

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Un approccio cognitivo alla rappresentazione dei concetti nelle ontologie formali

### **This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/141403> since 2020-04-17T13:20:07Z

*Published version:*

DOI:10.1422/75352

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Un approccio cognitivo alla rappresentazione dei concetti nelle ontologie formali

Antonio Lieto  
Università di Torino  
Dipartimento di Informatica  
Corso Svizzera 185, 10149 Torino  
[lieto@di.unito.it](mailto:lieto@di.unito.it)

PREPRINT - DRAFT VERSION In corso di stampa su *Sistemi Intelligenti*

La rappresentazione dei concetti è un settore centrale di ricerca in Intelligenza Artificiale (IA) e nelle Scienze Cognitive. In anni recenti, tale tematica ha ricevuto una grande attenzione nell'ambito dei formalismi per lo sviluppo delle ontologie computazionali e nel settore della ingegneria della conoscenza. Tuttavia, ad oggi, molti problemi classici (quali, ad esempio, quelli legati alla rappresentazione di informazioni prototipiche e al ragionamento non monotono) sono rimasti irrisolti. In questo articolo si propone una analisi della situazione e, nella parte finale, viene avanzata una proposta per il superamento di tali problemi basata su alcuni suggerimenti provenienti dalle scienze cognitive.

## 1 La rappresentazione dei concetti tra Scienze Cognitive e IA

Numerosi sono i punti di contatto tra Scienze Cognitive e Intelligenza Artificiale sul tema della rappresentazione dei concetti (si veda Cordeschi e Frixione 2011). Sin dagli anni '70 del secolo scorso, infatti, i risultati provenienti dalle scienze cognitive hanno avuto notevoli ripercussioni sia sullo studio dei sistemi concettuali naturali che artificiali. La teoria di Eleanor Rosch sui prototipi (Rosch 1975) - che ha dimostrato empiricamente la inadeguatezza della “teoria classica dei concetti”<sup>1</sup> - ha avuto, ad esempio, una

---

<sup>1</sup> Secondo la quale un concetto può essere definito semplicemente sulla base di condizioni necessarie e sufficienti.

influenza diretta sui primi ricercatori di Intelligenza Artificiale (IA) nel campo della rappresentazione della conoscenza (RC)<sup>2</sup>. In una prima fase di sviluppo di tale disciplina, infatti, molti pionieri tentarono di prendere in considerazione i suggerimenti provenienti dalla ricerca psicologica, progettando sistemi artificiali – quali ad esempio i *Frames* (Minsky 1975) e le Reti Semantiche (Quillian 1968) - in grado di rappresentare informazione “non classica” (prototipica) e di ammettere forme di ragionamento non monotono quali, ad esempio, meccanismi di eccezione al principio di ereditarietà<sup>3</sup>.

Tuttavia, tali sistemi erano di solito caratterizzati in modo impreciso e piuttosto approssimativo, in quanto mancavano di una definizione formale chiara. Di conseguenza, furono successivamente sacrificati a vantaggio di una classe di sistemi di rappresentazione della conoscenza che ha avuto (ed ha ancora) una grande diffusione in Intelligenza Artificiale. Più precisamente: la classe di formalismi derivante dalle cosiddette reti semantiche ad ereditarietà strutturata e dal sistema KL-ONE (Brachman e Schmoltze, 1985). I discendenti di tali formalismi, sono oggi conosciuti come *logiche descrittive* (*description logics*, *DL*) (Baader et al., 2002) e rappresentano la base del linguaggio OWL (Ontology Web Language) su cui si fonda la realizzazione delle ontologie formali per lo sviluppo del web semantico. In tali formalismi, una semantica tarskiana di tipo compositazionale<sup>4</sup> è direttamente associata alla sintassi del linguaggio. La scelta di privilegiare tali aspetti, se da un lato ha migliorato l'apparato formale delle rappresentazioni, ha lasciato, però, irrisolto sia il problema della rappresentazione prototipica dei concetti che quello relativo al ruolo giocato dal ragionamento non monotono.

---

<sup>2</sup> I risultati empirici ottenuti da Eleanor Rosch furono, in un certo senso, preceduti storicamente dall'analisi filosofica di Ludwig Wittgenstein che introdusse la nozione di «*somiglianze di famiglia*» per concetti come GIOCO che risultavano essere assolutamente indefinibili in base alla teoria classica dei concetti.

<sup>3</sup> Il principio di *ereditarietà* è un meccanismo inferenziale standard nei sistemi orientati ai concetti in base al quale ogni concetto, di norma, eredita gli attributi dei concetti più generali che lo sussumono. Normalmente i sistemi che derivano dalla tradizione delle reti ad ereditarietà strutturata non permettono la violazione di questo principio. I sistemi che consentono, invece, la rappresentazione di informazione prototipica devono consentire, al contrario, la possibilità di ammettere eccezioni alla ereditarietà. Deve, cioè, poter succedere che un concetto non erediti tutti gli attributi dei concetti da cui è sussunto.

<sup>4</sup> La compositionalità è un principio in base al quale il significato di un simbolo complesso *s* dipende funzionalmente dalla struttura sintattica di *s* e dal significato dei simboli primitivi che lo compongono. In altre parole, è possibile determinare in modo non ambiguo il significato di un simbolo complesso a partire dal significato dei suoi componenti. Tale concetto è stato introdotto per la prima volta da Gottlob Frege.

Proprio nel contesto dei formalismi per le ontologie formali il ruolo del ragionamento non monotono è attualmente un problema molto dibattuto. Sono in corso, infatti, numerose ricerche per l'applicazione di logiche non monotone finalizzate ad ottenere un ulteriore sviluppo delle capacità di ragionamento delle applicazioni di Semantic Web.

