratello abbia bisogno causa il mio obbligo di aiutarlo'.

Sartor afferma che se si accetta l'esistenza di *normative states of affairs* solo da un punto di vista funzionale, come semplice attrezzo teorico utile per migliorare le nostre abilità di ragionamento in ambito pratico<sup>51</sup>, non sorgono particolari domande e problemi. Il problema sorge invece se l'esistenza dei *practical states of affairs* viene posta sul piano ontologico, domandandosi, ad esempio, 'quali *states of affairs* rendono una credenza pratica vera?'

Non è accettabile la tesi che le proposizioni normative diventino vere con il fatto che l'agente ha il corrispondente stato conativo. Ciò viene reso evidente dall'osservazione che il cambiamento dello stato conativo, dovuto al sopraggiungere di nuove informazioni esige un nuovo *state of affairs* che falsifica il precedente<sup>52</sup>, oppure dall'osservazione che un altro agente, in possesso di informazioni diverse può ritenere vero un *normative state of affairs* diverso.

## 8. Gli stati cognitivi ottimali

Su queste basi Sartor introduce la nozione di *giustificabilità cognitiva* come *costitutiva* delle *verità normative*, dei *normative states of affairs*: "ideal practical cognition is constitutive of practical truths: Ideal practical

---

<sup>50</sup> "However, when one believes that one should adopt certain conative attitudes, one tends to assume that states of affairs exist corresponding to these attitudes, to wit, that there are *normative states of affairs*. This takes doxification one step further [...] Thus, to doxifying reasoners the fabric of the world appears to be enriched with a new dimension: the practical state of affairs corresponding to their practical beliefs. In particular, these states of affairs include the obligations and the permissions corresponding to shall and may instructions that the "doxifiers" believe to be binding (out of any of their particular concerns or roles)." G.Sartor, *op. cit.*, § 3.2.1...

<sup>51</sup> Sartor rileva che questa è la posizione di J. L. Pollock, *Cognitive Carpentry*, *cit.*, p. 275

<sup>52</sup> "Similarly, this perspective leads us to say that the proposition that some things (for instance, freedom, welfare, peace, etc.) are values holds true as long as one continues to like those things (to have a preference for them)." G.Sartor, *op. cit.*, § 3.3.1.

cognition can reach practical truths because it constitutes the very states of affairs that (doxified) practical propositions describe"<sup>53</sup>.

La conclusione di quanto prima osservato conduce al dilemma: o ammettiamo che esistono nel mondo esterno degli *states of affairs* che sono indipendenti dagli stati mentali delle persone, oppure accettiamo che non esistano, sicché ragionarvi sopra diventerebbe una vicenda completamente irrazionale e priva di senso<sup>54</sup>. Se si accettasse questa seconda posizione si accetterebbe anche che il normale agire di avvocati e giudici nei tribunali è irrazionale o assurdo, perché il loro agire è sulle credenze normative.

La soluzione proposta da Sartor è che i *practical states of affairs* vengano costituiti dall'ottimo cognitivo dell'adozione del corrispondente stato conativo, ovverosia dall'ottimo cognitivo della credenza che quello stato conativo sia binding<sup>55</sup>.

L'ottimo cognitivo o giustificabilità cognitiva<sup>56</sup> non riguarda lo stato mentale che un agente di fatto ha in un determinato momento, bensì quello che dovrebbe avere se avesse applicato perfettamente (correttamente) gli strumenti cognitivi a disposizione dell'agente. Seguono quindi le attinenti definizioni:

**"Definition 3.3.1 Inferential justifiability.**

*A mental state M is inferentially justifiable to agent j, if j would maintain M after applying perfect ratiocination to j's current mental states.*

[...]

**Definition 3.3.2 Rational justifiability.**

*A mental state M is rationally justifiable relative to agent j, if j would maintain M after applying perfect reasoning to j's current mental states*

---

<sup>53</sup> G. Sartor, *op. cit.*, § 3.3.4..

<sup>54</sup> Sartor qui respinge la tesi del realismo scandinavo sull'irrazionalità delle credenze normative. V. Pattaro 2000, Faralli 1987, citati nel testo, notiamo che un'impostazione probabilistica non si troverebbe di fronte al medesimo dilemma, ma approfondiremo il punto nelle conclusioni.

<sup>55</sup> "We believe that there is a way out, so that the existence of a normative state of affairs can be distinguished from people's current beliefs, without assuming that normative states of affairs have an existence similar to that of physical states of affairs. This is the idea that what constitutes a practical state of affairs is the *cognitive optimality* of the adoption of the corresponding conative state—or, which is the same, the cognitive optimality of the belief that this conative state is adoption-worthy." G.Sartor, *op. cit.*, § 3.3.2..

<sup>56</sup> Sartor usa sinonimicamente *cognitively optimal* e *justifiable*

(with the exclusion of any external cognitive input, but including not only ratiocination, but also heuresis and coherence evaluation).

[...]

**Definition 3.3.3 Cognitive justifiability.**

A mental state *M* is cognitively justifiable to agent *j*, when *j* would maintain *M* after a perfect inquiry starting from *j*'s current mental states."

oppure riformulando la 3.3.3

**"Definition 3.3.4 Cognitive justifiability.**

A mental state *M* is cognitively justifiable when any rational agent would maintain *M* after an optimal inquiry."<sup>57</sup>

## 9. Note d'analisi conclusive

In queste note conclusive aggiungo alcune mie osservazioni problematiche, senza peraltro negare o sminuire l'importanza della teoria di Sartor, né, soprattutto, l'avanzamento scientifico nel campo giuridico, peculiarmente nelle tecniche di oggettivazione del significato.

Pongo in premessa il mio sospetto, o pregiudizio, nei confronti di due lemmi utilizzati spesso da Sartor, razionalità e fede, che, a mio avviso, coincidono nella loro inflessibilità; io preferisco i due lemmi fiducia ed esattezza<sup>58</sup>, la prima perché ammette smentite, revisioni ed ulteriori giustificazioni su ciò che a sua volta la giustifica, ed ammette la prova dell'empiria, che costituisce l'unica oggettività pensabile e perseguibile, sia per il mondo esterno che per l'individuo e la sua soggettività, oggettivamente esistente né più né meno di un sasso. La seconda,

---

<sup>57</sup> "Therefore, a mental state is *cognitively justifiable* for a certain reasoner, if that reasoner would have that mental state, and remain in it, after having been provided with all relevant perceptual input, and having processed that input (plus the current mental states of the reasoner) according to all resources of perfect, or *optimal rationality* (there included ratiocination, heuresis, and coherence evaluation).

Since optimal rationality may be assumed to be the same for all reasoners, we can make this idea independent of the identity of the reasoner" ed in nota: "Though one may argue that even optimal cognition cannot discount different starting points: When reasoners *j* and *k* have different starting points (different current mental states) even perfect cognition would sometimes lead them to different outcomes" G.Sartor, *op. cit.*, § 3.3.2..

<sup>58</sup> Traggo spunto da un intervento di Roberto Calasso al convegno del Pen Club: *La fede la ragione e il Segretario dell'Esattezza*, pubblicato su La Repubblica del 27.04.06, pp.1 e 51.

l'esattezza, perché lascia il segno là dove è passata, sicché è sempre possibile ripercorrere il cammino all'indietro per trovare il bivio in cui si è imboccata la falsa strada, in cui si è compiuto quell'errore che l'esattezza *probabilisticamente* accetta e teorizza, cercando di minimizzarlo. Fiducia ed esattezza sono anche, per il diritto, i due rampini verso le scienze cognitive, l'esattezza, sotto forma di certezza, mi sembra essere ciò di cui abbisogni il diritto, e che gli studi delle scienze cognitive possono favorire con il loro studio dell'uomo e dei suoi processi psichici e cerebrali. La fiducia, perché è uno stato mentale di origine biologica, in sé senz'altro irragionevole se valutato nell'ottica del singolo individuo e della sua fitness, ma che costruisce quella socialità presupposta dal diritto, e la costruisce non su caratteristiche culturali che, qualsiasi esse siano, uniscono quanto dividono i gruppi di individui, bensì su caratteristiche biologiche, che, così come dice Sartor per la razionalità, agli esseri umani accade di possedere. È un guess, una scommessa, che non tiene conto di differenze culturali, si basa su di una caratteristica biologica universalmente condivisa, anche se non riesce a dar conto di se stessa se non in termini di probabilità.

Non è una differenza radicale rispetto alla posizione di Sartor, la razionalità lascia anch'essa le tracce del suo cammino ed esige esattezza, solo che essa, anche quella *defeasible*, è deterministica, pretende *anche* di concludere in modo esatto, non ammette la coesistenza dell'errore nella conclusione. Similmente per la fede: anch'essa è un atto irragionevole, come la fiducia, e unisce culturalmente gli individui in gruppi, solo che il cammino della conoscenza è interrotto dall'atto di fede, nulla, neppure l'empiria riesce a rimuoverlo, ed in questo le fedi dividono, la fiducia no. Forse il mio è solo uno scetticismo rispetto alla possibilità di completa giustificazione di quelli che Sartor chiama 'stati cognitivi ottimali'. Qui di seguito cerco di giustificare questa mia posizione sempre in confronto alla teoria di Sartor.