Nell'ambito della letteratura relativa ai sistemi di rappresentazione della conoscenza simbolici e orientati alla logica, esistono diversi approcci rigorosi che permettono di rappresentare eccezioni e che quindi, sarebbero, almeno in linea di principio, utilizzabili per rappresentare informazioni e ragionamenti non classici. Sono da leggere in questa ottica, ad esempio, le soluzioni proposte nell'ambito delle cosiddette Logiche Descrittive Approssimate (*Rough DLs*, si veda ad es. Schlobach et al 2007) - che puntano ad estendere le logiche descrittive classiche per consentire la modellazione di concetti vaghi mediante descrizioni approssimate ottenute tramite l'introduzione di operatori di tipo modale o mediante l'utilizzo di fuzzy DLs - e gli approcci orientati all'integrazione delle logiche fuzzy nelle logiche descrittive (per una rassegna si rimanda a Calegari e Ciucci 2007 e a Bobillo e Straccia 2009). Nell'ambito degli approcci ai prototipi basati su logica fuzzy, tuttavia, un noto problema è rappresentato dalle difficoltà che tali proposte incontrano con la composizionalità. Ovvero: la fuzzy set theory non è in grado di supportare una semantica composizionale il cui input è costituito di concetti prototipici (si veda Osherson e Smith 1981). Anche le diverse estensioni probabilistiche o non monotone del linguaggio OWL e dei relativi formalismi possono essere considerate come tentativi di trovare una soluzione ai problemi sopra elencati. Per quanto riguarda i cosiddetti approcci probabilistici ai linguaggi ontologici, molte delle proposte presenti in letteratura, come sottolineato da Lukasiewicz e Straccia (2008), sono focalizzate sulla combinazione del linguaggio OWL con formalismi basati su reti Bayesane. In particolare Da Costa e Laskey (2006) hanno proposto una generalizzazione probabilistica del linguaggio OWL (chiamata PR-OWL) la cui semantica è basata su reti bayesane multi entità (multi-entity Bayesian networks, MEBNs). Ding et al (2006), invece, hanno invece proposto una soluzione basata su reti Bayesane standard. Un problema generale di questi approcci consiste, a nostro avviso, nell'evitare l'arbitrarietà nell'assegnazione dei pesi nella traduzione che va da formalismi di rappresentazione tradizionali a probabilistici. Per quanto concerne, infine, le estensioni non monotone delle logiche descrittive (si veda ad es. Baader e Hollunder, 1995), il problema principale di tali proposte è rappresentato dalle difficoltà computazionali e dalla semantica estremamente complessa che e-

merge da tale integrazione (per una rassegna più esaustiva degli approcci citati si veda Frixione e Lieto 2010).

In buona sostanza, dunque, le logiche descrittive, mentre sono in grado di catturare l'aspetto compositivo della conoscenza concettuale, risultano essere ancora inadeguate a rappresentare la conoscenza prototipica. E una alternativa non classica – una logica descrittiva generale in grado di rappresentare informazioni in termini prototipici – non sembra ancora emergere. Come possibile via di uscita da questo *trade-off* che vede, da sempre, contrapposte le esigenze di composizionalità, da un lato, e di prototipicità dall'altro (su questo aspetto si rimanda a Frixione e Lieto, 2011) si propone, nelle pagine a seguire, una possibile soluzione basata su diversi suggerimenti provenienti dalla scienza cognitiva.

## **2 Alcuni suggerimenti dalle Scienze Cognitive**

Alcune tendenze recenti nell'ambito della ricerca psicologica sembrano favorire l'ipotesi che i processi di ragionamento negli esseri umani non costituiscono un fenomeno unitario. Allo stesso modo, i vari aspetti legati alle abilità concettuali potrebbero derivare da meccanismi differenti di rappresentazione. Dato che, a nostro avviso, una metodologia matura per affrontare il problema della rappresentazione concettuale in sistemi artificiali dovrebbe tenere conto dei risultati provenienti dalla psicologia cognitiva, noi proponiamo di seguire alcuni di questi suggerimenti nello sviluppo di sistemi orientati alla rappresentazione della conoscenza. Nello specifico si propone: (i, par. 2.1) di tenere in considerazione la distinzione tra due differenti tipi di sistemi di ragionamento sviluppata nell'ambito della cosiddetta *teoria dei processi duali di ragionamento e razionalità*, (ii, par. 2.2) di mantenere gli aspetti prototipici separati dalla rappresentazione compositiva dei concetti, (iii, par. 2.3) di sviluppare modelli ibridi dei concetti, basati su formati diversi di rappresentazione.

### **2.1 Un approccio duale ai processi di ragionamento**

La ricerca cognitiva sui concetti sembra suggerire che la rappresentazione concettuale non costituisce un fenomeno unitario da un punto di vista cognitivo. In questa ottica una possibile soluzione potrebbe essere ispirata ai risultati sperimentali provenienti dalla cosiddetta *teoria dei processi duali*

(Dual Process Theory) *del ragionamento e della razionalità*<sup>5</sup> (Evan e Frankish 2008). In tali teorie viene ipotizzata l'esistenza di due diversi tipi di sistemi cognitivi. Un sistema di primo tipo (tipo 1) filogeneticamente più vecchio, automatico, parallelo e veloce, e un sistema di tipo 2 più recente, sequenziale, lento e basato su regole esplicite. A nostro avviso ci sono, *prima facie*, buone ragioni per ritenere che, nei soggetti umani, il processo di classificazione<sup>6</sup> sia una forma di ragionamento di tipo 2. Al contrario, le eccezioni giocano un ruolo importante nei processi come la categorizzazione non monotona<sup>7</sup> e nei meccanismi di ereditarietà, che sembrano essere processi associabili ai sistemi di tipo 1: sono veloci, automatici e così via.

Dunque, una ipotesi ragionevole in grado di assecondare tale prospettiva, è quella in base alla quale un sistema di rappresentazione concettuale dovrebbe includere moduli diversi di ragionamento: uno monotono di tipo 2, coinvolto in compiti "difficili" come la classificazione, e uno non monotono, di tipo 1, coinvolto in forme di ragionamento - come la categorizzazione - che ha a che fare con la gestione delle eccezioni. Tale soluzione va nella direzione di una rappresentazione "duale" dei concetti nelle ontologie formali, con la realizzazione di sistemi di ragionamento ibrido (monotono e non monotono).

## 2.2 Una proposta pseudo-fodoriana

Secondo Jerry Fodor i concetti non possono essere rappresentazioni prototipiche, in quanto i concetti devono essere composizionali e i prototipi, invece, non lo sono. D'altra parte, in virtù delle critiche alla "teoria classica", i concetti non possono neppure essere definizioni. Pertanto Fodor ritiene che la maggior parte dei concetti siano rappresentazioni atomiche (ossia, simboli senza struttura interna). Il loro contenuto è determinato dalla loro relazione

---

<sup>5</sup> Tale teoria pone l'accento sui diversi tipi di processi cognitivi coinvolti nei processi di ragionamento e non è da confondersi con la "dual code theory" secondo la quale gli esseri umani usano contemporaneamente sia rappresentazioni verbali che non-verbali (immagini) per rappresentare le informazioni (si veda. Sternberg, 2003; pp. 216-9 & 234-36).

<sup>6</sup> La classificazione è una forma di ragionamento monotono e deduttivamente corretto che consiste nell'inferire relazioni di tipo sottoclasse/superclasse (relazioni ISA) dalle informazioni implicitamente codificate all'interno di una base di conoscenza. Nelle logiche descrittive tali relazioni vengono indicate utilizzando, di solito, il termine "sussunzione".