Nel modello teorico di Sartor la razionalità dipende, almeno in parte, da fattori non razionali e non spiegati né in questa né in altre teorie, come Sartor stesso accetta. Questo è, in realtà, un rilevante punto a favore della teoria di Sartor, un avanzamento rispetto alle tradizionali impostazioni dell'intelligenza artificiale. Sartor ritiene che il ragionamento si debba inserire dopo le risposte intuitive, testandole e raffinandole, mentre al contrario qualora il ragionamento fosse la sola o dominante forma di cognizione i risultati sarebbero estremamente poveri<sup>59</sup>. Afferma che è facile constatare quanto il ragionamento sia

---

<sup>59</sup> G.Sartor, *op. cit.*, § 1.5.1..

pienamente dipendente dalla cognizione implicita: ne sono esempio la costruzione di peretti, il riconoscimento di facce, le misurazioni ecc.. Sempre nella sua elaborazione noi non possiamo controllare i meccanismi della cognizione implicita, ciononostante ci dobbiamo fidare essendo il risultato dell'evoluzione, essa è tuttavia imprecisa e poco affidabile, ma ciò non implicherebbe che la razionalità sia sempre superiore. È vero, per Sartor, esattamente il contrario, la cognizione implicita è sia l'antecedente evoluzionistico che il prerequisito funzionale per il possesso della ragione. Sartor enuncia tre motivi<sup>60</sup>:

1. è necessaria l'efficienza dell'intuito in un ambiente dinamico
2. il fine tuning del comportamento dell'agente e l'acquisizione incrementale di nuovi schemi comportamentali sulla base dell'esperienza, che è fornita dal meccanismo dei riflessi condizionati, è necessario per il ragionamento
3. infine non sappiamo come potremmo acquisire attraverso il ragionamento o la computazione i risultati che otteniamo usando i nostri organi mentali inconsci in compiti quali la percezione, l'analogia, la comprensione del linguaggio.

La ragione s'inserisce quindi su delle informazioni già processate (ed in modo non compreso), testandole e raffinando l'output, è un ulteriore organo mentale, che offre una possibilità in più rispetto al solo processamento dell'informazione dei moduli mentali non raziocinativi, offrendo quindi un migliore adattamento alla realtà.

Concordo su quest'impostazione, ma la teoria deve tener presente questa realtà *all'interno* della teorizzazione stessa, ed a mio avviso l'unico modo per farlo è inserire l'errore come *necessariamente presente* anche nel risultato del procedimento raziocinativo. Sartor dovrebbe ammettere che tale risultato è sempre e solo un risultato *probabilisticamente* vero, così come la teoria degli errori ci insegna, e che quindi ogni *normative state of affairs*, può essere *solo* probabilisticamente un ottimo. Sartor però non accetta l'impostazione probabilistica<sup>61</sup>.

Un secondo rilievo riguarda la scelta della mente e non del cervello come causa o origine della razionalità.

---

<sup>60</sup> "Correspondingly, a rational agent cannot be an agent all of whose behaviour is determined by reasoning, but rather an agent that adds a reasoning capability on the top of the faculties for implicit cognition. Relying on reason alone would produce death or madness, and thus a complete failure both in practice and theory, rather than the triumph of rationality."1.5.3..

<sup>61</sup> G.Sartor, *op. cit.*, § 2.2.11..

Anche quest'ancoraggio alla mente è un avanzamento delle ricerche di intelligenza artificiale rispetto al passato, Sartor chiama questo avanzamento *'the mentalistic turn'* rispetto alle posizioni precedenti basate su linguaggio, che a loro volta rappresentarono *'the linguistic turn'*<sup>62</sup>. È così possibile prendere in considerazione un più ampio spettro di facoltà cognitive umane. Resta però il problema che nella mente vengono ricompresi aspetti che riguardano l'individuo, la sua struttura genetica e biologica, come anche aspetti della realtà - l'ambiente e la cultura - che, interagendo con la struttura biologica la vanno a formare. In questo modo non si comprende bene se la razionalità riguardi la natura delle cose o la natura dell'uomo, cioè se quegli schemi di ragionamento e/o i loro contenuti siano propri dell'individuo o se, invece, riguardino la realtà esterna all'individuo, che, lo 'costringerebbe' all'agire secondo schemi razionali non appartenenti all'individuo. In particolar modo ciò riguarda l'inferenza *defeasible*, che Sartor ammette nella sua teorizzazione. Qui notiamo che, è ben vero che anche le scienze naturali ammettono l'uso di questo tipo d'inferenza, lo fanno però sempre in congiunzione con la prova empirica di quel che si afferma vero o esistente. Le scienze naturali analizzano dati appartenenti alla realtà esterna all'individuo, non alla sua mente, la prova della verità, la truth-preservingness, è lasciata alla realtà esterna. Se invece gli stessi tipi di ragionamento si utilizzano senza prova empirica allora, nelle scienze naturali, essi diventano invalidi o tuttal'più euristici. Nel caso della mente, quale prova empirica è possibile? A causa del fatto che la mente mescola dati appartenenti all'individuo e dati esterni appare difficile distinguere tra i due, oppure appare difficile provare i dati appartenenti all'agente.