<sup>7</sup> La categorizzazione è un processo inferenziale che consiste nella assegnazione di una specifica istanza ad una certa classe. Nei processi di categorizzazione non monotona il processo di assegnazione ad una classe è un processo inferenziale non deduttivo basato sulla tipicità. Nell'ambito del machine learning ci si riferisce, di solito, al processo di categorizzazione utilizzando il termine "classificazione".

con il mondo e non dalla loro struttura interna e/o dalle loro relazioni con altri concetti (Fodor, 1987). Naturalmente, Fodor riconosce l'esistenza degli effetti prototipici. Tuttavia ritiene che le rappresentazioni prototipiche non siano parte dei concetti. Le rappresentazioni prototipiche permettono di individuare il riferimento dei concetti, ma non possono essere identificate con essi. Si consideri, ad esempio, il concetto CANE. Ovviamente, nella nostra mente, abbiamo qualche rappresentazione prototipica associata a questo concetto (in base alla quale, ad esempio, di solito i cani scodinzolando, abbaiano, eccetera). Tuttavia questa rappresentazione non coincide con il concetto di CANE: per Fodor il concetto CANE è un simbolo atomico non strutturato.

In questo lavoro prendiamo a prestito da Fodor l'ipotesi che le rappresentazioni composizionali e gli effetti prototipici siano demandati a componenti differenti della architettura cognitiva. Riteniamo, infatti, che sia legittimo ipotizzare una componente composizionale delle rappresentazioni che non ammette eccezioni e non mostra effetti prototipici. In un sistema artificiale, tale componente potrebbe essere rappresentata, ad esempio, come una base di conoscenza che utilizza una logica descrittiva (DL). Riteniamo però plausibile che una rappresentazione prototipica delle categorie possa essere responsabile di altri processi cognitivi, come ad esempio la categorizzazione non monotona, senza che ciò interferisca con il comportamento inferenziale della componente composizionale. Si noti tuttavia che la nostra proposta non è interamente "fodoriana", almeno rispetto ai tre punti seguenti:

i. Non si prende in considerazione il problema della natura del contenuto semantico delle rappresentazioni concettuali. Fodor adotta una teoria causal-informazionale del significato, secondo cui il contenuto dei concetti sarebbe costituito da una relazione nomica mente-mondo. Ma, in generale, il problema filosofico della natura del contenuto intenzionale delle rappresentazioni può essere trascurato come irrilevante per i nostri fini attuali.

ii. Fodor sostiene che i concetti sono composizionali, e che le rappresentazioni prototipiche, non essendo composizionali, non possono essere concetti. In questa sede, invece, non si prende alcuna posizione su quale parte del sistema che proponiamo sia autenticamente "concettuale". Sospettiamo infatti che la nozione intuitiva di concetto non sia unitaria dal punto di vista cognitivo. Sia la componente composizionale, sia quella prototipica contribuiscono a determinare il "comportamento concettuale" del sistema, vale a dire, hanno un ruolo nel determinare quelle abilità che noi di solito descriviamo nel termine di possesso di concetti.

iii. Secondo Fodor, la maggior parte dei concetti è atomica. In particolare, egli ritiene che quasi tutti i concetti che corrispondono ad elementi del lessico

non abbiano struttura. A nostro avviso, invece, molti concetti lessicali, sebbene non possano essere definiti nei termini della teoria classica, possono esibire qualche forma di struttura, e tale struttura potrebbe essere rappresentata, ad esempio, per mezzo di una logica descrittiva.

### **2.3. Teorie dei concetti nelle Scienze Cognitive**

In campo psicologico esistono diverse teorie sulla natura dei concetti. Di solito esse vengono raggruppate in tre grandi classi: teorie dei prototipi, teorie degli esemplari e teorie della teoria (Murphy, 2002). Ciascuna di esse aspira a rendere conto di alcuni aspetti degli effetti di tipicità che emergono nella concettualizzazione.

Secondo le teorie dei prototipi, la conoscenza riguardo a una data categoria viene rappresentata mentalmente nei termini di una rappresentazione della istanza “migliore” di quella categoria. Ad esempio: il concetto GATTO coinciderebbe con la rappresentazione di un gatto prototipico (un gatto con il pelo, la coda e quattro zampe, che miagola e fa le fusa, eccetera). Nella versione più semplice di questo approccio, i prototipi vengono rappresentati come liste (eventualmente pesate) di attributi caratteristici.

Secondo le teorie degli esemplari, una data categoria è codificata mentalmente tramite la memorizzazione di un certo insieme di rappresentazioni di esemplari specifici. In questo caso, la nostra rappresentazione mentale del concetto GATTO consisterebbe nella rappresentazione di un certo numero di esemplari di gatto che abbiamo incontrato nel corso della nostra vita.

Le teorie della teoria adottano un atteggiamento di tipo olistico nei confronti dei concetti. Secondo alcune versioni di questa impostazione, i concetti sono analoghi ai termini teorici in una teoria scientifica. Ad esempio, il concetto GATTO si identifica in base al ruolo che esso gioca nella nostra teoria mentale della zoologia. Secondo altre versioni, i concetti stessi vanno identificati con micro-teorie. Ad esempio, il concetto GATTO corrisponderebbe a una micro-teoria mentale sui gatti.

Queste impostazioni non risultano essere mutuamente esclusive. Piuttosto, esse sembrano avere successo nello spiegare classi diverse di fenomeni cognitivi, e molti ricercatori ritengono che esse siano tutte necessarie per spiegare i dati che provengono dalla ricerca empirica (si veda ad es. Murphy 2002). In questa prospettiva, noi proponiamo di integrare alcune di queste teorie nell'ambito della rappresentazione computazionale dei concetti. Più precisamente, si sostiene l'opportunità di combinare, nell'ambito delle rappresentazioni categoriali da implementare in un sistema artificiale, tanto rap-



presentazioni basate su prototipi che rappresentazioni basate su esemplari (Frixione e Lieto, 2012a e 2012b), in modo da poter rendere in conto, in una prospettiva integrata, degli effetti di tipicità (per una proposta simile sviluppata nell'ambito dell'apprendimento automatico, si veda Gagliardi, 2008). Non prenderemo in considerazione invece le teorie della teoria, poiché, rispetto a prototipi ed esemplari, esse sono definite in maniera più vaga, per cui un loro trattamento computazionale risulterebbe, almeno per il momento, più problematico.