Richard Dawkins, biologo inglese, da alcuni anni ha evidenziato come il possedere la capacità di imitare ed apprendere possa anche essere visto come una debolezza dell'individuo, perché ciò lo espone all'influenza delle informazioni che gli vengono fornite dall'esterno<sup>63</sup>. In

---

<sup>62</sup> G.Sartor, *op. cit.*, § 3.3.4..

<sup>63</sup> R. Dawkins, *Il gene egoista*, Mondadori, Milano 1995; R. Dawkins, *Il cappellano del Diavolo*, Raffaello Cortina Editore, Milano 2004; O.R. Goodenough - R. Dawkins, *The St. Jude Mind Virus*, in: *Nature*, 371, 1994, p.23.; ma vedi anche P.J. Richerson, R. Boyd, *Not by genes alone, how culture transformed human evolution*, University Of Chicago Press, Chicago 2005; L.L.Cavalli Sforza, *Similarities and Dissimilarities of Sociocultural and Biological Evolution*, in: F.R.Hodson, D.G.Kendall, P.Tautu, Eds., *Mathematics in the Archaeological and Historical Sciences*, Edinburgh University Press, Edimburgo 1971, p. 535 ss. Vedi: L.L. Cavalli-Sforza, M.W. Feldman, *Cultural Transmission and Evolution: a*



tal caso, se l'ultima parola spetta alla nostra natura biologica, come del resto Sartor accetta, allora la nostra decisione (*want*) non deriva dalla formazione di un piano o dall'adozione di una strategia, bensì da *come* le informazioni provenienti dall'esterno sono riuscite ad avere effetto sulla formazione del primo livello (*like-dislike*) di stati conativi, portando quindi ad esiti irrazionali qualsiasi piano razionale basato su di essi. Così, ad esempio, si sceglie di fumare, nonostante tutto, oppure si sceglie di uccidersi con un atto di terrorismo suicida ecc.. Come potentemente l'informazione possa agire su questi presupposti biologici dell'individuo ne è stato esempio il nazismo: all'epoca la scelta della persecuzione del popolo ebraico appariva razionale, con tutte le motivazioni che venivano addotte e che apparivano convincenti. Non diversamente agisce oggi l'informazione sulla formazione della volontà individuale nell'America post 9.11, nonché sulle masse di mussulmani che reclamano un'impostazione religiosa integralista. Concordo con Sartor sul fatto che la razionalità sia un potente mezzo per limitare la forza con la quale l'informazione può influire sulla dotazione biologica individuale, però dubito che, scegliendo la mente e non il cervello come sede della razionalità, si riesca a distinguere ciò che condiziona da ciò che è condizionato. Non dividere la realtà esterna dall'individuo porta a non individuare le rispettive caratteristiche: quelle dell'informazione appartenente al mondo esterno e quelle dell'informazione appartenente all'individuo. Occorre studiare quindi separatamente i due tipi d'informazione.

In particolar modo ciò è vero per il diritto in cui il fine perseguito non è quello della massimizzazione della fitness individuale bensì anche quello della fitness del gruppo. Trovare una giustificazione al diritto, considerando una soltanto delle caratteristiche biologiche dell'individuo (la razionalità), può essere pericoloso se non si separa lo studio dell'informazione nei due versanti esterno ed interno, così come fondare su una fede, qualsiasi essa sia, anche quella nell'oggettività della ragione, il diritto. Per violare i confini segnati da Kelsen è necessario, e soprattutto utile, avere degli ottimi motivi, che non possono essere altri che quelli di aver trovato una fondazione del diritto che ne chiarisca da un lato la sua funzione, guardando al futuro, e che si agganci alla sua origine empirica, guardando al passato, questa è la posta sulla quale si

---

*Quantitative Approach*, Princeton University Press, Princeton, 1981; L.L. Cavalli-Sforza, *Geni, popoli e lingue*, Adelphi, Milano 1996; L.L. Cavalli Sforza, *L'evoluzione della cultura*, Codice Edizioni, Torino 2004.

gioca la partita delle scienze cognitive applicate al diritto, lo scritto di Sartor è un primo forte passo in questa direzione, ma non è ancora decisivo.

Per concludere vorrei riformulare la frase di Sartor, posta in epigrafe a questo scritto come segue: In view of that, I would like to conjecture by endorsing my trust in the certainty of law.