### **2.3.1 Prototipi ed Esemplari**

I modelli basati su prototipi ed esemplari presentano, tra loro, differenze significative. Nei modelli basati su esemplari, ad esempio, si assume che le stesse rappresentazioni siano impiegate sia nei compiti di categorizzazione, sia nei compiti di identificazione (ossia nel riconoscimento di istanze specifiche, del tipo “questo è il Colosseo”). Ciò contrasta con i modelli basati sui prototipi, i quali assumono che questi processi cognitivi coinvolgono tipi differenti di rappresentazioni. Inoltre, un prototipo cattura solo aspetti cognitivamente salienti di un concetto, mentre la rappresentazione di un esemplare include *in toto* la conoscenza di una specifica entità. Ad esempio: la rappresentazione prototipica del concetto CANE è basata su un sottoinsieme di tratti cognitivamente rilevanti associati ai cani (come il fatto che di solito i cani scodinzolano, abbaiano ecc.). Viceversa, la rappresentazione dell'esemplare Fido potrebbe contenere informazioni periferiche specifiche come, ad esempio, l'informazione che Fido ha contratto il cimurro.

Tuttavia, le rappresentazioni basate sugli esemplari possono essere utili in molte situazioni. Vari esperimenti mostrano come istanze di un concetto che sono abbastanza dissimili dal prototipo, ma che assomigliano a qualche esemplare noto, possano essere caratterizzate dai soggetti velocemente e con sicurezza. Ad esempio, un pinguino è abbastanza dissimile dal prototipo di UCCELLO. Quindi, se la categorizzazione si basasse sul confronto con un prototipo, gli esemplari di pinguino dovrebbero essere categorizzati soltanto con fatica e con un basso grado di confidenza. Tuttavia, se io già conosco un certo esemplare di pinguino, e se so che si tratta di un uccello, non avrò problemi a categorizzare come UCCELLO un pinguino sconosciuto. Questo aspetto diventa particolarmente rilevante per quei concetti i cui membri differiscono in modo significativo l'uno dall'altro (come accade ad esempio con UTENSILE, o con VEICOLO).

Rispetto ai prototipi, le rappresentazioni basate su esemplari sono più facili e più veloci da acquisire. In alcune situazioni, può accadere che non ci sia abbastanza tempo per estrarre un prototipo dalle informazioni disponibili. Inoltre, le rappresentazioni basate su esemplari semplificano la rappresentazione di concetti che non sono linearmente separabili (Medin e Schwanenflugel 1981).

Un altro aspetto di divergenza è rappresentato dal trattamento del processo di categorizzazione. Per risolvere questo compito, sia i modelli basati su prototipi che quelli basati su esemplari hanno la necessità di calcolare il grado di similarità tra una rappresentazione *target*, da un lato, e una rappresentazione di un prototipo o di un insieme di esemplari dall'altro. Il processo avviene nel modo seguente: una rappresentazione prototipica/basata su esemplari di una data categoria viene recuperata dalla memoria e comparata con la rappresentazione *target*. Sulla base di tale comparazione il processo decide se la rappresentazione *target* appartiene o meno alla categoria. Tuttavia, nonostante questa prima fase comune ad entrambi gli approcci, ci sono in fasi successive, alcune differenze sostanziali. In base al modello basato su prototipi, infatti, il calcolo di similarità è di solito lineare. Ossia, se una proprietà è condivisa dal *target* e dal prototipo, allora la similarità tra queste due rappresentazioni aumenta indipendentemente dal fatto che ci siano o meno altre proprietà condivise. Ad esempio: supponiamo che il prototipo di SUPEREROE sia caratterizzato dalla seguenti proprietà: essere una brava persona, lottare per la giustizia, avere super poteri. Supponiamo inoltre che il concetto *target* condivida anche solo una delle tre caratteristiche elencate (poniamo che si tratti della proprietà "essere una brava persona"). Ciò sarebbe comunque sufficiente a far aumentare la similarità tra i due concetti (anche se, nel caso specifico, la singola proprietà condivisa non è assolutamente discriminante rispetto alla effettiva appartenenza alla classe dei supereroi). In base al modello basato su esemplari, invece, il calcolo della similarità è non lineare. Quindi: una proprietà condivisa dal concetto *target* e dall'esemplare viene considerata rilevante e fornisce un contributo al calcolo di somiglianza tra concetti solo se ci sono altre proprietà condivise tra le due rappresentazioni. Ad esempio: se tra il concetto *target* e l'esemplare Superman appartenente alla classe SUPEREROE ci fosse, come unica proprietà condivisa, essere giornalista, ciò non sarebbe ancora sufficiente a far "avanzare", nella scala di similarità, il concetto *target* verso l'esemplare Superman (e quindi verso la classe SUPEREROE). Tale proprietà condivisa assume una rilevanza nell'ambito del calcolo di similarità solo se tra *target* ed esemplare ci sono altre proprietà condivise (supponiamo, ad esempio, le caratteristiche di esse-

re un giornalista e quella di avere super poteri). Questa differenza nelle procedure di calcolo della similarità concettuale dipende dalla natura stessa delle rappresentazioni: mentre, infatti, i prototipi possono essere considerati come concetti delle rappresentazioni le cui caratteristiche sono tutte “salienti” da un punto di vista cognitivo (e quindi più centrali e più discriminative), gli esemplari, come detto in precedenza, possono esprimere anche condizioni particolarissime che necessitano di una fase di controllo più accentuata. Su questi aspetti si veda ad esempio il par. 3.3 del cap. 4 di (Machery 2009).

La dicotomia tra teorie della categorizzazione basate su prototipi e su esemplari non emerge solo in campo psicologico, ma è riscontrabile anche in discipline come l’apprendimento automatico e la classificazione automatica (Witten e Frank, 2005). In particolare, nell’apprendimento automatico questi due approcci sono stati adottati per la realizzazione di tipi diversi di classificatori. Ad esempio, il *Nearest Prototype Classifier* (NPC) è basato sui prototipi, mentre il *Nearest Neighbour Classifier* (NNC) è basato sulla teoria degli esemplari<sup>8</sup> (si veda Gagliardi, 2011). Recentemente, in questo ambito di ricerca, sono state presentate diverse proposte di classificatori ibridi finalizzati al superamento della dicotomia, e volti ad integrare i vantaggi provenienti da entrambi gli approcci. In questa ottica, prototipi ed esemplari non sono più considerati come impostazioni contrastanti, bensì come due casi limite delle tecniche di *Instance Based Learning*, utilizzate per la categorizzazione da parte di sistemi sia naturali, sia artificiali. Infatti, nonostante le differenze poste in evidenza dalla letteratura psicologica, i prototipi e gli esemplari, presi insieme, possono essere considerati come aspetti diversi di un unico meta fenomeno: il fenomeno di *tipicità* che prende forma nei processi di categorizzazione (Murphy, 2002). A nostro avviso, tale approccio ibrido può essere di grande aiuto per affrontare il problema della rappresentazione della tipicità nel campo delle ontologie formali.

### **3. Una architettura cognitiva per le ontologie formali**

In questo paragrafo presentiamo la proposta di una possibile architettura cognitiva per la rappresentazione di concetti non classici nell’ambito delle ontologie formali o computazionali. Tale architettura ha come scopo quello di tenere in considerazione gli aspetti cognitivi precedentemente individuati. Essa, pertanto, è basata su: modularità delle rappresentazioni concet-

---

<sup>8</sup> Questa distinzione, in realtà, non è così netta in quanto anche i sistemi con più prototipi per classe possono rappresentare correttamente classi non linearmente separabili come ad es. il classificatore NMPC (Nearest Multiple Prototype Classifier).

tuali (i), dualità dei processi di ragionamento (ii), possibilità di esprimere rappresentazioni “tipiche” basate su prototipi e su esemplari (iii).

Per quanto concerne la modularizzazione delle rappresentazioni e la dualità dei processi di ragionamento, l'architettura prevede, per uno stesso concetto, una componente composizionale<sup>9</sup> - basata su logiche descrittive classiche - e una componente tipica che rappresenta, per gli stessi concetti della base di conoscenza DL, informazione tipica di senso comune e che permette, almeno in linea di principio, modalità di ragionamento approssimato o, comunque, non classico.

I concetti nella componente composizionale sono rappresentati come in figura. 1.

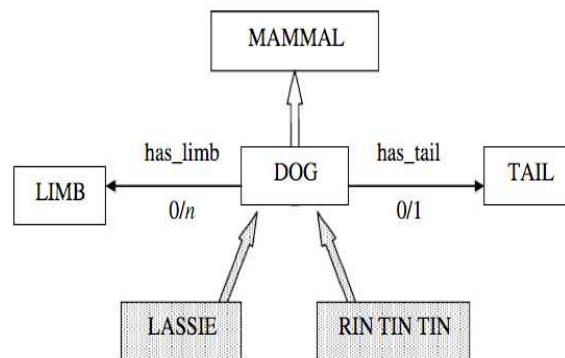


Fig. 1

Ogni concetto può essere sussunto da un certo numero di superconcetti e può essere rappresentato mediante attributi che lo collegano ad altri concetti all'interno della base di conoscenza. In termini di logica dei predicati, i concetti corrispondono a predicati ad un argomento e gli attributi a relazioni a due argomenti. A ciascun attributo possono essere associate delle restrizioni

---

<sup>9</sup> E' bene ricordare che “composizionale” non è sinonimo di “definitorio”. Non è detto che un sistema composizionale debba esprimere delle definizioni; ad esempio in una logica descrittiva si può descrivere un concetto senza definirlo (nel senso tecnico di esprimere un insieme di condizioni individualmente necessarie e congiuntamente sufficienti).

sui valori possibili che può assumere. Dato un concetto, gli attributi e le relazioni espresse rappresentano, per esso, condizioni necessarie e/o sufficienti. I concetti possono avere, inoltre, anche un numero imprecisato di istanze rappresentate come concetti individuali all'interno della tassonomia. Ad esempio: in una base di conoscenza concernente il concetto CANE, istanze di questo concetto possono essere gli individui {Fido, Lassie, Rin Tin Tin ecc...}.

Per quanto concerne la componente tipica, invece, i prototipi sono rappresentati come strutture dati esterne alla base di conoscenza DL. Tali strutture possono essere, ad esempio, liste di coppie - eventualmente pesate - di tipo attributo/valore. Alcuni attributi prototipici possono corrispondere a proprietà già specificate, per lo stesso concetto, nella base di conoscenza DL ma che vengono ulteriormente ristrette. Ad esempio: se nella componente DL l'attributo avere *AvereCoda*, per il concetto CANE, è espresso con una restrizione di tipo 0/1 (che indica che il valore che può assumere tale proprietà è 0 oppure 1), la componente prototipica del concetto CANE può contenere l'informazione che i cani - tipicamente - hanno esattamente 1 coda. Altri attributi prototipici, invece, possono non avere alcuna controparte nella base DL e, quindi, aggiungere informazione completamente nuova per quel concetto.

Le due componenti dell'architettura, dunque, esprimono sostanzialmente livelli diversi di informazione per uno stesso concetto. Consideriamo un altro esempio: il concetto di ACQUA. In questo caso, nella componente DL ci sarà una rappresentazione della sostanza che corrisponde anche alla sua definizione in termini di condizioni necessarie e sufficienti basata sulle conoscenze della chimica (cioè, che l'acqua è H<sub>2</sub>O<sup>10</sup>). Nella parte prototipica, invece, sarà rappresentata l'informazione che descrive l'ACQUA come un liquido tipicamente insapore, inodore e incolore.

A nostro avviso lo schema di rappresentazione della componente prototipica, e il suo collegamento alla rappresentazione DL, trova una sistematizzazione possibile nel contesto dei linguaggi ontologici per il web, e più nello specifico nella prospettiva dei Linked Data (Bizer, Heath and Berners-Lee 2009). Tale approccio, nell'ambito della comunità scientifica del web semantico, sta assumendo un ruolo prominente. Uno dei principali obiettivi dei Linked Data riguarda l'integrazione di differenti rappresentazioni di dati (localizzati, di solito, in differenti sorgenti<sup>11</sup>) all'interno di un unico framework

---

<sup>10</sup> Per essere precisi, l'acqua (intesa come sostanza del senso comune, ossia, per intenderci, quella che tipicamente "esce da un rubinetto") è una soluzione, di cui H<sub>2</sub>O costituisce il solvente.

<sup>11</sup> Tale situazione è tipica dei sistemi informativi distribuiti e/o federati e rappresenta, ad oggi, la situazione più comune in cui ci si imbatte quando ci si trova a dover svolgere compiti di *data integration*.

rappresentazionale. Tale integrazione rende possibile l'allargamento della spazio di risposte possibili fornite ad una query mediante la realizzazione di "ponti semantici" tra pezzi differenti di informazioni e sorgenti di dati. Tale integrazione è possibile attraverso specifici costrutti forniti dai linguaggi della comunità del Semantic Web quali, ad esempio, OWL.

Riteniamo che l'implementazione di rappresentazioni basate su esemplari e su prototipi possa trarre vantaggio dall'approccio dei Linked Data. Per quanto concerne i prototipi, infatti, essi sono rappresentabili mediante il cosiddetto Open Knowledge-Base Connectivity (OKBC) protocol<sup>12</sup> il cui *knowledge model* è supportato dalle cosiddette Frame Ontologies. Dato che è possibile esportare, senza perdere l'informazione prototipica rappresentata, tali ontologie in linguaggi basati su RDF, la connessione tra le due rappresentazioni (quella composizionale e quella prototipica) può essere realizzata utilizzando i formalismi standard forniti dalla comunità del semantic web nell'ambito della prospettiva dei Linked Data (utilizzando, ad esempio, i costrutti sintattici *owl:EquivalentClass*, *owl:SameAs* o costrutti similari di linking). Nel caso, ad esempio, del costrutto *owl:EquivalentClass*, il modello di collegamento che può essere fornito tra la rappresentazione DL e quella prototipica ha, in pseudo-codice OWL, la forma seguente:

```
<owl:Class rdf:ID="DLClassNameofConceptX">  
  
<owl:EquivalentClass  
rdf:resource="URI_FrameOntology/PrototypeClassNameofConceptX"/>  
  
</owl:Class>13
```

<sup>12</sup> OKBC (<http://www.ai.sri.com/~okbc/>) è uno degli elementi su cui si è basato il linguaggio OIL (su cui, a sua volta, si basa il linguaggio OWL che è dato dalla integrazione di DAML + OIL). OIL (che in buona sostanza è una estensione dell'OKBC knowledge model) ha rappresentato storicamente uno dei primi tentativi di integrare, in linguaggi per il semantic web, le caratteristiche di modellazione dei sistemi frame-based con una semantica formale e un ragionamento efficiente proveniente dall'ambito delle Logiche Descrittive (per maggiori dettagli si rinvia a Corcho O. e Gómez-Pérez, 2000 e Horrocks et. al. 2003). Nel nostro caso la scelta di utilizzare il protocollo OKBC per le rappresentazioni della componente prototipica è stata dettata dalla possibilità di esportare tali informazioni in linguaggi RDF based. Ciò permette di avere compatibilità con i linguaggi ontologici standard utilizzati per descrivere le rappresentazioni nella componente composizionale.

<sup>13</sup> Si noti che, assumendo l'utilizzo del linguaggio OWL Full (tra le cui proprietà espressive rientra quella di poter utilizzare le classi come istanze) sarebbe corretto utilizzare il costrutto *owl:sameAs* (di norma utilizzabile, ad es. in OWL DL, per esprimere identità tra istanze) anche per indicare le identità tra classi. Ad ogni modo, a scanso di equivoci, è necessario sottolineare che, nella proposta avanzata, le connessioni tra le due rappresentazioni sono da intendersi solo tra classi di concetti classici con classi di concetti prototipici e tra istanze di concetti classici con istanze di concetti prototipici.

Per quanto concerne gli esemplari, invece, essi possono essere direttamente rappresentati nella base di conoscenza DL come specifiche istanze di concetti. Tuttavia può anche darsi che alcune informazioni particolari legate ad esemplari di un concetto possano essere rappresentate all'esterno della componente DL, sempre mediante Linked Data<sup>14</sup>.

Con questo tipo di architettura ibrida, processi diversi di ragionamento possono seguire percorsi differenti. Ad esempio, la categorizzazione e la classificazione monotona possono coinvolgere solo la base di conoscenza ontologica in DL, mentre la categorizzazione non monotona, basata sugli elementi di typicalità, potrebbe avvantaggiarsi di strutture esterne alla componente composizionale dell'ontologia connesse mediante il framework dei Linked Data.

La possibilità di effettuare, almeno in linea di principio, ragionamento non monotono solo all'esterno della componente composizionale è, a nostro avviso, uno degli aspetti principali di questa proposta. Tra le altre cose, infatti, questa soluzione rende possibile evitare problemi di inconsistenza logica all'interno della componente composizionale<sup>15</sup>, introducendo, al tempo stesso, nella base di conoscenza ontologica "duale", la possibilità di espandere le tipologie di ragionamento permesse.

Il ragionamento distribuito, infatti, è una caratteristica comune a molti sistemi di rappresentazione della conoscenza. Tuttavia, nel campo delle ontologie formali, esso è stato spesso limitato a casi in cui è semplicemente permesso effettuare una stessa tipologia di ragionamento (ad es. la classificazione basata su logica deduttiva classica) su basi di conoscenza multiple e modulari. A nostro avviso l'implementazione di un processo di categorizzazione ibrido, non monotono, basata su prototipi ed esemplari dovrebbe, invece, tenere in considerazione i suggerimenti provenienti dall'ambito del *machine learning* che, come evidenziato in precedenza, spingono in una direzione volta a superare la dicotomia prototipi/esemplari al fine di ottenere un approccio ibrido alla categorizzazione basata su typicalità. In questa ottica una possibilità consiste ad esempio nell'impiego dell'algoritmo ibrido *PEL-C* (*Prototype Exemplar Learning Classifier*) sviluppato da Gagliardi (2008), che va in questa direzione<sup>16</sup>.

---

<sup>14</sup> Questo potrebbe essere il caso in cui viene collegata informazione "non simbolica" come, ad esempio, immagini.

<sup>15</sup> Questo è stato uno dei problemi principali riscontrati sia nei sistemi basati su frame che in altri approcci ibridi di rappresentazione della conoscenza.

<sup>16</sup> Nel Semantic Web esiste una vasta letteratura sulla tema della Concept Induction (si veda ad es. D'Amato et al. 2008, Fanizzi et al. 2011). Tali approcci si ripropongono di ottenere forme di ragionamento non monotono su basi di conoscenza DL. Sebbene i lavori presenti in questo ambito siano di fondamentale importanza per lo sviluppo di reasoning non classico su

Il vantaggio di associare conoscenza prototipica ai concetti non è solo limitata al compito della categorizzazione. Consideriamo, ad esempio, il compito del *property checking*. Il property checking è un task che consiste nel rispondere a domande del tipo “La classe A possiede la proprietà b?”. Supponiamo, ad esempio, che un utente esegua una query informazionale<sup>17</sup> su una base di conoscenza duale (realizzata così come proposto) relativa al dominio della Frutta allo scopo di sapere quale tipo di AGRUME è giallo (che corrisponde alla domanda “esiste qualche agrume che la proprietà di essere giallo?”). Intuitivamente, la risposta attesa dal sistema, da parte dell’utente, è LIMONE. Tuttavia, nella base di conoscenza DL dell’architettura duale, non esiste alcun agrume che possiede la proprietà “essere giallo” come condizione definitoria. “Essere giallo”, infatti, non è una condizione necessaria per essere un limone e, quindi, tale proprietà non è associata al concetto LIMONE nella ontologia DL. Tuttavia, da un punto di vista cognitivo, la proprietà essere giallo è rilevante ai fini della caratterizzazione di un LIMONE. In base al nostro approccio ibrido, tale proprietà può essere rappresentata tra le informazioni “tipiche” associate al concetto LIMONE<sup>18</sup> (o in

---

ontologie, si è preferito, in questa fase, porre l’enfasi principalmente su prospettive - come quella indicata - che tengono già conto, in nuce, della differenza riscontrata in psicologia tra prototipi ed esemplari

<sup>17</sup> In base alla letteratura dell’Information Retrieval, le *query informazionali* sono differenti da quelle *transazionali* e *navigazionali*. Nelle *query informazionali* l’intenzione dell’utente è di ottenere una specifica informazione riguardante un dato oggetto.

<sup>18</sup> Si noti che, in base all’approccio proposto, questa distinzione tra le proprietà rappresentate nelle due diverse componenti costituisce il solo modo corretto di rappresentare diverse tipologie di informazioni per uno stesso concetto. In altri termini: non sarebbe corretto inserire la proprietà essere giallo in una ipotetica base di conoscenza DL relativa agli agrumi dato che essa non rappresenta una condizione necessaria/sufficiente per definire questo concetto né i rispettivi sottoconcetti (es. Limone). Il fatto che, attualmente, proprietà prototipiche vengano inserite in basi di conoscenza DL classiche rappresenta, da nostro punto di vista, un errore di modellazione. Consideriamo un esempio (preso da Motik et al. 2006) proveniente dal mondo della medicina. Supponiamo che la conoscenza da modellare sia a seguente: “le persone hanno, di solito, il cuore a sinistra, ma alcune persone (i destrocordiaci) hanno il cuore a destra”. Motik et.al affermano che tale conoscenza non può essere modellata in termini classici da una base di conoscenza OWL in quanto gli assiomi  $UOMO \sqsubseteq CuoreASinistra$ ,  $DESTROCARDIACO \sqsubseteq UOMO$  e  $DESTROCARDIACO \sqsubseteq \neg CuoreASinistra$  renderebbero il concetto DESTROCARDIACO insoddisfacibile pena l’insorgere di una contraddizione. Dal nostro punto di vista sarebbe un errore modellare il dominio in questi termini. La procedura corretta per tenere in considerazione l’“eccezione” rappresentata dal concetto DESTROCARDIACO sarebbe, in base al nostro approccio, la seguente: la classe UOMO andrebbe rappresentata in termini di condizioni necessarie e sufficienti nel modulo DL, la proprietà CuoreASinistra andrebbe rappresentata come proprietà prototipica degli esseri umani (e quindi andrebbe inserita nella componente rappresentazionale tipica riferita al concetto UOMO) mentre la condizione “atipica” di essere un destrocordiaci può essere modellata definendo la classe DESTROCARDIACO come:  $DESTROCARDIACO \sqsubseteq$



termini di valore dell'attributo *colore* del corrispondente prototipo o sottoforma di conoscenza legata a specifici esemplari). In tal modo diventerebbe possibile recuperare l'informazione desiderata dalla componente "tipica" della rappresentazione (formata da prototipi e/o esemplari). Data una query in linguaggio SPARQL<sup>19</sup> del tipo:

```
SELECT ? Agrume
      WHERE {? Agrume : HaColore : GIALLO
            }
```

il risultato fornito dalla rappresentazione DL dovrebbe essere nullo, mentre la risposta "corretta" (corretta dal punto di vista di quelle che sono le intenzioni cognitive e informative dell'utente) sarebbe generata dalla componente prototipica della rappresentazione ibrida.

#### 4. Risultati iniziali e lavoro futuro

La proposta qui descritta è stata sperimentata su basi di conoscenza costruite ad hoc. I primi risultati ottenuti da una serie di test pilota hanno mostrato un miglioramento di prestazioni - rispetto alle strutture ontologiche classiche a livello unico - in compiti di information retrieval (nello specifico di concept retrieval) per query di tipo informazionale basate su proprietà tipiche (del tipo di quella illustrata nell'esempio del Limone). Tali risultati iniziali sono il frutto della ratio con cui è stata costruita la componente rappresentazionale. Pare, cioè, che suddividere, già in fase di modellazione, le componenti informazionali prototipiche da quelle classiche permetta - rispetto ai modelli tradizionali in cui tali informazioni sono inserite nella stessa base di conoscenza - di recuperare sempre l'informazione prototipica associata ad un determinato concetto. Il motivo di tale risultato è da ricercare nel fatto che, mentre nelle rappresentazioni classiche l'inserimento dell'informazione prototipica era demandato alla sensibilità dell'ingegnere della conoscenza (rappresentando, comunque, come illustrato nell'esempio del concetto di destrocardiaco, un errore di modellazione), in questo caso è esplicitamente previsto un doppio livello di informazioni disponibili per uno stesso concetto e, di conseguenza, la componente tipica viene sempre considerata (e pertanto

---

UOMO (in DL) e con la proprietà CuoreADestra come condizione necessaria e sufficiente per essere un destro cardiaco.

<sup>19</sup> Il linguaggio SPARQL (l'acronimo sta per **S**tandard **P**rotocol and **R**DF **Q**uery **L**anguage) è, dal 2008, uno standard del W3C raccomandato per l'interrogazione di basi di conoscenza ontologiche.

questa tipologia di informazioni è sempre disponibile/recuperabile) . Tali risultati iniziali, ad ogni modo, ne-cessitano di indagini ulteriori (ad es. è necessario considerare basi di conoscenza di maggiore dimensioni) e più approfondite.

Una fase di sviluppo futuro consiste nella realizzazione della componente relativa ai processi di ragionamento non monotòno (e alla relativa integrazione con la componente di reasoning classico) di cui si vuole dotare la componente tipica. In questo caso, come è noto, in fase information retrieval, il vero punto di collegamento tra i livelli informativi previsti dalle due componenti e le relative conoscenze inferite sulla base delle diverse forme di ragionamento considerato, saranno i cosiddetti entailment regimes di SPARQL. Gli entailment regimes sono in sostanza estensioni della semantica di SPARQL che permettono, rispetto alla computazione standard del calcolo dei risultati (basato sul matching tra la tripla richiesta in input e il relativo sottografo RDF della base di conoscenza), di materializzare la conoscenza RDF implicita prima di una query permettendo, quindi, di recuperare informazioni aggiuntive. Data l'incompleta realizzazione del modulo di ragionamento non monotòno questa importante fase di valutazione non è stata attualmente contemplata. Tuttavia, dato che diversi standard del W3C, inclusi i linguaggi RDF e OWL, sono in grado di fornire interpretazioni semantiche per grafi RDF che consentono di inferire RDF statements addizionali, e dato che, in base alla soluzione proposta, le due componenti informative sono tra loro integrabili mediante tali linguaggi, tale valutazione rappresenta un importante lavoro futuro da portare a compimento per avere un quadro maggiormente esaustivo con cui poter considerare l'efficacia della soluzione proposta.

### **Ringraziamenti**

Il lavoro di ricerca finalizzato alla individuazione della soluzione proposta è stato svolto con Marcello Frixione (che mi preme ringraziare anche per i preziosi suggerimenti forniti per la stesura di questo articolo). Un ringraziamento va anche ai due revisori anonimi di questo lavoro che, con i loro commenti, hanno contribuito a migliorarne il contenuto.

## Riferimenti bibliografici

1. Baader F., Calvanese D., McGuinness D., Nardi D. and Patel-Schneider P. (2003). *The Description Logic Handbook: Theory, Implementations and Applications*. Cambridge University Press.
2. Baader, F., Hollunder, B., 1995. Embedding defaults into terminological knowledge representation formalisms. *J. Autom. Reasoning* **14**(1), 149–180.
3. Bizer, C., Heath, T., Berners-Lee, T., (2009). Linked Data - The Story So Far, *International Journal on Semantic Web and Information System*, 5 (3), pp. 1-22.
4. Bobillo, F., Straccia, U., (2009). An OWL Ontology for Fuzzy OWL 2. *Proc. 18th International Symposium on Methodologies for Intelligent Systems (ISMIS-09), Lecture Notes in Computer Science*, Springer Verlag, Berlin-New York.
5. Brachman, R. and Schmolze, J.G. (1985). *An overview of the KL-ONE knowledge representation system*, *Cognitive Science*, 9, 171-216.
6. Calegari S., Ciucci. D. (2007), “*Fuzzy Ontology, Fuzzy Description Logics and Fuzzy OWL*”, Proceedings of WILF 2007, Vol 4578, LNCS.
7. Corcho O. , Gómez-Pérez A., (2000), A Roadmap to Ontology Specification Languages, in *Proceedings of EKAW00. 12th International Conference on Knowledge Engineering and Knowledge Management*.
8. Cordeschi R., Frixione M. (2011) Rappresentare i concetti: filosofia, psicologia e modelli computazionali, *Sistemi Intelligenti*, 1, 25-40.
9. D’Amato C., Nicola Fanizzi N., and Esposito (2008) “*Analogical Reasoning in Description Logics*“ in F. P.C.G da Costa et al. (Eds.): URSW 2005-2007, LNAI 5327, pp. 330–347, 2008. Springer-Verlag Berlin.
10. Da Costa P.C.G., Laskey, K.B., 2006. PR-OWL: A ramework for probabilistic ontologies. *Proc. FOIS-2006*, 237–249.
11. Ding, Z., Peng, Y., Pan, R., (2006). BayesOWL: Uncertainty modeling in Semantic Web ontologies. In Z. Ma (ed.), *Soft Computing in Ontologies and Semantic Web*, vol. 204 of *Studies in Fuzziness and Soft Computing*, Springer.
12. Evans, J. St.B.T. and Frankish, K. (eds.) (2008). *In Two Minds: Dual Processes and Beyond*. New York, NY: Oxford UP.
13. Fanizzi N., (2011), “*Concept Induction in Description Logics Using Information-Theoretic Heuristics*” in *International Journal on Semantic Web and Information Systems*, 7(2), 23-44.
14. Fodor, J., (1987), *Psychosemantics*, Cambridge, MA: The MIT Press.

15. Frixione, M., Lieto, A. (2010). The computational representation of concepts in formal ontologies: Some general considerations, *Proceedings of KEOD 2010*, Valencia, Spain, October 25-28th 2010.
16. Frixione, M., Lieto, A. (2011). "Representing Concepts in Artificial Systems: A Clash of Requirements", in Proceedings of HCP 2011, Genoa, Italy.
17. Frixione, M., Lieto, A. (2012a). Representing Non Classical Concepts in Formal Ontologies: Prototypes and Exemplars, in *New Challenges in Distributed Information Filtering and Retrieval*, Studies in Computational Intelligence. Volume 439. Springer, 2012.
18. Frixione, M., Lieto, A. (2012b). "Prototypes vs Exemplars in Concept Representations", in Proceedings of KEOD 2012, Barcelona, Spain.
19. Gagliardi, F. (2008). "A Prototype-Exemplars Hybrid Cognitive Model of "Phenomenon of Typicality" in *Categorization*, in *Proc. 30th Annual Conf. of the Cognitive Science Society, Austin, TX, 1176-1181*.
20. Gagliardi F. (2011), Instance-based classifiers applied to medical databases: Diagnosis and knowledge extraction, *Artificial Intelligence in Medicine*. 52(3):123-139. doi:10.1016/j.artmed.2011.04.002
21. Horrocks I., Patel-Schneider P. F., and Van Harmelen F., (2003), "From SHIQ and RDF to OWL: The Making of a Web Ontology Language. *J. of Web Semantics*, 1(1):7-26, 2003.
22. Lukasiewicz, L., Straccia, U., 2008. Managing uncertainty and vagueness in description logics for the Semantic Web. *Journal of Web Semantics*, 6, 291-308.
23. Machery, E., 2009. *Doing without Concepts*. Oxford, UK: Oxford University Press.
24. Medin, D.L., Schwanenflugel, P.J., 1981. Linear separability in classification learning. *J. of Exp. Psyc.: Human Learning and Memory*, 7, 355-368.
25. Minsky, M. (1975). A Framework for Representing Knowledge. *The Psychology of Computer Vision*, P. H. Winston (ed.), McGraw-Hill.
26. Motik B., Horrocks I., Riccardo Rosati R., Ulrike Sattler U., Can owl and logic programming live together happily ever after?, In Proc. of the 5th Int. Semantic Web Conf. (ISWC 2006).
27. Murphy, G.L., 2002. *The Big Book of Concepts*. Cambridge, MIT Press.
28. Osherson, D.N., Smith, E.E., 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition*, 11, 237-262.
29. Quillian, M. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing*, 227-270. Cambridge, MIT Press.
30. Rosh, E. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology* 104: 573-605.

31. S. Schlobach, M.C.A. Klein, L. Peelen, (2007), Description logics with approximate definitions: precise modeling of vague concepts, in: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007), pp. 557–562.
32. Sternberg, R.J. (2003) Cognitive psychology. Third edition. Thomson Wadsworth. Belmont CA.
33. Witten, I.H., Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. (2nd Ed.) Kaufmann, San Francisco, CA.
34. Wittgenstein, L., (1953). *Philosophische Untersuchungen*. Oxford, Blackwell.